Environ Res Tec, Vol. 8, Issue. 2, pp. 447-455, June 2025

Environmental Research and Technology
https://dergipark.org.tr/en/pub/ert
DOI: https://10.35208/ert.1520902

Environmental Research & Technology

Research Article

Exploring the relationship between rainfall and crop yield and best practices adoption using participatory approach

Arunya KALAIYARASI GANESAPANDIAN*[1] Krishnaveni MUTHIAH[2]

[1]*Centre for Water Resources, Anna University, Guindy, Chennai, India*
[2]*Institute for Ocean Management, Anna University, Guindy, Chennai, India*

ARTICLE INFO

Article history
Received: 31 Jul 2024
Accepted: 05 Oct 2024

Key words:
Precipitation, ground truthing, participatory approach, İmplementation of results.

ABSTRACT

Crop yield is a standard measurement for the amount of agricultural production. Sustainable agriculture demands an increase in crop yield. This study deals with rainfed agriculture; hence, precipitation becomes the driving factor for crop yield. Heat maps are used to examine the rainfall and crop yield correlations. ML is an essential tool in decision-making, and many ML algorithms are available for prediction. This study uses the ML algorithms to predict whether the crop yield will increase with increased rainfall. Logistic regression, Decision tree classifier, Random Forest classifier, and XGBoost classifier are the algorithms chosen for analysis. Altogether this region consists of forty crops but focuses on five predominant annual crops. Implementation-based results are the universal goal of every research which society needs. The chances of implementation are associated with two major components: the reliability of the results and society's willingness. Analysis of these components needs ground truthing and Participatory Rural Appraisal, respectively. Farmers and villagers filled out a questionnaire about the details required for this study. The survey was an active approach to collecting necessary information from the participants. The survey showed positive results among one hundred and fifty samples from six blocks. Finally, cashew nut, sugarcane, and turmeric showed good dependency on the precipitation, and around 88% of villagers are ready to implement the results derived from ML algorithms.

Cite this article as: Arunya KG, Krishnaveni M. Exploring the relationship between rainfall and crop yield and best practices adoption using participatory approach. Environ Res Tec 2025;8(2) 447-455.

INTRODUCTION

Many models have evolved and validated for crop yield predictions, which is one of the significant issues in agriculture, as [1] found. Though crop yield depends on many factors [2], such as climate, seed variety, irrigation, soil fertility, and fertilizer, this study tests the dependency of crop yield on rainfall. Many statistical methods are available for yield loss using all the above parameters [3]. Rainfall is a predominant factor that affects agricultural production [4]. Since rainfall is a complex atmospheric process, it takes work to predict. Difficulty in predicting rainfall affects people in many different ways, such as tourism, pilgrimage, and many other factors [5], including agriculture.

Yield prediction has greater significance in terms of yield measuring, yield evaluation, and coordinating harvest supply. [6]. Many studies have analyzed various techniques used for crop yield prediction and concluded that the data mining techniques are the best [7]. Computational models such as deep learning can predict by adequately adjusting the input parameters. However, sometimes, the output may fit the data well during the training but will provide poor results due to overfitting or overtraining [8]. Machine Learning algorithms are less computationally expensive than these computational models. Researchers used algorithms such as Neural Net-

***Corresponding author.**
*E-mail address: arunya0517@gmail.com

works, Linear Regression, Support Vector Machines, Random Forest, and Gradient Boosting Trees to predict crop yields [1].

Participation and empowerment are new words in the field [9]. To achieve sustainable adaptation, we need to enhance the site-specific knowledge of the locals. The traditional research method focuses little on community capacity building; however, disseminating predicted results is more critical [10]. It is essential to increase sensitivity in the minds of the locals towards their responsibility to society's social and physical conditions. Participatory Rural Appraisal (PRA) makes the local's joint action increase awareness and responsibility [11]. PRA's main aim is to develop context-based solutions with the combined effort of professionals, universities, state agency officials, and local communities [12]. Hence using PRA, this study analyses the possibilities for implementation with the help of the locals.

MATERIALS AND METHODS

Study Area

Paravanar basin is one of the minor River basins in Neyveli, Cuddalore District of Tamil Nadu. The latitude and longitudinal extent of Paravanar River Basin are 11°18" to 11°45" North latitude and 79°18" to 79°45" East longitude which is shown in figure 1. The Penniyar River basin bounds the Paravanar River basin in the North, the Vellar River basin in the South and West, and the Bay of Bengal in the East.

The word "Paravanar" in Tamil means spreading River. Numerous small streams are spreading over this basin area, forming the River Paravanar. The sedimentary rock type covers the Paravanar River Basin, occurring from tertiary to recent age. The Basin receives a high amount of rainfall during North – East Monsoon. In this basin, district, taluk headquarters, and major towns are well connected with Road and Rail networks. The average annual rainfall of this basin is 1197.70mm.

Two sub-basins cover the Paravanar River basin. The upper part of the Paravanar River is called the Paravanar sub-basin, and the lower part of the Paravanar River is called the Uppanar sub-basin. The areas covered by these two sub-basins are Paravanar Sub Basin (435.016 Sq. Km) and Uppanar Basin (437.325 Sq. Km), which covers a total area of 872.34Sq.Km.

The "Neyveli Lignite Corporation" is located within this basin. Mine 1, Mine 1A & Mine 2 are the three opencast mines operated in this area. The Perumal Tank utilizes the groundwater pumped from the mines via the Walajah tank for irrigation.

Test taken on the command area of the Perumal Tank indicates that the water used for irrigation from both Perumal Tank and Walajah Tank has no adverse effect on the soils to hinder crop growth.

The topographic trend is falling towards the south and southeast direction. Paravanar is a plain and upland terrain basin devoid of hills and related morphologies. Geologically the basin is covered with sedimentary formation. About 70% of the basin area covered Cuddalore sandstone of the Tertiary Age, which consists of laterite, sandstone, clay, silt, and sand. River alluvium and coastal alluvium of recent age cover 30% of the basin area.

The Paravanar River has yet to fully develop and is only present during certain seasons. It comes from various streams in the highlands northwest of the Neyveli lignite corporation area, specifically from the Semmakottai Reserve forest near Raghavankuppam and Kovilankuppam villages, at an elevation of approximately 100 meters above MSL. The River Paravanar is joined by several small streams in its upper reaches, which flow eastward and eventually drain into the Walajah Tank. The Sengal Odai and the Kanniyakovil odai are two streams originating in Mulaikuppam village and Southeast Neyveli Township, respectively. These two streams merge in Ellaikudi village, where the River Middle Paravanar begins and flows into the Perumal tank. Some streams that flow eastward from the Walajah tank eventually reach the Satapadi tank, and the excess water from this tank forms the

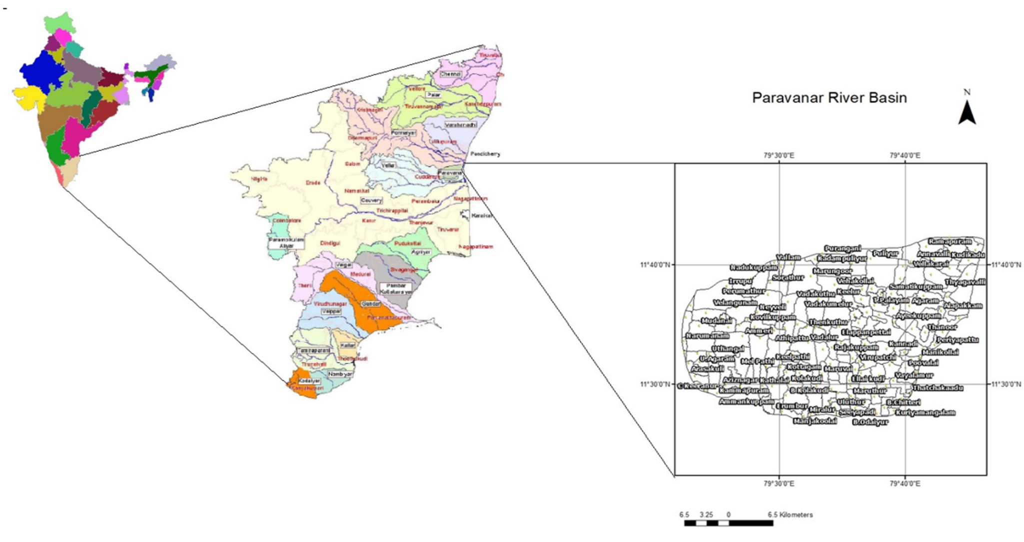

Figure 1. Paravanar River Basin

Uppanar River. As it flows northeast, the Uppanar River is fed by excess water from the Perumal tank near Periyapakkam, Anniyampettai, and Pundiyankuppam. Eventually, the river merges with the Bay of Bengal to the north of Semabadakuppam and south of Cuddalore O.T. The Uppanar River spans roughly 24 kilometers and has 37 tanks within its basin, overseen by WRD/PWD, with 16 of these tanks having an ayacut of over 100 hectares. The main tanks responsible for receiving and distributing water for irrigation are the Perumal Tank and Walajah Tanks.

The crops chosen for this study are the crops that are cultivated throughout the year and depend majorly on rainfall. Hence the yearly rainfall chose for this analysis. They follow crop rotation also, which depends on the rain and the crop harvested during the previous year. This basin cultivates the following crops which are areca nut, arhar, bajra, banana, black pepper, cashew nut, castor seeds, coconut, coriander, cotton, cowpea, dry chilly, garlic, gram, ground nut, guar seed, horse gram, jowar, maize, masoor, moong, onion, potato, ragi, mustard, rice, sannhamp, sesamum, minor millets, sugarcane, sunflower, sweet potato, tapioca, tobacco, turmeric, urad and other cereals, kharif, and rabi pulses. Farmers cultivate these crops in six seasons: autumn, kharif, rabi, summer, winter, and yearly. The data is collected from 1998 to 2020. As the data quantity increases, ML accuracy also increases, so it is advisable to use a considerable amount of data.

Data Collection

The data used for rainfall is the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Climate Data Record (PERSIANN CDR) from National Centres for Environmental Information – National Oceanic and Atmospheric Administration for the location 11°18" to 11°45" North latitude and 79°18" to 79°45" East longitude. The PERSIANN algorithm on Gridsat - B1 Infrared satellite data produces the data from 1983 with 0.25° resolution, including data from most geostationary satellites. Validation of satellite rainfall data is mandatory for any analysis [13]. Therefore, the data is validated using the rain gauge station data, and the correlation is 0.76.

India's Agriculture and farmers welfare ministry contains the agriculture crop production data. The data consists of 1998-present data with more than 40 crops cultivated in that region. The primary five annual crops are the main focus of this study. They are Cashew nut, Coriander, Sugarcane, Sweet Potato, and Turmeric. Cashewnut, Sugarcane and Turmeric are annual crops whereas Coriander and Sweet Potato are seasonal crops (June to September). Hence the correlation study for the annual crops are analysed with the annual rainfall and the seasonal crops are analysed with the southwest monsoon rainfall. The data validation uses ground truthing with the farmers through a questionnaire survey conducted during Rapid Rural Appraisal.

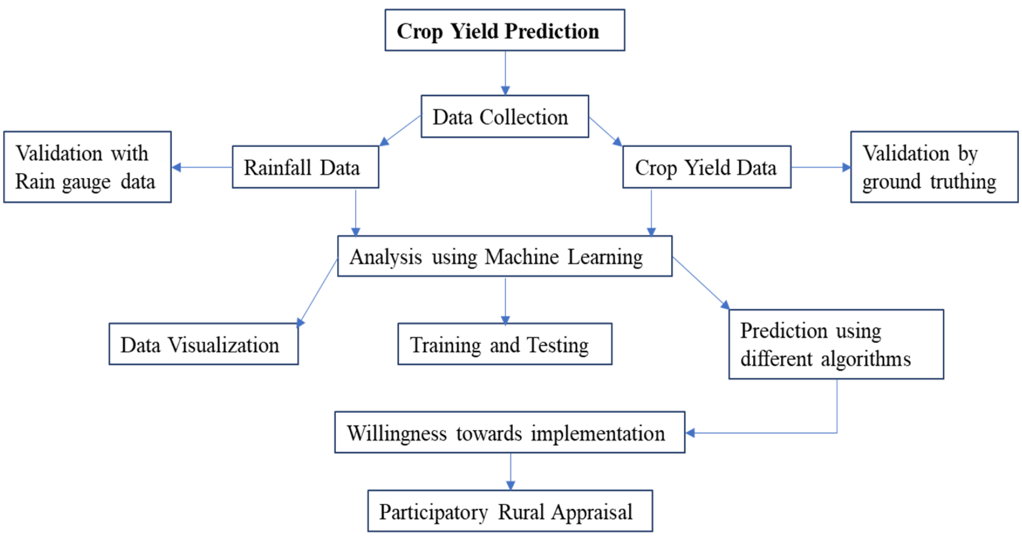

Figure 2. Methodology Flow Chart

ML Analysis

The overall methodology is explained in a flow chart denoted as figure 2. The field of machine learning falls under the umbrella of artificial intelligence and involves training machines to learn independently without external assistance. Through machine learning, machines can analyze past data to improve their future decision-making processes. This practice uses various tools, techniques, and algorithms to achieve better results and explore new knowledge from existing data sets. The choice of machine learning algorithms depends on the type of prediction needed, whether it is an estimate or classification [4]. Supervised learning is often chosen for prediction as it provides labels for data division and desired output, offering insights into fields such as health and education [14]. To learn, the algorithm compares its work with the trained output, identifies mistakes, and makes necessary adjustments. The present research employs multiple classifiers such as Logistic Regression, Decision Tree Classifier, Random Forest Classifier, and XGBoost Classifier to examine whether increased rainfall positively affects crop yield.

Logistic regression

Logistic regression is a frequently used ML technique that follows the supervised learning technique. It predicts the categorical dependent variable from the independent variables. It is similar to linear regression, but it fits the data with an S-curve which predicts two extreme values. Logistic regression is a good predictor with low incidences and simple predictors [15]. Hence it is chosen for our prediction.

Decision tree classifier

The decision tree is also a supervised technique similar to logistic regression. A decision tree predicts basic tests efficiently and works under the principle of Boolean logic [16]. Each tree comprises nodes and branches. This study concentrates on only a fundamental analysis of this classifier chosen for prediction.

Random forest classifier

Random forest classifier comprises many simple decision tree classifiers. [17] Researchers found that this classifier is an ensemble learning technique, and the individual trees follow the same technique as the decision tree in the ensemble. According to the pseudo-code, every tree in the forest contributes to a voting system to determine the classification, with the tree that receives the most votes being chosenjij overcomes the problem of overfitting data in the decision tree classifier. It takes the average of various subsets of the decision tree, which improves the accuracy of random forest classifier prediction.

XGBoost classifier

It is also called the eXtreme Gradient Boosting algorithm, created to decrease the time of model prediction for large datasets. Researchers developed this for highly accurate predicting [18]. Researchers in this study found that this classifier dominates the structured datasets on predictive modeling [19].

Village Cluster based Participatory Rural Appraisal (PRA)

Understanding the crop species is vital for its improvement [20]. Collaborative analysis can identify the needs and strengths of the local people and provide a wide range of information [21]. The study will be explained to the local farmers in detail before the survey, and then the questionnaire will be given to the farmers. We divided the study area into 6 clusters and the farmer groups into small and large-scale farmers from all six blocks of the study basin. The main aim is to make the locals to adopt the best-practice, The best-practice in the study make the locals represent and analyze the study results about their crop production and make their plans accordingly. Also, ground truthing helps in validating the predicted results. Ground truthing means validating the predicted results by doing a questionnaire survey from the local farmers. Also, the farmers helped to identify their willingness to implement the predicted results through a questionnaire survey.

RESULTS AND DISCUSSION

The results were showed as maps and tables. ML predicts the crop yield for the selected crops with the accuracy scores and are shown in table 1. The heat maps were generated for all the five crops to test the dependency of different factors which were shown in figures 3-7.

And the preliminary survey details from the participatory approach with the gender, age group and the type of farmer is showed in table 2. Ground truthing results and the willingness for implementation from the survey has been displayed in table 3 and figure 8 respectively.

Table 1. Accuracy scores for different algorithms

S.No	Crop	ML Algorithms	Accuracy Scores
1	Cashewnut	Logistic Regression	0.82
		Decision Tree	0.91
		Random Forest	0.91
		XGBoost Classifier	0.86
2	Coriander	Logistic Regression	0.71
		Decision Tree	1.0
		Random Forest	0.9
		XGBoost Classifier	0.66
3	Sugarcane	Logistic Regression	0.88
		Decision Tree	0.85
		Random Forest	0.85
		XGBoost Classifier	0.88
4	Sweet Potato	Logistic Regression	1.0
		Decision Tree	0.62
		Random Forest	0.81
		XGBoost Classifier	0.9
5	Turmeric	Logistic Regression	0.91
		Decision Tree	0.91
		Random Forest	0.91
		XGBoost Classifier	0.66

Table 2. Primary details from a questionnaire survey

S.No	Description	Percentage
1	Gender	
	Male	58
	Female	42
2	Age group	
	21-45	28
	45-60	57
	>60	15
3	Type of Farmers	
	Small Scale (<5 acres)	32.5
	Large Scale (>5acres)	67.5

Table 3. Ground truthing on the dependency of crop yield over rainfall

S.No	Crops	Results
1.	Cashewnut	Increase
2.	Coriander	Increase
3.	Sugarcane	No notable changes
4.	Sweet Potato	Increase
5.	Turmeric	Increase

ML Predictions

ML models can make two kinds of decisions which can be descriptive and predictive. Descriptive models help to gain knowledge about the data, whereas predictive models help to make predictions from the data [22]. Descriptive and predictive analysis are both essential for this study. The descriptive analysis uses heat maps to learn about the data and its dependency on rainfall. The following Figures represent the heat maps for all the crops, in the following Figures. Four algorithms, such as Logistic Regression, Decision Tree, Random Forest, and XGBoost, help in predictive analysis. This study identifies which of the four algorithms can accurately predict this region and data. Hence, this study chose the accuracy scores as the evaluating factor.

A heat map is a graphical representation of data in a two-dimensional color matrix with columns and rows. It is used to visualize the strength of correlation among the variables. In this study, the heat map contains the factors of the year of rainfall and yield, area of crop cultivated, production in tonnes, crop yield, rainfall in millimeters (mm), and rainfall in meters (m). It has a scale that ranges from -1 to +1. This scale represents the values of correlation coefficients. The major concentration is identifying the correlation between crop yield and rainfall.

Figure 3 shows the heat map of cashew nuts. The cashew nut shows good dependency on rainfall with yield and also, and it states that the area is also a contributing factor to the yield. The heat map contains a scale with a range of -1 to +1. The value falls between 0.4 to 0.6 for the dependency of rainfall over the yield.

The coriander shows less dependency on all the terms in the heat map than the cashew nut except the rainfall with yield, as shown in Figure 4. The value of dependency falls somewhere between 0.25 to 0.5 on the scale.

Also, from Figure 5, the heat map of sugarcane needs to show a better dependency on all of the factors. It is evident from the scale, too, that yield dependency over rainfall is -0.25 from the heat map.

Figure 6 shows that the Sweet potato also show good relationships with the factors in the heatmap, and the value from the heat map is 0.4 to 0.6 for the dependency of yield over rainfall.

Inference from Figure 7 shows a good relationship between turmeric yield and rainfall with a value that falls between 0.4 to 0.6 from a scale of -1 to +1. Hence, the descriptive modeling shows that cashew nuts and turmeric yield depends on rainfall.

Four ML algorithms show their accuracy scores for all five crops. The algorithm with the highest accuracy score best predicts crop yield with increased rainfall. The bifurcation of the data is as follows: 80% of the data is for training, and the remaining 20% is for testing. The algorithm predicts whether the yield will increase if rainfall increases. Since this study has yet to accurately predict future rainfall, we have identified the best algorithm for it. We will estimate the amount of rainfall in the future by conducting further studies.

The Cashewnut accuracy scores are 0.82, 0.91, 0.91, and 0.86 for logistic regression, decision tree classifier, random forest classifier, and the XGBoost classifier, respectively. Hence it is identified that the production of cashew nuts in the study region depends on the rainfall.

The accuracy scores for coriander are 0.71, 1.0, 0.9, and 0.66, respectively. The accuracy score of 1 denotes an error value due to the low amount of dataset for training and testing. Hence, the coriander production depend on the annual rainfall in the study region.

The accuracy scores for sugarcane are 0.88, 0.85, 0.85, and

0.88, respectively. However, since the current yield does not correlate with the rainfall, increasing the dataset record length needs further analysis.

The accuracy scores are estimated as 1.0, 0.62, 0.81, and 0.9, respectively, for sweet potatoes. The random forest and XG Boost predict pretty accurately, and the heat map shows that the sweet potato yield directly depend on rainfall.

Moreover, the accuracy scores are 0.91 for the first three algorithms and 0.66 for the XGBoost classifier in turmeric. Therefore, logistic regression, decision tree, and random forest classifiers can predict the turmeric yield for this study region. The heat map estimates the yield of turmeric depending on the annual rainfall.

Village Cluster based PRA for Best-Practices Adoption

Even with crop yield prediction, the implementation could only be possible with the help of the end users or the farmers. Here comes the role of Participatory Rural Appraisal (PRA). The PRA method asses the willingness for best-practices adoption. The assessment includes a site visit, direct farmer interaction, and a questionnaire survey. Among the interactions made with more than 150 farmers, this study identifies that around 85% are eagerly willing to adopt if the prediction accuracy is about 85-90%. Successful implementation of these predicted results will increase crop yield, thereby improving the chances of other farmers for adopting the best-practices.

The PRA employs a method of sampling known as cluster sampling or 'two-stage sampling.' The first stage involves selecting a sample of geographical areas, while the second involves selecting a sample of respondents within those areas. This method divides the population into clusters of similar units based on geographical proximity. Totally six blocks in this study region with more than 50 villages, including the mine sites, are covered during this study. They blocks are Cuddalore, Panruti, Kurinjipadi, Kammapettai, Mel Bhuvanagiri and Parangipettai. More than 50% of farmers are male, with an age group of 45-60. This region contains many large-scale farmers compared to small-scale farmers.

Table 3 explains the ground truthing of the predicted results; it matches precisely with cashew nut, coriander, and turmeric. The main question they asked for ground truthing is, "Whether the yield of the particular crop will increase if there is an increase in rainfall?". Cashewnut yield is increasing, whereas coriander yield has no notable change, even if there is an increase in rainfall. Moreover, they said the sugarcane yield would also increase, but it does not match the ML prediction. Sweet potato yield will increase with rainfall during the early cropping period. Furthermore, the ML prediction also does not show the dependency of Sweet Potato yield over rainfall. Finally, the turmeric yield also increases with an increase in rainfall.

A questionnaire survey helps this study to calculate their willingness towards implementation based on the predicted results by the ML by explaining the findings to them. We identified that 78% could readily accept and cultivate crops based on predictions. 18% of people are willing to accept only after successful implementation results, and 4% are unwilling to implement based on rainfall dependency since they are unsure about their cropping pattern. The following pie chart shows the results in Figure 8.

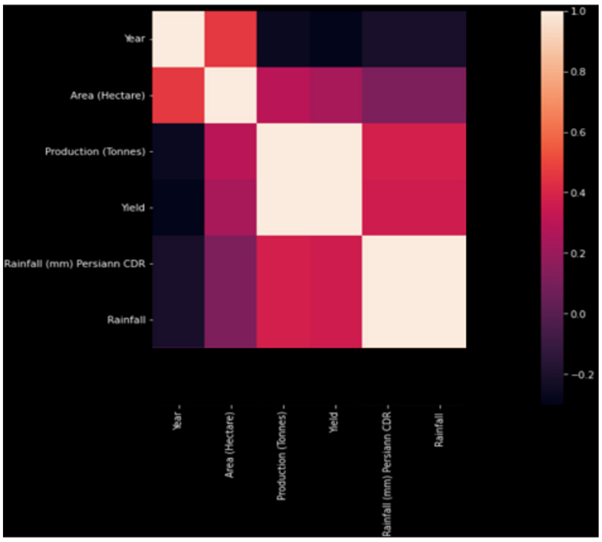

Figure 3. Heat map of Cashew Nut

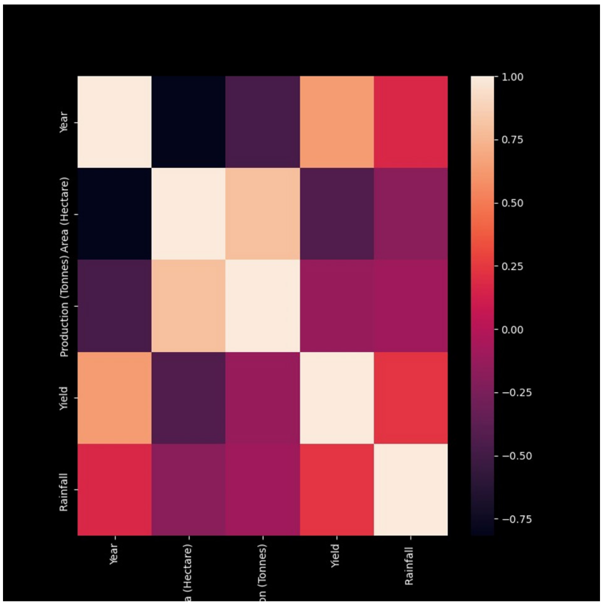

Figure 4. Heat map of Coriander

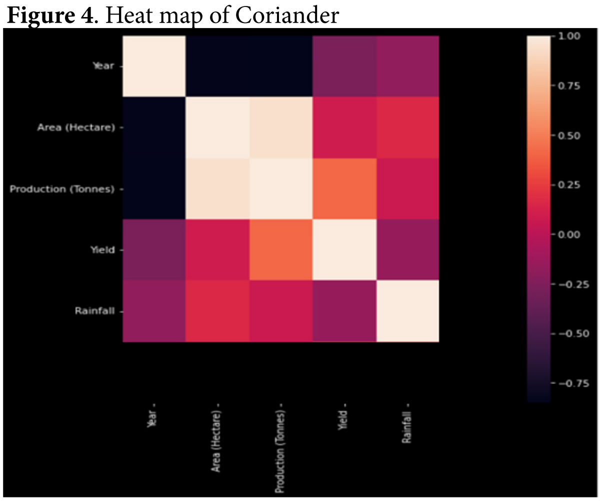

Figure 5. Heat map of Sugarcane

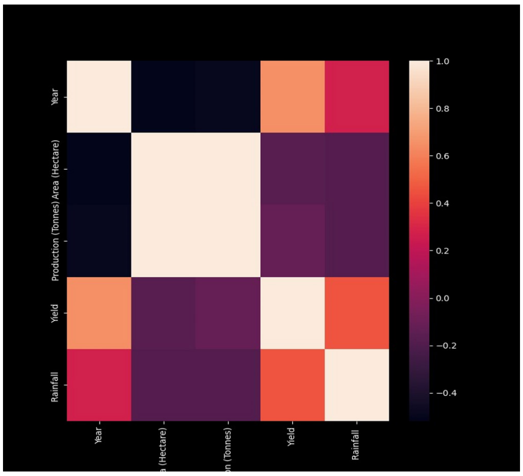

Figure 6. Heat map of Sweet Potato

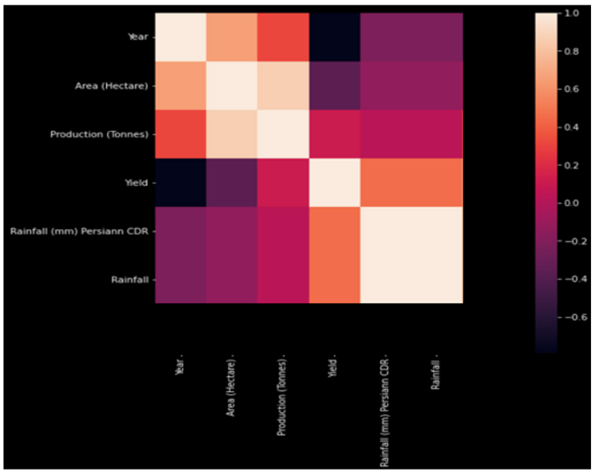

Figure 7. Heat map of Turmeric

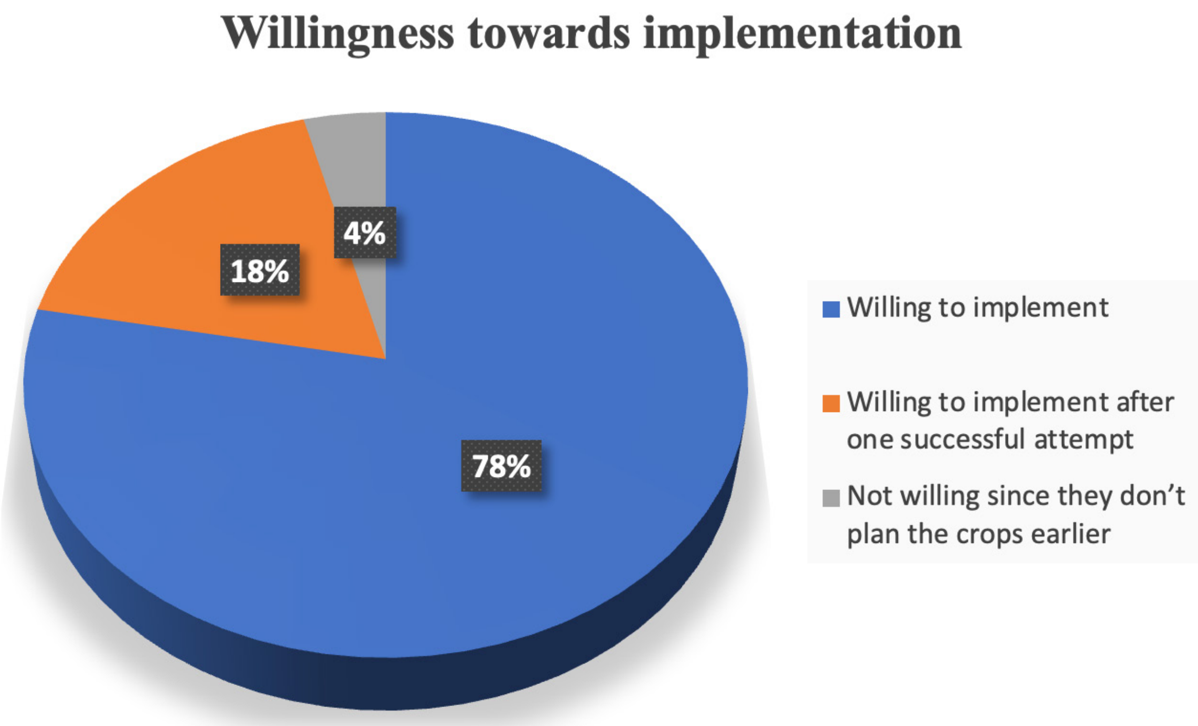

Figure 8. Willingness to implementation of results

CONCLUSION

On the whole, this study tests the dependency of crop yield in the Paravanar River basin on the annual rainfall using logistic regression, decision tree, random forest, and XG-Boost classifiers of ML algorithms. The results are validated using field observations called the participatory rural appraisal technique, and also the results are correlated well. Hence the ML predictions are found to be similar to the ground results. The drawback felt in this study is the data record length; still, these predictions require vast datasets. That will increase the accuracy of the predictions. Hence, using more than 50 years of data for these predictions is recommended. ML predictions based on the heat maps and accuracy scores are best interpreted and identified. Ground truthing with the farmers enhances the reliability of this study with the help of a questionnaire survey, and all the ML predictions match the field data, except for sugarcane and sweet potato. Sugarcane has an increase in yield if there is an increase in rainfall. The above statement is a fact inferred from the farmer's survey, but in ML, it does not show any increase in yield. Increasing the length of the data can rectify the error. Sweet potato shows no increase in ML prediction, whereas, from the survey, this study identifies that the yield will increase if rainfall increases during the sowing season. Otherwise, the rainfall does not impact the yield of sweet potatoes. Hence, crops such as cashew nut, sugarcane, and turmeric depend on annual rainfall in this region.

Moreover, the best algorithm for cashew nuts is a decision tree and random forest with an accuracy score of 0.91. The sugarcane prediction's best algorithm is logistic regression and XGBoost, with an accuracy score of 0.88. The best algorithm for turmeric prediction is logistic regression, decision tree, and random forest, with an accuracy score of 0.91. The participatory rural appraisal identifies whether the farmers and the villagers are eager to implement the results of ML predictions. This analysis shows positive results, with 88% of the people ready to cultivate crops based on this prediction. Further extension of this study involves predicting the rainfall so that the results can help in decision-making and successful implementation based on the dependency of predicted rainfall.

Data Availability Statement

It is stated that the published publication includes all data collected or developed during the study. The data sources are mentioned for the secondary data used in this study.

ACKNOWLEDGEMENT

Anna University, Department of Science and Technology, Chennai, India.

DATA AVAILABILITY STATEMENT

The author confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

USE OF AI FOR WRITING ASSISTANCE

Not declared.

ETHICS

There are no ethical issues with the publication of this manuscript.

REFERENCES

[1] Klompenburga TV, Kassahuna A, and Catalb C, "Crop yield prediction using machine learning: A systematic literature review," Computers and Electronics in Agriculture, vol. 177, 2020.

[2] Xu X et al., "Design of an integrated climatic assessment indicator (ICAI) for wheat production: a case study in Jiangsu Province, China," Ecological Indicators, vol. 101, pp. 943–953, 2019.

[3] Pradyumn Kumar et al., "Assessment of crop loss caused by Chilo partellus in maize," Indian Journal of Agricultural Sciences, vol. 91, no. 2, pp. 218–221, 2023.

[4] Steve Oberlin, "Machine Learning, Cognition and Big Data," CA Technology Exchange, United States., 2012.

[5] Niveditha Nath, "From Pilgrim Landscape to Pilgrim Road: Tracing the transformation of the Char Dham Yatra in Colonial Garhwal," Journal for the study of religion, nature and culture, pp. 419–437, 2018.

[6] Ramos PJ, Prieto FA, Montoya EC, and Oliveros CE, "Automatic fruit count on coffee branches using computer vision," Computer Electronics and Agriculture, vol. 137, pp. 9–22, 2017.

[7] Beulah R, "A survey on different data mining techniques for crop yield prediction," International Journal of Computer Science Engineering, vol. 7, no. 1, pp. 738–744, 2019.

[8] Weigend A, "An overfitting and the effective number of hidden units. Lawrence Erlbaum Associates," Hillsdale, pp. 335–342, 1993.

[9] Chandra G, "PARTICIPATORY RURAL APPRAISAL Issues and Tools for Social Science Research in Inland Fisheries," Central Inland Fisheries Research Institute. Bulletin, vol. 163, pp. 286–302, 2010.

[10] Lilian A Omondi, "Learning together: Participatory rural appraisal for coproduction of climate change knowledge," Action Research, vol. 21, no. 2, 2020.

[11] Rudi Saprudin Darwis, Risna Resnawaty, and Eva Nuriyah, "Increasing the Sensitivity of Local Leadership in Citarum River Management through Participatory Rural Appraisal (PRA) techniques in Rancamanyar Village," Kumawula Journal of Community Service, vol. 3, no. 1, pp. 48–59, 2020.

[12] Pankaj Kumar, Dheeraj Kumar, Sachin Kumar, Jitendra Kumar, Kiran Pal, and Nikhil Jadhav, "Historical Perspective of Watershed Management in India: A Participatory Rural Appraisal (PRA) based Assessment," Asian Journal of Agricultural Extension, Economics & Sociology, vol. 40, no. 10, pp. 406–418, 2022.

[13] Alejo LA and Alejandro AS, "Validating CHIRPS ability to estimate rainfall amount and detect rainfall occurrences in the Philippines.," Theoritical Applications of Climatology, vol. 145, pp. 967–997, 2021.

[14] Jainendra Singh, "Big Data Analytic abd Mining with Machine Learning Algorithm," International Journal of Information and Computation Technology, vol. 4, no. 1, pp. 33–40, 2014.

[15] Nusinovici S et al., "Logistic regression was as good as machine learning for predicting major chronic diseases," Journal of Clinical Epidemiology, vol. 122, pp. 56–69, 2020.

[16] Jijo BT and Abdulazeez AM, "Classification Based on Decision Tree Algorithm for Machine Learning," Journal of Applied Science and Technology Trends, vol. 2, no. 1, pp. 20–28, 2021.

[17] Peter DC and Arandejelovic O, "Precision medicine

in digital pathology via image analysis and machine learning," Artificial Intelligence and Deep Learning in Pathology, pp. 149–173, 2021.
[18] Jerome H Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," Annals of statistics, pp. 1189–1232, 2001.
[19] Ramraj S, Nishant U, Sunil R, and Shatadeep B, "Experimenting XGBoost algorithm for prediction and classification of different datasets," International journal of control theory and applications, vol. 9, pp. 615–662, 2016.
[20] Muralidhara MB et al., "Survey, collection and characterization of Indian Avocado (Persea Americana) germplasm for morphological characters," Indian Journal of Agricultural Sciences, vol. 93, no. 2, pp. 139–144, 2023.
[21] Katiha P, Vass KK, Sharma AP, and Bhaumik U, "Issues and Tools for Social Science Research in Inland Fisheries," Central Inland Fisheries Research Institute. Bulletin, vol. 163, 2010.
[22] Alpaydin E, Introduction to Machine Learning, 2nd Edition. MIT Press, Cambridge, MA, 2010.