J Krishi Vigyan 2024, 12(4) : 956-964
DOI : 10.5958/2349-4433.2024.00164.7
ISSN 2319-6432 (Print)
ISSN 2349-6433 (Online)

Weather-Based Rice Yield Prediction in Kerala Using ANN, SMLR and Normal Regression

Lincy Davis P[1], Ajithkumar B[2], Riya K R[3], Arjun Vysakh[4] and Kavya Babu[5]
Department of Agricultural Meteorology, College of Agriculture
Kerala Agricultural University, Thrissur 680 651 (Kerala)

ABSTRACT

Rice is the staple food crop in Kerala which accounts for nearly all of the state's food grain production and is mainly cultivated under rainfed conditions during the *Kharif* season. The critical influence of weather on rice productivity necessitates accurate and timely yield forecasts to aid agricultural planning. This study aimed to develop district-level rice yield prediction models for Kerala by analyzing the effects of essential weather variables: temperature, rainfall, relative humidity, and solar radiation on yield outcomes. Three statistical methods were employed: normal regression, Artificial Neural Networks (ANN) and Stepwise Multiple Linear Regression (SMLR). Among these models ANN demonstrated the highest predictive accuracy, with superior R^2 values across all districts, indicating its robustness in modeling yield variability based on weather parameters. The ANN model has the ability to capture complex, nonlinear relationships among weather variables underscores its reliability as a tool for rice yield forecasting in all districts of Kerala. This enhanced forecasting potential holds substantial value for proactive agricultural planning and decision-making, allowing stakeholders to better manage resources and mitigate climate risks to ensure stable rice production in the state.

Key Words: Rice, Yield prediction, Statistical model, SMLR, ANN.

INTRODUCTION

Rice is the Kerala's staple food cultivated extensively across the state and accounts for nearly all of its food grain production, with a recorded yield of 6.3 lakh tonnes (GOK, 2022). Yoshida *et al* (1981) highlighted rice's sensitivity to climate factors such as temperature, rainfall, and solar radiation, crucial for its productivity in tropical regions. Being a subtropical crop, rice thrives at temperatures between 20 to 40°C but can suffer under extremes. Bhattacharya and Panda (2013) found that rice yield rose with increased rainfall in subtropical areas but excessive rain with strong winds during flowering could reduce yield. Khavse *et al* (2014) found that weather accounts for about 67% of crop yield variation, underscoring the importance of reliable yield forecasts to guide effective agricultural planning.

Forecasting yield before harvest aids policymakers in managing food supplies and preparing for surplus or scarcity (Dutta *et al,* 2001). Yield prediction methods typically involve empirical statistical and crop simulation models. While simulation models are process-based and provide precise results, they require extensive input data, limiting their application over broad spatiotemporal ranges. Statistical models, on the other hand, need less data and can serve as a practical alternative and are generally developed using crop and weather data through simple regression (Lobell and Burke, 2010; Shi *et al,* 2013), and need calibration for reliability (Das, 2018). Rai *et al* (2013) demonstrated that models incorporating multiple weather variables improved rice yield predictions. This study aims to develop and evaluate statistical models, including Normal Regression, Artificial Neural Networking (ANN), and Stepwise Multi-Linear Regression (SMLR), to forecast *Kharif* rice yields in Kerala.

Corresponding Author's Email - lincy.davis@kau.in
1.Associate Professor & Head, 2,3 Assistant Professor, 4,5 PhD students
Dept. of Agrl. Meteorology,College of Agriculture(COA),Kerala Agricultural University(KAU), Vellanikkara(VLK), Thrissur

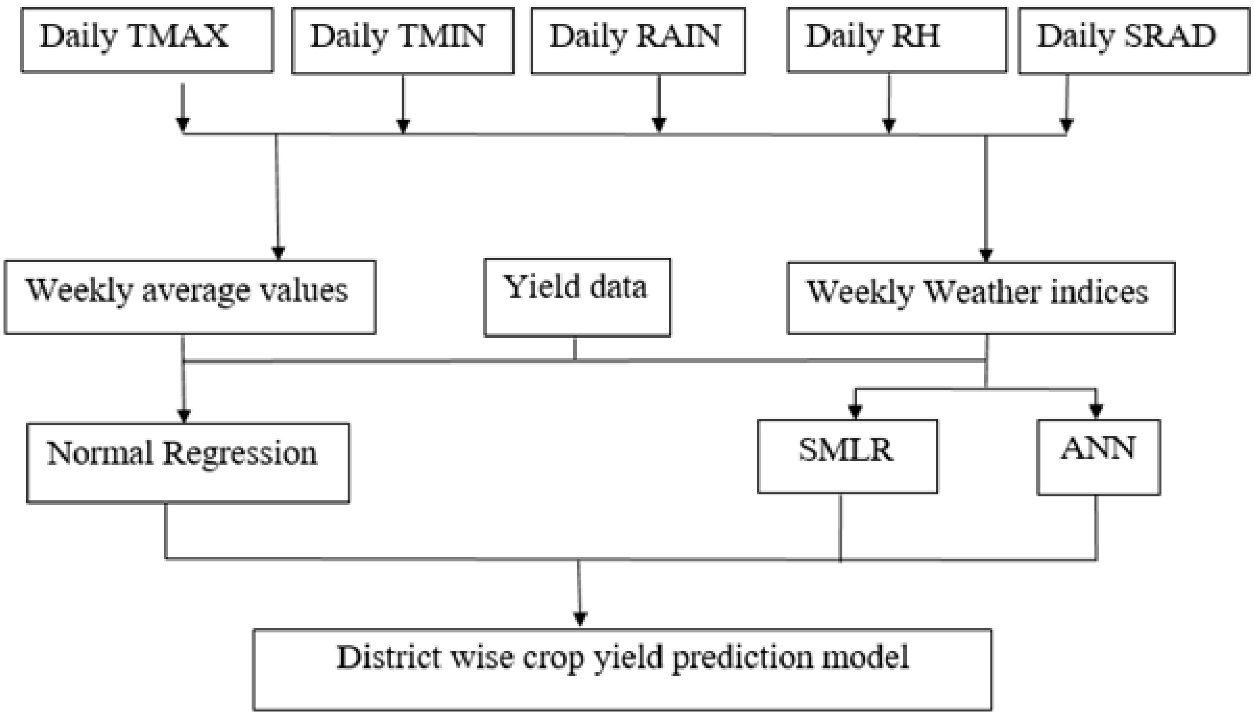

Fig.2. Flow chart representing the development of the statistical model

MATERIALS AND METHODS

Study area

Kerala, situated in southwestern India, lies between 8°18' and 12°48' N latitude and 74°52' and 77°22' E longitude, encompassing a total area of 38,863 square kilometers (Ajithkumar, 2015). The state is divided into 14 districts, with weather data analyzed across each one for this study. Kerala is further classified into five distinct agro-climatic zones: the Northern, Central, High-Range, Problematic, and Southern zones.

Data and methodology

Rice yield data for the *Kharif* season between 1998 and 2021 was gathered from Kerala's Agri Eco-stat website (ecostat.kerala.gov.in) and the Ministry of Agriculture and Farmers Welfare's Area and Production website (aps.dac.gov.in). Data includes all Kerala districts except Wayanad, as no *Kharif* rice cultivation occurs there. The daily weather data from 1998 to 2021 (23 years) was obtained from NASA's Prediction of Worldwide Energy Resources (NASA/POWER; power.larc.nasa.gov) for each district of Kerala, excluding Wayanad. Key weather parameters—maximum temperature, minimum temperature, rainfall, and relative humidity—were recorded daily. These daily data points were then aggregated into weekly values, focusing on the weeks within the Kharif season (20th to 37th Standard Meteorological Week, SMW) to calculate weather indices. Figure 2 illustrates the process flow.

Calculation of weather indices

Two weather indices were developed for each weather variable: the simple index (Z10) and the weighted index (Z11). The simple weather index was calculated by summing individual weekly weather data values (see Equation 1). For the weighted weather index, the sum of the product of each weekly weather value and its correlation with yield was taken (see Equation 2). Table 1 provides an overview of the weather variables and the combinations used in creating these indices.

Table 1. Simple and weighted weather indices.

Parameter	Unweighted weather indice	Weighted weather indices
Tmax	Z10	Z11
Tmin	Z20	Z21
Rain	Z30	Z31
RH	Z40	Z41
SRAD	Z50	Z51
Tmax * Tmin	Z120	Z121
Tmax * Rain	Z130	Z131
Tmax * RH	Z140	Z141
Tmax * SRAD	Z150	Z151
Tmin * Rain	Z230	Z231
Tmin * RH	Z240	Z241
Tmin * SRAD	Z250	Z251
Rain * RH	Z340	Z341
Rain * SRAD	Z350	Z351
RH * SRAD	Z450	Z451

$$Z_{10}=\sum_{k=1}^{n} W_k \quad \text{.......... Equation1}$$

$$Z_{11}=\sum_{k=1}^{n} X_k W_k \quad \text{.......... Equation2}$$

W= Weekly weather

X= Correlation coefficient between yield and weather variables

Statistical Models for yield prediction

Normal Regression

Normal regression was conducted using SPSS (Statistical Package for the Social Sciences), a versatile tool used across various fields for data analysis and transformation (Hinton, 2004). The model equation was formulated using the relationship between yield and weather variables from 1998 to 2021.

Artificial Neural Networking (ANN)

ANN was applied using a three-layer structure comprising input, hidden, and output layers (Fig. 3). Each layer consists of interconnected neurons or nodes, with the number of nodes in the input and output layers depending on the data set. The key challenge in ANN modeling lies in selecting the optimal number of hidden neurons. In this study, the 'train' function from the 'caret' package (method 'nnet') in R Studio (Kuhn, 2008) was used to determine these nodes. Thirty-two variables served as input layers, while yield was the response variable. From a total of 24 observations, 17 were designated as training data (for calibration), and 7 as testing data (for validation or prediction).

Stepwise Multi Linear Regression method (SMLR)

Stepwise Multiple Linear Regression (SMLR) is an iterative method used to construct a regression model by sequentially selecting the most significant weather variables for the final model. As described by Singh *et al* (2014), this method involves adding or removing explanatory variables one at a time, testing for statistical significance after each iteration to identify the optimal set of predictors. In this study, SMLR was conducted using SPSS software, with 20 years (1998-2017) of yield data and 30 corresponding weather indices used for model calibration. The stepwise process yielded multiple models, from which the one with the highest R^2 value was chosen to formulate the equation for yield prediction.

RESULTS AND DISCUSSION

Normal Regression

Using the Normal Regression method, yield prediction models were developed for various districts in Kerala, as displayed in Table 1. The coefficient of determination (R^2) values for these models ranged between 0.143 in Alappuzha

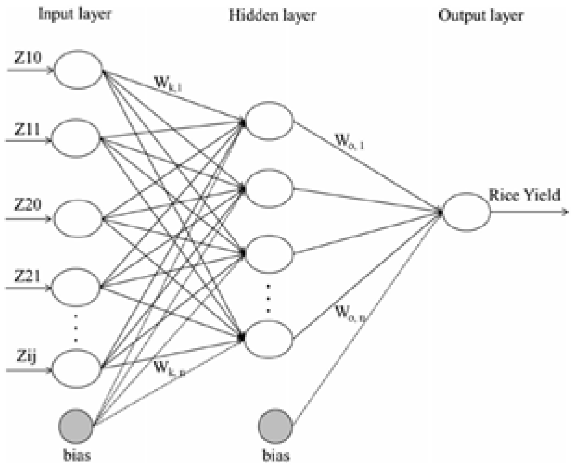

Fig. 3. Schematic representation of ANN

to 0.657 in Idukki. Figure 4 illustrates the combined actual and predicted weather data across all districts.

Stepwise Multiple Linear Regression (SMLR)

Yield prediction models for each district in Kerala, developed using SMLR, are summarized in Table 2. The coefficients of determination (R^2) were significant at the 1% probability level for all districts except Wayanad, with R^2 values ranging from 0.573 in Kozhikode to 0.925 in Thiruvananthapuram. These models identify the key weather variables impacting yield in each district.

In Thiruvananthapuram, yield was influenced by minimum temperature, time series, and the interaction between minimum temperature and solar radiation. Kollam sees yield influenced by time and the interaction of relative humidity with solar radiation. Pathanamthitta yield was shaped by maximum and minimum temperatures and combinations of minimum temperature with relative humidity and rainfall with solar radiation. Alappuzha yield depends on time, relative humidity combined with maximum temperature, and solar radiation.

In Kottayam, influential factors include time, solar radiation, and the interaction of minimum temperature with solar radiation. Idukki yield was driven by maximum temperature and the combination of minimum temperature with relative humidity. Ernakulam was influenced by minimum temperature, time, and the interaction of maximum temperature with solar radiation. Thrissur showed yield dependence on maximum temperature, time, and its relation to relative humidity.

Palakkad yield was affected by time and the interaction of maximum temperature with relative humidity and rainfall. Malappuram yield was highly impacted by maximum temperature, especially in combination with relative humidity and solar radiation. In Kozhikode and Kannur, time and the interaction of rainfall with solar radiation were influential factors. Lastly, Kasargod yield was impacted by minimum temperature, time, and the combination of minimum temperature and solar radiation with relative humidity.

Artificial Neural Networks (ANN)

ANN models were developed to predict crop yield across Kerala districts using weather variables (denoted as Z variables). The models' performance was optimized by selecting between 2 to 9 hidden neurons, and predictive accuracy was assessed using the R^2 (coefficient of determination) and RMSE (root mean square error) (Table 3). The R^2 ranged from 0.784 (Kottayam) to 0.998 (Kannur), while RMSE values ranged from 8.485 (Kannur) to 293.3171 (Alappuzha). The comparison of actual and predicted yield has been shown in Figure 6.

Thiruvananthapuram: Maximum temperature and its combination with relative humidity positively influenced yield. However, minimum temperature, relative humidity, and the interaction between rain and solar radiation had a negative effect.

Kollam: Minimum temperature and the combination of maximum temperature with relative humidity were positive influences, while rain and solar radiation negatively affected yield.

Pathanamthitta: Maximum temperature in combination with minimum temperature and rainfall with solar radiation positively impacted yield, while minimum temperature and combinations of maximum temperature with relative humidity had negative effects.

Alappuzha: Maximum temperature, relative

Table 1. Yield prediction models for different districts of Kerala using Normal regression method.

District	Equation	R^2 value
Thiruvananthapuram	-24814.525+[(-72.457*SRAD)+ (1117.9*Tmax) + (-886.66 *Tmin) + (-0.42*Rain) +(229.814*RH)]	0.402
Kollam	-9394.00022+[(-50.9268717*SRAD)+(268.791473*Tmax)+ (-128.5654598*Tmin)+(-0.210876242*Rain)+ (94.76093136*RH)]	0.154
Pathanamthitta	15935.97461+[(151.4430698*SRAD)+(218.620536*Tmax) + (-102.2300204*Tmin) + (0.411234259 * Rain) + (-97.2348 *RH)]	0.266
Alappuzha	-18698.22033+[(-174.1216152*SRAD)+ (-1285.11114*Tmax)+(2273.5951*Tmin)+(-0.596109923*Rain) +(17.68290016*RH)]	0.143
Kottayam	-41430.91748+[(-46.8756684*SRAD)+(1389.712194*Tmax)+ (-1408.48217*Tmin)+(-0.83358467*Rain) +(447.0837839*RH)]	0.482
Idukki	10592.7036+[(115.0532401*SRAD)+(93.01397708*Tmax)+ (-280.0157007*Tmin)+(0.382678687*Rain)+ (-28.58750841*RH)]	0.657
Ernakulam	4378.98573+[(-39.83024951*SRAD)+(-51.08197118*Tmax)+ (-25.17810597*Tmin)+(-0.003984433*Rain) +(3.30839941RH)]	0.499
Thrissur	-4475.196083+[(-157.4413573*SRAD)+(261.8551*Tmax)+ (-36.76465031*Tmin)+(-0.42951265*Rain)+ (34.82497276*RH)]	0.556
Palakkad	-8039.61987+[(-155.24489*SRAD)+(438.3396118*Tmax)+ (-257.7852899*Tmin)+(0.343905266*Rain)+ (68.14430203*RH)]	0.318
Malappuram	119.2199927+[(- 103.530073*SRAD)+(38.04420186*Tmax)+(148.3815033*Tmin)+ (-0.149697116*Rain)+ (-6.00067641*RH)]	0.365
Kozhikode	1877.362669+[(- 131.5805962*SRAD)+(515.9208771*Tmax)+(555.7013247*Tmin)+ (-0.1001878*Rain) +(13.77365569*RH)]	0.349
Kannur	-38350.38441+[(-143.7347362*SRAD)+(1095.848947*Tmax)+ (-558.9084821*Tmin)+(0.004723617*Rain) +(288.8652*RH)]	0.362
Kasaragod	-27085.38714+[(-98.92614686*SRAD)+(987.8737241*Tmax)+ (-448.1913974*Tmin)+(0.0538434*Rain) +(167.4537938*RH)]	0.457

SRAD= solar radiation (MJm^{-2}), Tmax - Maximum temperature (℃), Tmin- Minimum temperature (℃), Rain- Rainfall

humidity, and the product of rain with solar radiation positively influenced yield, while the combination of rain and solar radiation had a negative influence.

Kottayam: Maximum temperature and solar radiation combined with relative humidity had positive effects, while minimum temperature and rain with relative humidity were negative influences.

Idukki: Relative humidity, solar radiation, and the combination of maximum temperature with solar radiation had positive effects, while both maximum and minimum temperatures negatively impacted yield.

Ernakulam: Relative humidity and combinations of maximum temperature, rain, and relative humidity showed positive effects, while minimum temperature and the combination of rain with solar radiation were negative influences.

Thrissur: Relative humidity, as well as

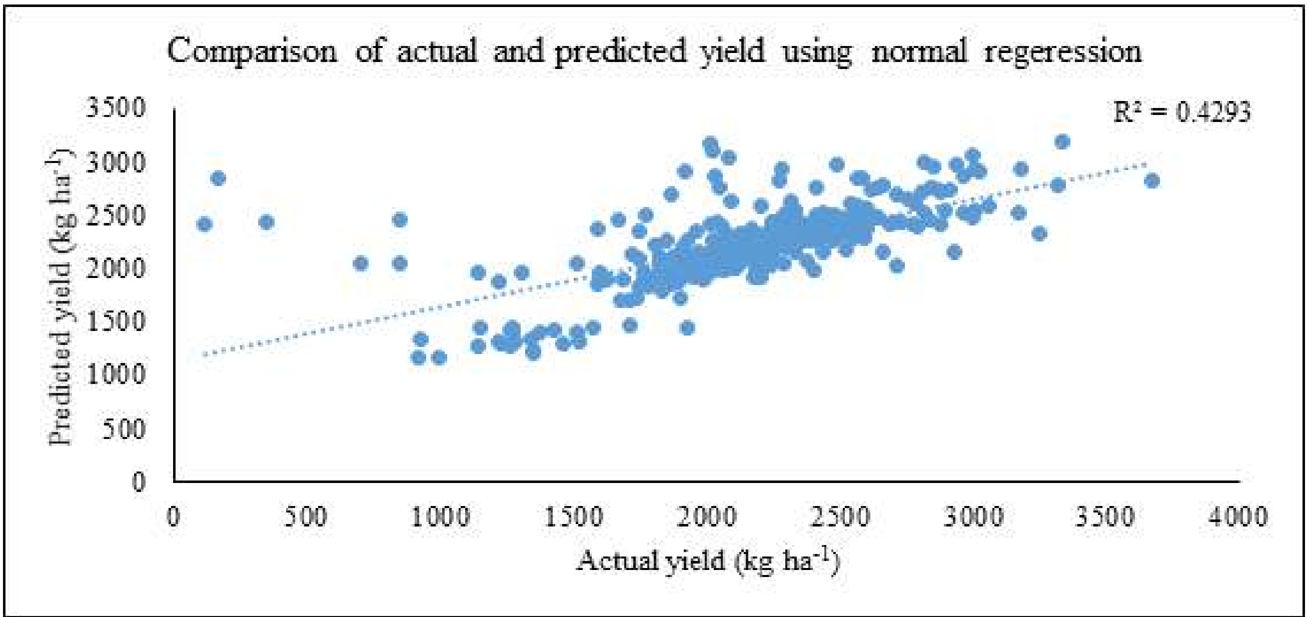

Fig. 4. Comparison of actual and predicted yield by normal regression

Table 2. Yield prediction models for different districts of Kerala using SMLR

Districts	Equation	R^2 value
Thiruvananthapuram	6887.346+(40.129*Time)+(0.53*Z251) +(28.845*Z21)	0.925
Kollam	1834.602+(22.814*Time) +(0.241*Z451)	0.75
Pathanamthitta	21688.97+(4.432*Z241) +(256.647*Z21) + (-0.08*Z351) +(-156.086*Z11)	0.888
Alappuzha	22852.28+(1.306*Z451) +(81.268*Time) +(3.332*Z141)	0.791
Kottayam	109.983+(45.653*Time) +(253.301*Z51) + (-8.932*Z251)	0.876
Idukki	6418.853+(43.203*Z11) +(1.219*Z241)	0.77
Ernakulam	4788.089+(0.293*Z151) +(7.922*Time) +(-5.901*Z20)	0.817
Thrissur	1186.741+(96.06*Z11) + (0.016*Z450) +(17.413*Time) +(1.267*Z141)	0.916
Palakkad	6157.388+(41.239*Time) +(0.115*Z131) + (-0.102*Z140)	0.937
Malappuram	4794.502+(1.697*Z141) +(18.828*Time) + (-0.124*Z140) +(0.388*Z151)	0.901
Kozhikode	1332.789+(0.093*Z351) +(21.323*Time)	0.573
Kannur	1782.151+(41.151*Time) +(0.051*Z351)	0.832
Kasaragod	5213.971+(79.516*Z21) +(17.67*Time) +(0.284*Z451) +(-0.199*Z240)	0.872

combinations of maximum temperature, rain, relative humidity, and solar radiation, positively affected yield, while the combination of relative humidity and solar radiation had a negative effect.

Palakkad: Combinations involving maximum temperature, rain, relative humidity, and solar radiation positively influenced yield, but relative humidity alone and the interaction of maximum and minimum temperatures had negative impacts.

Malappuram: Minimum temperature, rainfall, solar radiation, and their combination with relative humidity had a positive effect, but minimum temperature alone was a negative influence.

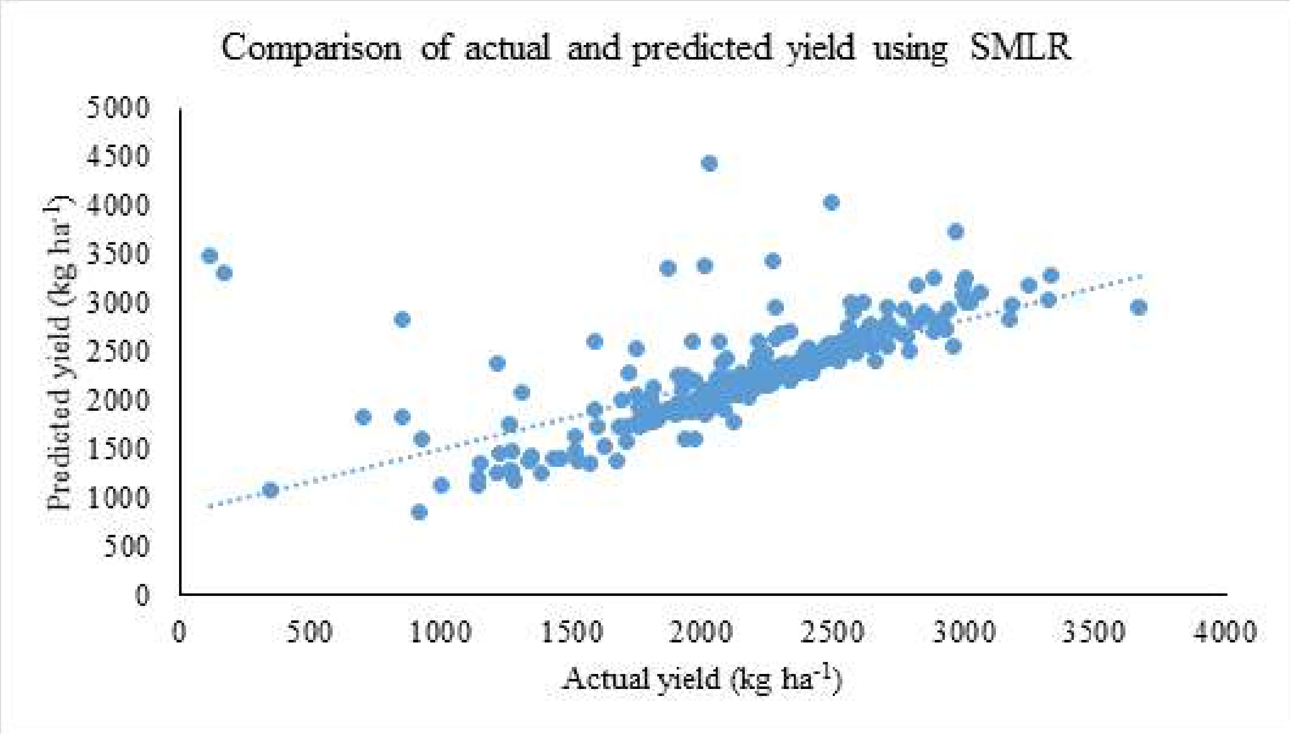

Fig. 5. Comparison of actual and predicted yield by SMLR

Table 3. ANN model: Number of hidden neurons, R^2 and RMSEC

District	No. of Hidden neurons	R^2 ($p < 0.01$)	RMSEC (kg ha^{-1})
Thiruvananthapuram	9	0.997	14.68938
Kollam	6	0.996	16.21406
Pathanamthitta	7	0.846	106.2033
Alappuzha	8	0.838	293.3171
Kottayam	2	0.801	267.8696
Idukki	5	0.898	53.3485
Ernakulam	9	0.994	16.29997
Thrissur	4	0.989	20.79367
Palakkad	9	0.975	29.32695
Malappuram	6	0.934	36.10754
Kozhikode	9	0.986	23.20222
Kannur	8	0.998	8.485507
Kasargod	8	0.997	13.03234

Kozhikode: Maximum temperature combined with relative humidity positively affected yield, while rain and solar radiation had negative effects.

Kannur: Minimum temperature and relative humidity were strongly positive influences, whereas combinations of maximum temperature, minimum temperature, and relative humidity had negative effects.

Kasargod: Minimum temperature, rainfall, and solar radiation positively influenced yield, while combinations of rainfall, relative humidity, and solar radiation showed a negative effect.

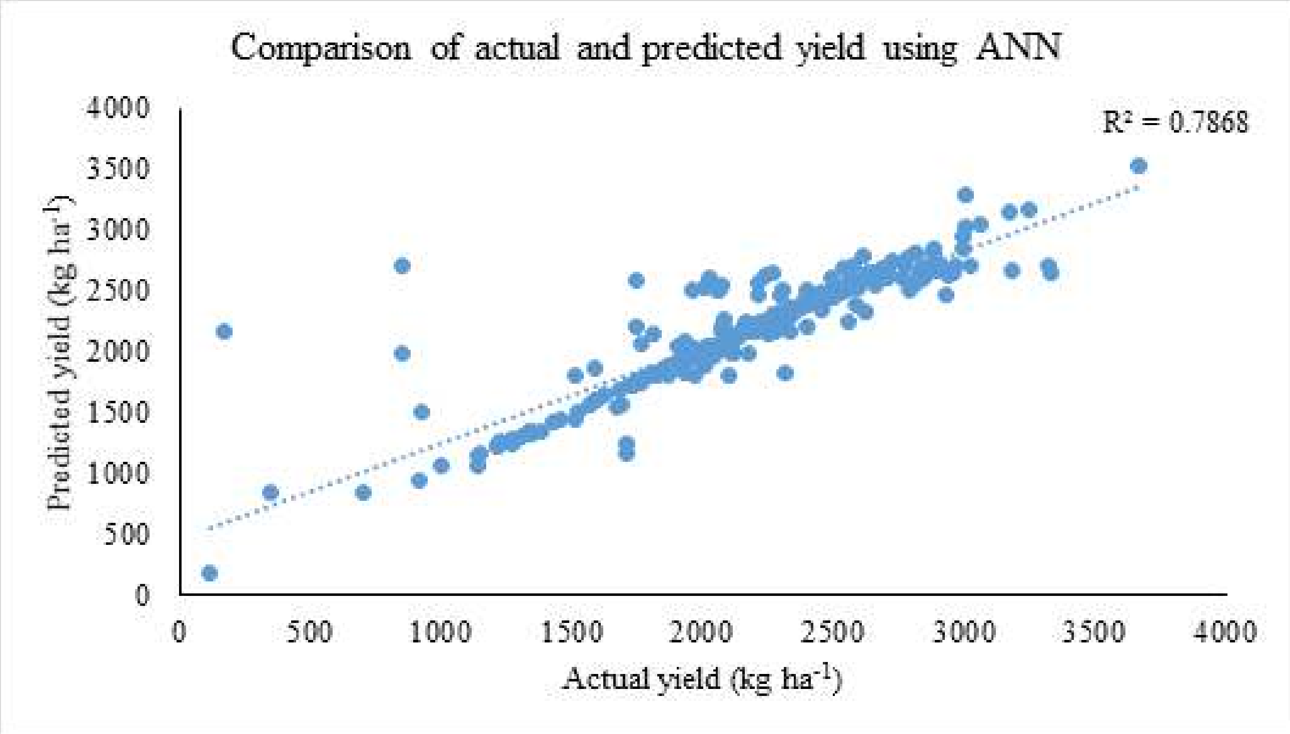

Fig. 6. Comparison of actual and predicted yield by ANN

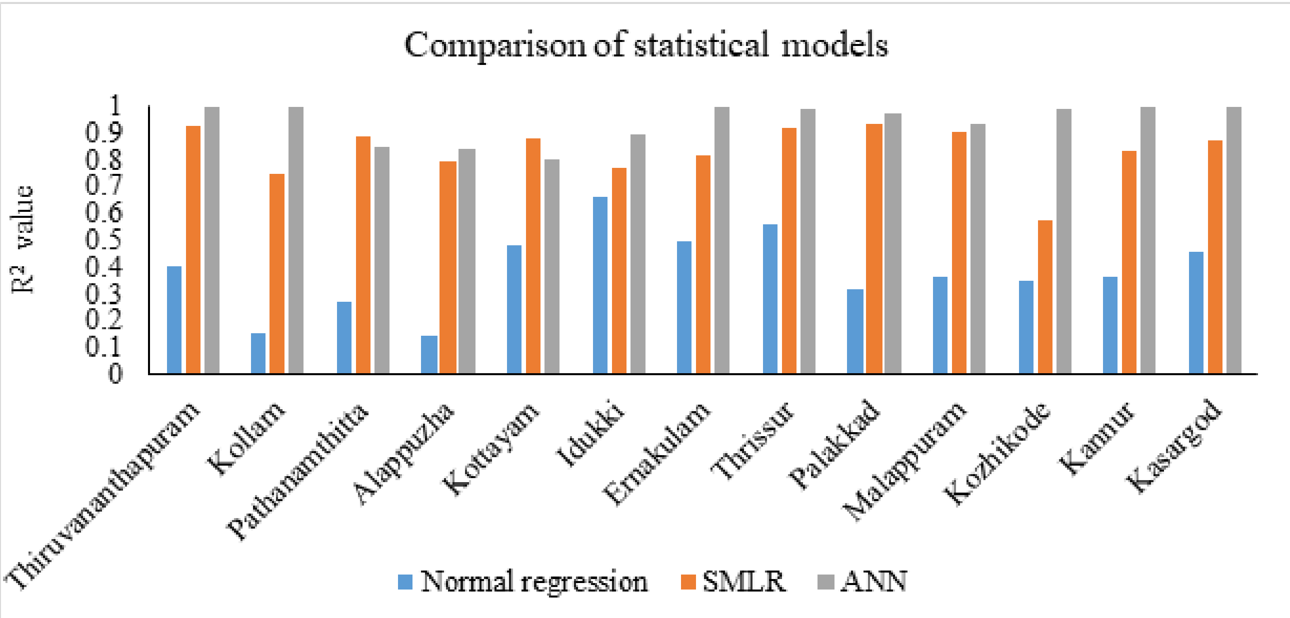

Fig. 7. Comparison of statistical models

Comparison of yield obtained from R2 value of Normal regression, ANN and SMLR

Across all districts in Kerala, the R^2 values were consistently higher in the ANN models compared to those from Normal Regression and SMLR. This indicates that ANN provided a more accurate fit for the yield prediction, making it the most effective modeling approach for all districts in Kerala (Fig. 7).

CONCLUSION

This study developed yield prediction models for rice across Kerala's districts using three statistical methods: Normal Regression, Artificial Neural Networks (ANN), and Stepwise Multiple Linear Regression (SMLR). Weather parameters such as temperature, rainfall, relative humidity, and solar radiation were used to assess their impact on yield. ANN consistently showed the highest R^2 values among the models, indicating superior predictive accuracy across all districts. Thus, the ANN model proved to be the most effective approach for accurately predicting rice yield in Kerala, highlighting its potential as a reliable tool for agricultural forecasting in the region.

REFERENCE

Ajithkumar B, Rao V U M, Sreekala P, Chandrasekhar Vishnu and Sekhar E (2015). Agroclimatic Atlas of Kerala. https://www.researchgate.net/publication/341788375_Agroclimatic_Atlas_of_Kerala

Bhattacharya T and Panda RK (2013). Effect of climate change on rice yield at Kharagpur, West Bengal. *Int J Food Agric Vet Sci 4*(2): 6-12.

Das B, Nair B, Reddy VK and Venkatesh P (2018). Evaluation of multiple linear, neural network and penalised regression models for prediction of rice yield based on weather parameters for west coast of India. *Int J Biometeorol* **62**(10):1809-1822.

Dutta S, Patel N K and Srivastava S K (2001). District-wise yield models of rice in Bihar based on water requirement and meteorological data. *J Indian Soc Remote Sensing* **29** :175-182.

GOK [Government of Kerala](2022). *Agricultural Statistics* 2021-22 [on-line]. Available: https://www.ecostat.kerala.gov.in/publication-detail/agricultural-statistics-2021-22

Hinton PR, Brownlow C and McMurray I (2004). *SPSS explained*. Psychology Press.

Kuhn M (2008). Building predictive models in R using caret package. *J Stat Softw* **28**:1–26

Lobell DB and Burke MB (2010). On the use of statistical models to predict crop yield responses to climate change. *Agric For Meteorol* **150**: 1443–1452. https://doi.org/10.1016/j.agrformet.2010.07.008

Rai KK, N P V, Bharti BVS, SA K (2013). Pre - harvest forecast models based on weather variable. *Adv Biores* **4**:118–122

Shi W, Tao F, Zhang Z (2013) A review on statistical models for identifying climate contributions to crop yields. *J Geogr Sci* **23**:567–576. https://doi.org/ 10.1007/s11442-013-1029-3

Singh RS, Patel C, Yadav MK and Singh KK (2014). Yield forecasting of rice and wheat crops for eastern Uttar Pradesh. *J Agrometeorol* **16**:199–202

Yoshida S, T Satake and D Mackill (1981). *High-temperature Stress*. International Rice Research Institute Paper No. 67: 1–15.

Received on 31/10/2024 Accepted on 20/11/2024