Original Research Article

Comparison of linear and non-linear models for coconut yield prediction in Coimbatore using weather parameters and external factors.

Abstract:

Coconut is the world's most significant plantation crop, and it is grown in practically every country. As Coimbatore is the leading producer of coconut in Tamil Nadu, followed by Thanjavur and Kanyakumari, this study is centred on the Coimbatore area. West Coast Tall is a popular cultivar that produces more than other types. The West Coast Tall (WCT) cultivar was used in this research. In this paper, five models were developed such as Multiple Linear regression (MLR), Ridge, Least Absolute Shrinkage Selection Operator (LASSO), Elastic net (ELNET) regression methods and Artificial Neural Networks (ANN). Further, we validate this model using field-level data from TNAU coconut research farm for two years. The purpose of this communication is to find the best fit model for prediction of coconut yield using weather parameters and external factors in Coimbatore district. The models were selected based on different performance metrics such as RMSE, MAPE, MAE, and R^2.

Keywords: *Coconut, Yield Prediction, ANN, Penalized regression models*, *Numerical Analysis.*

1. INTRODUCTION:

Coconut (*Cocos nucifera L.*), an important perennial oil-producing crop of the humid tropics, is widely grown in countries. Coconut is the world's most prolific crop, and India has the largest area under coconut. Most of India's coconut area and production are in four southern states, namely Kerala, Tamil Nadu, Karnataka and Andhra Pradesh (C. Palaniswami et al., 2008). The coconut tree is renowned as the "Tree of Heaven" because every portion of the coconut palm may be utilised in some way. The coconut fruit and flowers are produced throughout the year, making it unique among plantation crops. The development of a coconut fruit takes approximately 44 months from the onset of inflorescence primordium to full maturity(Rajagopal et al., 1996). Crop yield is affected by both genotype and environment. Among the environmental variables, weather has a significant impact on crop output potential.Though, a number of regression models based on weather parameters have been used to predict coconut yield (Peiris et al. 2008, Mahesha et al., 1992, Pathmeswaran et al., 2018), but comparison of multiple statistical models has received very much less attention. Multiple linear regression (MLR), on the other hand, is appropriate for smaller datasets, but its application is limited when the number of predictors exceeds the number of samples (Balabin et al., 2011). MLR results in overfitting of data when the number of samples is less than the number of predictors and presence of multicollinearity (Verma et al., 2016). To overcome this, variable selection and use of penalized regression methods such as SMLR, Ridge, LASSO, ELNET regression and advanced regression models is recommended for developing yield prediction model.As part of their study, Jayashree et al. (2015) examined six models for coconut yield prediction, including multilayer perceptron, support vector machine, decision tree, Naive Bayes, data-driven nonlinear Hebbian, and fuzzy cognitive map as a combination of soil and weather variables. Bappa Das *et al.*, (2020)have predicted yield of coconut using only weather indices and compared different multivariate techniques and the study revealed that the elastic net regression method had been more accurate for predicting coconut ~~yiel~~ yield.According to our literature survey, no model had been reported by using both weather parameters and external factors for coconut yield predictions. In this paper, both weather parameters and external factors are taken into account to achieve prediction accuracy. The objective of this study intends to find the best fit model for prediction of coconut yield in Coimbatore district.

2. MATERIALS AND METHODOLOGY:

2.1 Study area and data source:

The daily weather parameters were collected in Agro-Climatic Research Centre, Tamil Nadu Agricultural University, Coimbatore. The daily weather data are taken on an average to form year wise weather data. The external factors were collected year wise from Coconut Research Station, Aliyar for the years (2010 - 2016). The primary data for validation had been collected from Coconut Farm, Tamil Nadu Agricultural University for two years (2021 and 2022).

2.2. Methodology:

2.2.1Ridge Regression:

Ridge regression reduces the magnitude of the regression coefficients by imposing a penalty on their size. The ridge coefficients minimize a penalized residual sum of squares,

$$\widehat{\beta}\,(Ridge) = \ arg\,min_{\beta}\,||y - X\beta||^2 + \ \lambda||\beta||^2$$

Here λ ≥ 0 is a complexity parameter that controls the amount of shrinkage: the larger the value of λ, the greater the amount of shrinkage. When there are many correlated variables in a linear regression model, their coefficients can become poorly determined and exhibit high variance (Trevor Hastie et al. 2008, Textbook on Elements of Statistical learning).

2.2.2LASSO Regression:

"LASSO" is an acronym that stands for Least Absolute Shrinkage and Selection Operator. In this method, the regression coefficients are reduced towards zero by penalizing the regression model with the penalty, which sums the absolute coefficients. With Lasso regression, the penalty results in some coefficient estimates being exactly equal to zero, with a minor contribution to the model (Tibshirani, 1996). Compared with ridge regression, lasso regression produces simpler and more interpretable models that incorporate fewer predictors.

$$\widehat{\beta}\,(Lasso) = \ arg\,min_{\beta}\,||y - X\beta||^2 + \ \lambda||\beta||$$

where, $||\beta|| = \sum_{j=1}^{p} |\beta_j|$ and λ is penalty on β. The LASSO shrinks some parameters to zero for suitable lambda.

2.2.3ELNET Regression:

ELNET regression stands for elastic net regression which is the combination of penalties from both LASSO and ridge regression (Hoerl and Kennard, 1970) which makes the regularization of statistical model better. The L1 part of the penalty produces a sparse model during the regularisation operation. The quadratic component of the penalty (L2), on the other hand, makes the L1 part more stable on the route to regularisation, eliminates the quantity limit of variables to be chosen, and promotes the grouping effect.Therefore, it minimizes the impact of different features while not eliminating all of them (Cho et al., 2009).

$$\hat{\beta}(Enet) = \left(1 + \frac{\lambda_2}{n}\right)\left\{arg\ min_{\beta}||y - X\beta||^2 + \lambda_1||\beta|| + \lambda_2||\beta||^2\right\}$$

where, λ_1 and λ_2 are LASSO and ridge regression penalities.

Lambda must be optimized in each of the three methods listed above, A cross-validation with leave-one-out was used to select the lambda values that minimized the average mean squared error (Piaskowski et al., 2016). The overall strength of the penalty is controlled by tuning parameter λ (Hastie and Qian, 2014). Analysis of the data was performed using the R package 'glmnet' (Friedman et al., 2009).

2.2.4Artificial Neural Network:

A neural network is a massively parallel network of interconnected simple processors (neurons) in which neuron accept a set of inputs from other neurons and computes an output that is propagated to the output nodes. Thus, a neural network can bedescribed in terms of individual neurons, the network connectivity, the wights associated with the interconnections with neurons, and the activation function of neuron. The neuron receives a set of n inputs, x_i, $i = 1,2,\dots,n$, from its neighbouring neurons and a bias which is equals to 1. Each input has a weight (w_i) associated with it. The weighted sum of the inputs determines the state or activity of a neuron and is given by

$$a = \sum_{i=1}^{n+1} w_i x_i = W^T X$$

Where, $X = \{x_1 x_2 \dots x_n\ 1\}^T$. The output of the neuron is commonly described by a sigmoid function as

$$f(a) = \frac{1}{1 + e^{-a}}$$

The schematic diagram of single artificial neural network is as follows:

In artificial neural network, three separate functional processes take place. The product Xw is created by multiplying input X by weight w. Next, the net input (Xw+b) is then formed by adding bias to weighted input. In this case, the function is shifted by a factor of b because of the bias. In the end, the net input is fed into an activation function that generates the output Y.

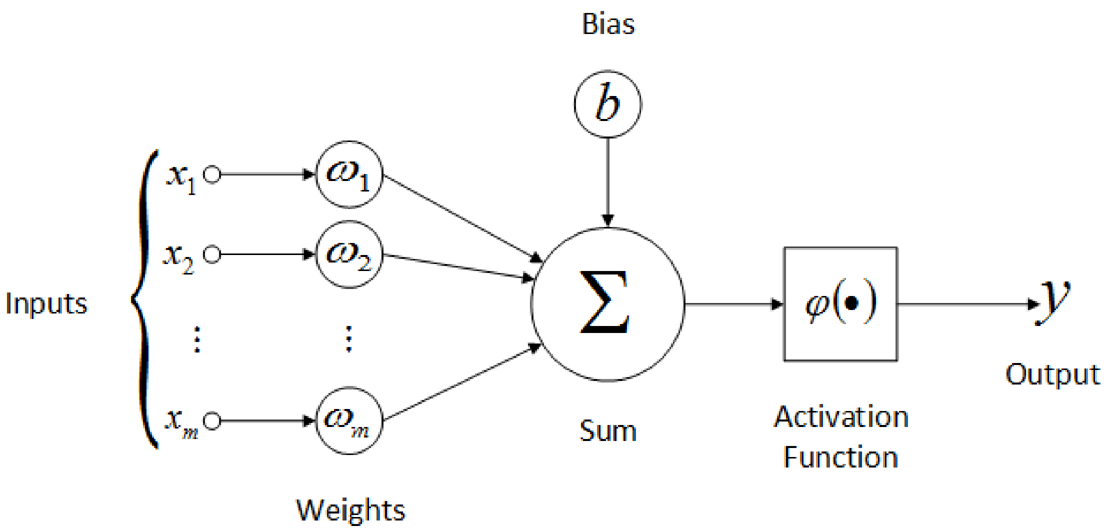

Fig 1. Schematic diagram of ANN.

2.3Model performance metrices:

The performance of the statistical models was evaluated using coefficient of determination (R^2), Root mean squared error (RMSE), mean absolute percentage error (MAPE), and Mean absolute error (MAE) by following formulas:

$$R^2 = 1 - \frac{\sum_{j=1}^{n}\left(Y_j - \hat{Y}_j\right)^2}{\sum_{j=1}^{n}\left(Y_j - \bar{Y}_j\right)^2}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{j=1}^{n}\left(Y_j - \hat{Y}_j\right)^2}{n}}$$

$$\text{MAPE} = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{Y_j - \hat{Y}_j}{Y_j}\right| * 100$$

$$\text{MAE} = \frac{\sum_{j=1}^{n}\left|Y_j - \hat{Y}_j\right|}{n}$$

where, Y_j – Actual yield, $\hat{Y}_j$ – Model yield respectively, n-number of years.

The coefficient of determination (R^2) and RMSE approaches 0 and also the lesser MAPE and MAE values indicates that the statistical model is the best model for prediction.

3. RESULT AND DISCUSSION:

3.1 Statistical Analysis:

The intercept and coefficients of various predicted models are shown in the Table 1. All the fitted models in the table revealed that the variables relative humidity, average rainfall, leaf length and copra content had positive contribution according to yield and the variables such as minimum temperature, plant height, stem girth and female flower were negatively related to yield.

Table 1. Intercept and coefficients of different fitted models:

	Ridge	LASSO	ELNET	MLR
Intercept	1115.0753	609.5997	1913.9497	0.0025
Minimum Temperature	-28.5280	0.2371	0.4741	0.01
Relative Humidity	12.5400	1.3780	1.6911	19.606
Average Rainfall	0.6308	0.0154	0.1429	1.4735
Plant height	-0.2195	-1.2270	-1.3083	-1.1593
Stem girth	-4.2695	-10.1670	-26.7849	-26.072
Leaf length	0.9958	4.9749	4.7098	4.8373
Female flower	-1.8312	-0.0038	-6.0430	-10.928
Copra content	1.3102	3.4187	6.9698	7.4639

Figures (2-6) shows the graphical representation of observed yield and predicted yield of Ridge, LASSO, ELNET, MLR and ANN respectively. In figure 4, the predicted yield is far away from the observed yield during all the time which leads us to conclude that ridge regression is not suitable for predicting coconut yield in Coimbatore district. Among the figures the MLR model (Fig. 8) predicted yield is very much closer to observed yield than the other prediction models.

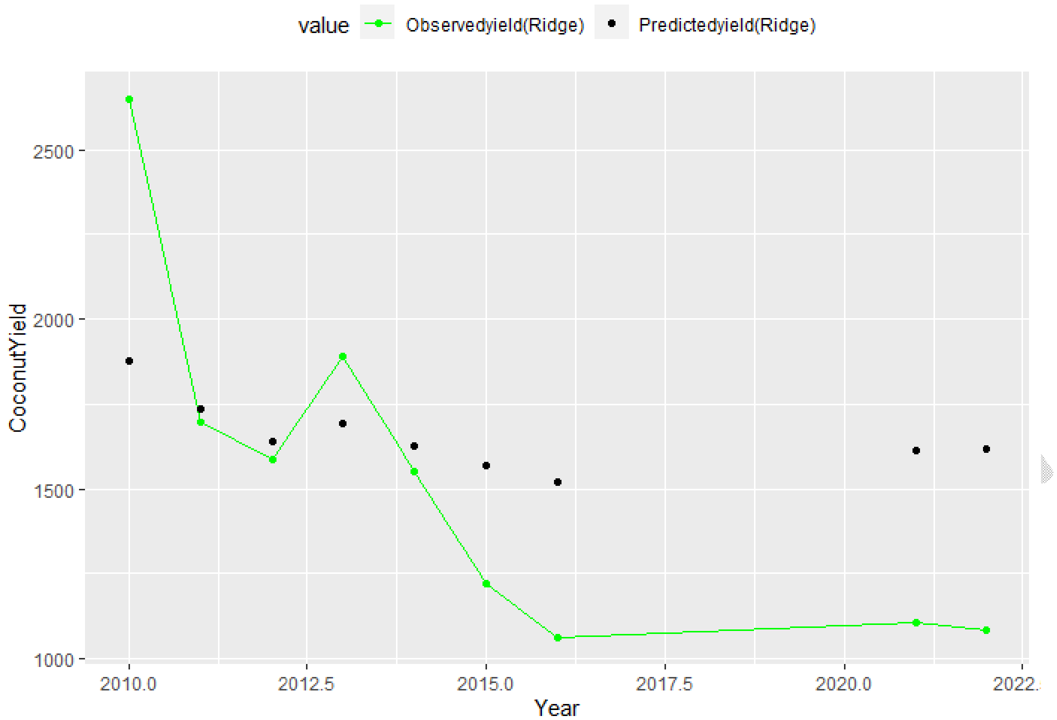

Fig 2. Visualization of observed and predicted yield by ridge regression.

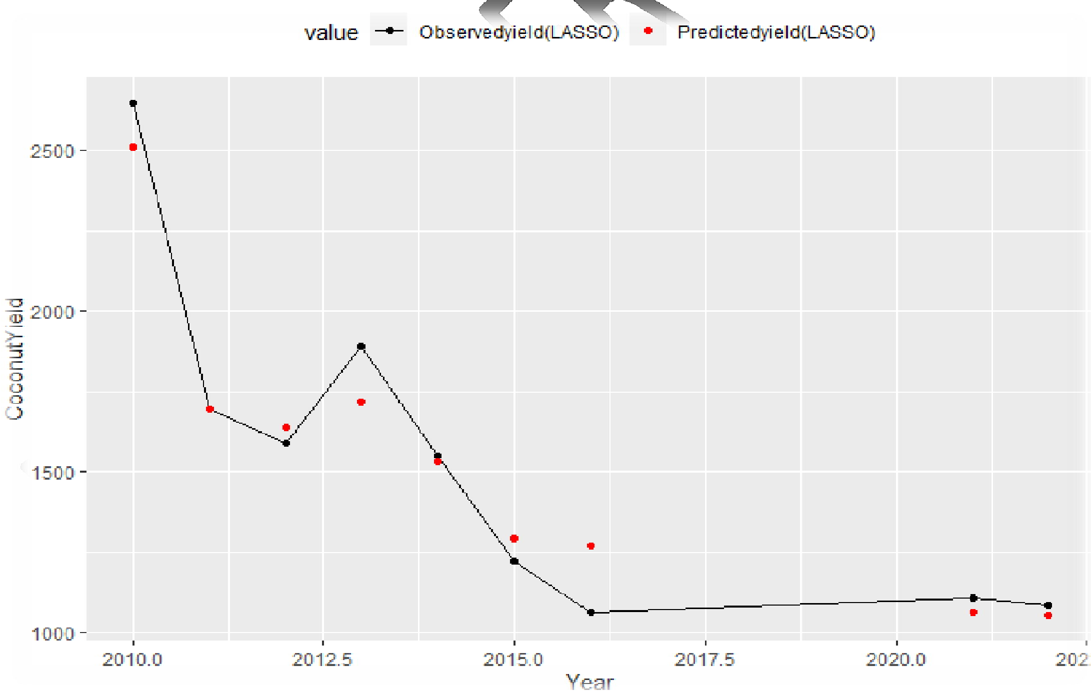

Fig 3. Visualization of observed and predicted yield by LASSO regression.

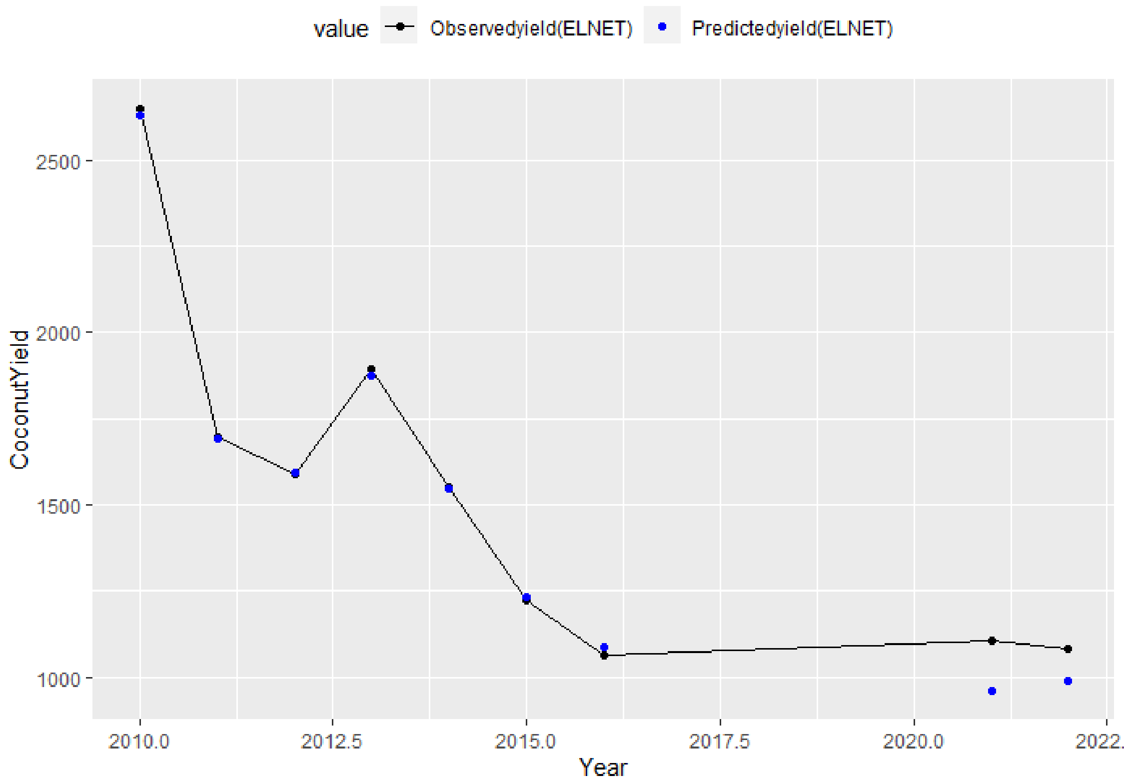

Fig 4. Visualization of observed and predicted yield by ELNET regression.

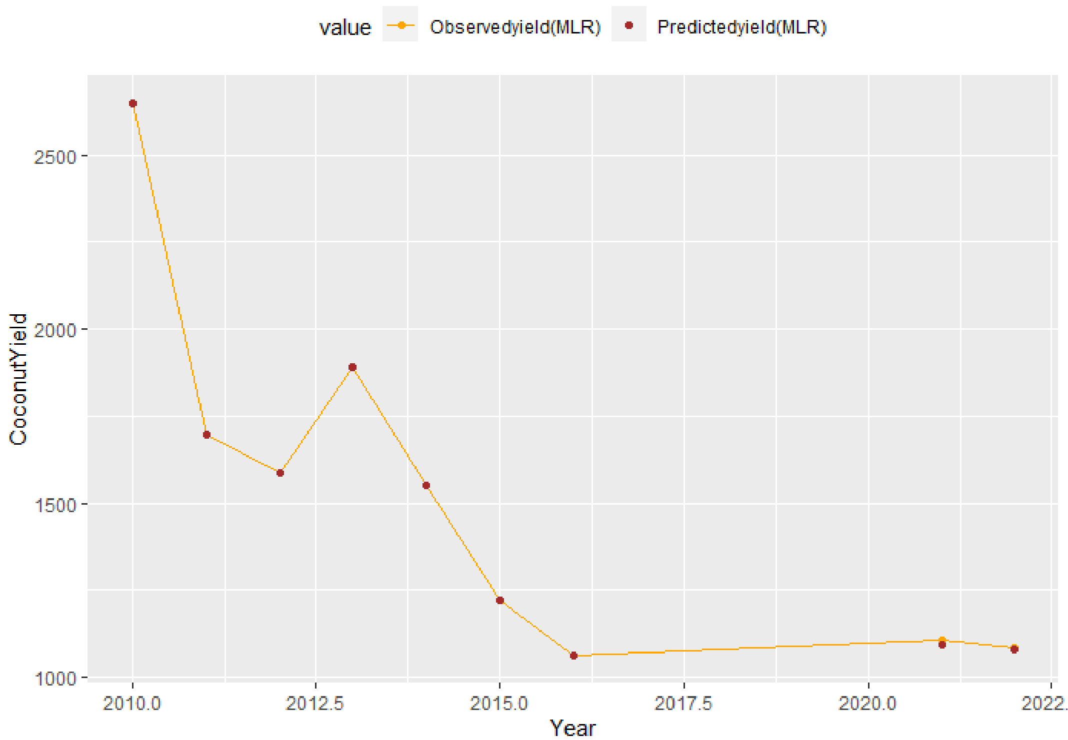

Fig 5. Visualization of observed and predicted yield by MLR.

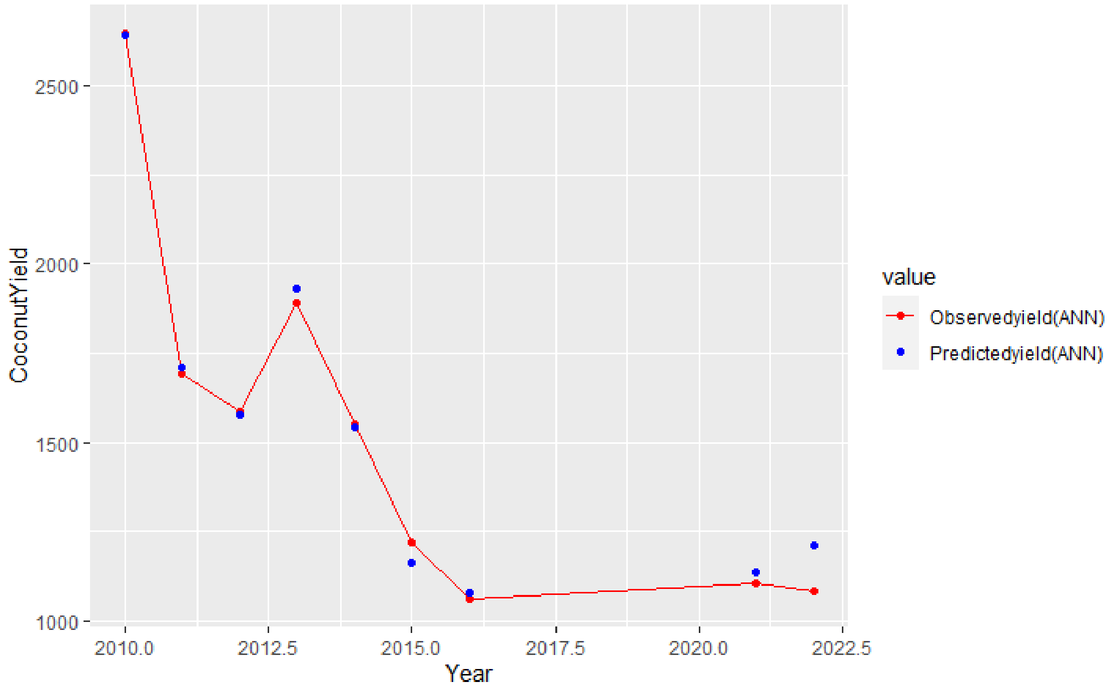

Fig 6. Visualization of observed and predicted yield by ANN.

3.2 Artificial Neural Network:

The number of neurons in each layer and the hidden layers in a particular problem was the major factor in constructing the neural network. So, it is most important to select the optimum number of hidden layers in a neural network. In R software, we used 'caret' package to tune the number of hidden layers. RMSE was used as a performance evaluator to find the optimal connection weights for neural network training. The diagrammatic representation of Artificial Neural Network was given by Fig 7.

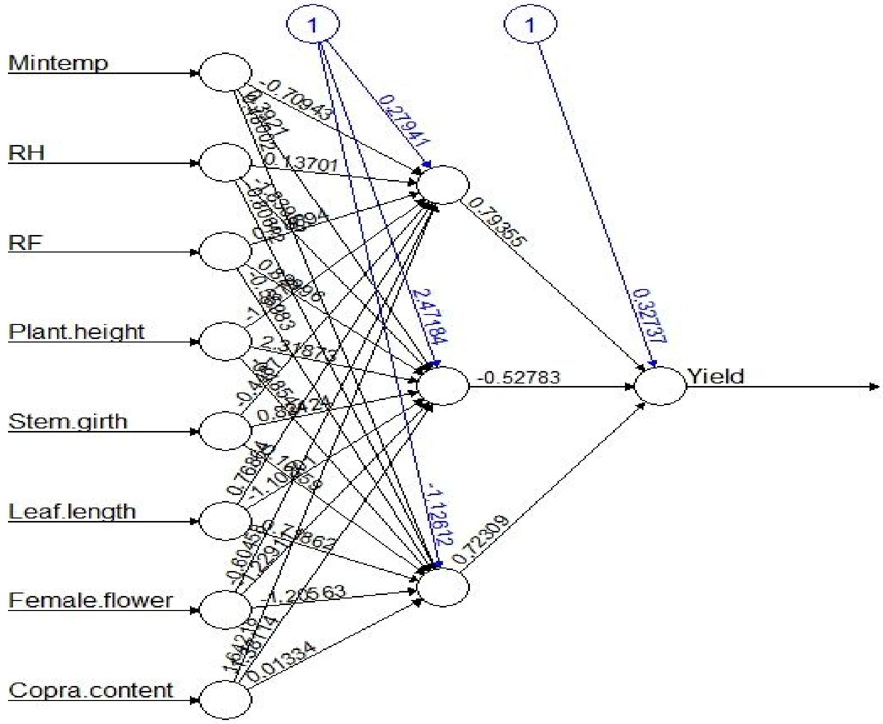

Fig 7. Graphical representation of ANN.

3.3 Intercomparison of models:

Table 2. Performance metrics for various models:

Models	RMSE	R^2	MAPE	MAE
Ridge	410.3293	0.8614	24.09	331.6014
LASSO	106.8821	0.9627	5.66	81.6873
ELNET	58.9887	0.9898	3.04	35.8974
ANN	92.5574	0.9268	7.31	79.63
MLR	3.93	0.99	0.16	1.7136

The models were compared using differentperformance metrics. The coefficient of determination (R^2) is high for MLR model (0.99) followed by ELNET (0.989) and LASSO (0.962). The RMSE, MAPE and MAE values was found to be least in MLRwhich is 3.93, 0.16%, 1.713 followed by ELNET and LASSO. Figures 8 and 9 were used to compare the model based on R^2 and mean absolute percentage error which clearly confirms that the model MLR is the best fit model.

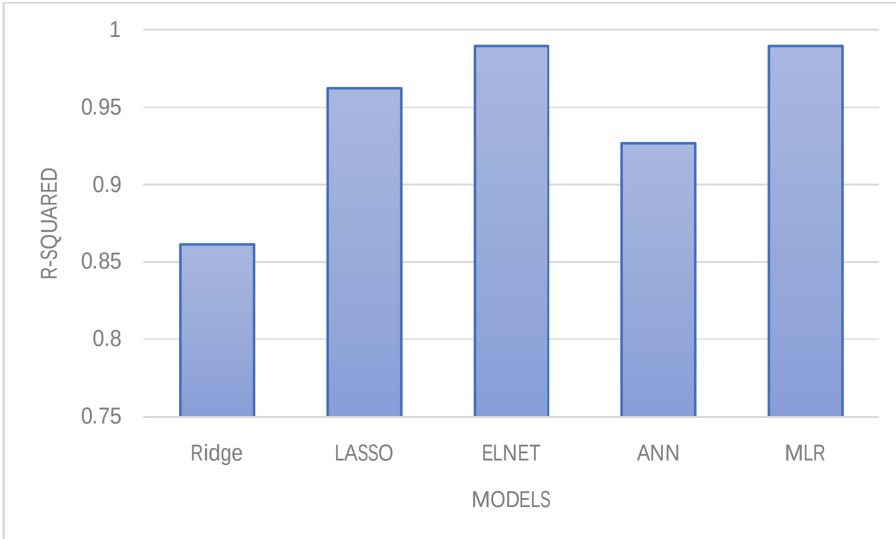

Fig 8. Comparison of different models using R-squared.

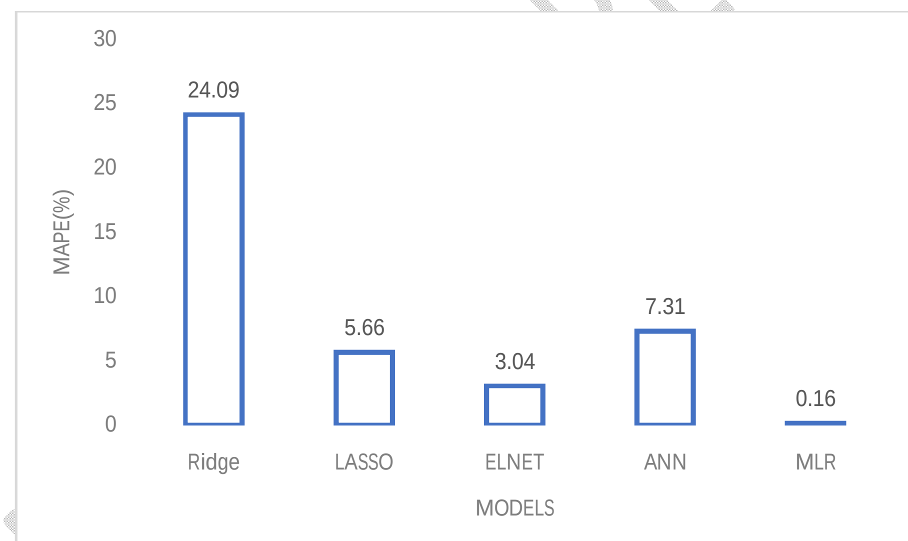

Fig 9. Comparison of different models using Mean absolute percentage error.

4. CONCLUSION:

In this paper, five different methods were employed in predicting the coconut yield and the models were compared and ranked by different goodness of fit measures. The study revealed that the model from MLR was found to be the best fit model for the available data for coconut yield prediction. So, MLRmodel is recommended for prediction of coconut yield for future perspectives.

5. REFERENCES:

1. Bappa Das., Bhakti Nair., Vadivel Arunachalam. and K. Viswanatha Reddy. 2020. Comparative evaluation of linear and nonlinear weather-based models for coconut yield prediction in the west coast of India. *Int. J. Biometeorol.,*64:1111–1123
2. Balakrishnan, M and K. Meena. (2020). Coconut yield prediction using back propagation and Split algorithm in bay islands. *CLIO an Annual Interdisciplinary Journal of History*.,6(2): 549-557.
3. Mahesha, A. and N. Lakshman. (1993). Influence of weather variables on coconut (Cocos nucifera L.) yield. *Mausam*.,44(1):102-104.
4. Peiris, T.S.G., Hansen, J. W. and Lareef Zubai, 2008. Use of seasonal climate information to predict coconut production in Sri Lanka. *Int. J. Climatol.*,28: 103–110.
5. Balabin RM, Lomakina EI, Safieva RZ (2011) Neural network (ANN) approach to biodiesel analysis: analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy, Fuel 90(5):2007–2015.
6. Pathmeswaran, C., Lokupitiya, E., Waidyarathne, K.P. and R.S. Lokupitiya. 2018. Impact of extreme weather events on coconut productivity in three climatic zones of Sri Lanka. *Eur J Agron*.,96:47–53.
7. Cho, S., Kim., Oh, S., Kim, K., and Park, T. 2009. Elastic net regularization approaches for genome wide association studies of rheumatoid arthiritis. *BMC proceedings.* 3(7):1-6.
8. Tibshirani, R. 1996. Regression shrinkage and selction via the lasso. *Journal of the royal Statistical Society:* Series B (*Methological*)., 58(1):267-268.
9. Hoerl, A. E., and Kennard, R. W. 1970. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1): 55-67.
10. Friedman, J., Hastie, T., and Tsbshirani, R. 2009. Glmnet: Lasso and elastic net regularized generalized linear models. *R package version.,* 1(4): 1-24.
11. Piaskowski, J. L., Brown, D., and Campbell, K.G. 2016. Near infrared calibration of soluble stem carbohydrates for predicting drought tolerance in spring wheat. *Agronomy Journal.,* 108(1): 258-293.
12. Hastie, T., and Qian, J. 2014. Glmnet vignette., 9:1-30.
13. Jayashree LS, Palakkal N, Papageorgiou EI, Papageorgiou K. 2015. Application of fuzzy cognitive maps in precision agriculture: a case study on coconut yield management of southern India's Malabar region. *Neural Comput & Applic* 26:1963–1978.
14. Singh., R. S., Patel, C., Yadav, M. K., and Singh, K. K. 2014. Yield forecasting of rice and wheat crops for eastern Uttar Pradesh. *Journal of Agrometerology.,* 16(2):199.
15. Draper, N. R and Smith H. 1998. Applied regression analysis. John Wiley & Sons, Hoboken

APPENDIX: R code for ANN:

```
set.seed(132)
data<-data.frame(data)
train_id<-data[1:7,]
test_id<-data[8:9,]
```

```
max = apply(data , 2 , max)

min = apply(data, 2 , min)

scaled = as.data.frame(scale(data, center = min, scale = max - min))

trainNN = scaled[1:7,]

testNN = scaled[8:9 , ]

library(neuralnet)

n<-neuralnet(Yield~.,data = trainNN,hidden=3,linear.output = TRUE)

n$result.matrix

plot(n)

predict_testNN1 = compute(n, testNN[,-1])

predict_testNN1

predict_testNN = (predict_testNN1$net.result * (max(data$Yield) - min(data$Yield))) + min(data$Yield)

predict_testNN

plot(test_id$Yield, predict_testNN, col='blue', pch=16, ylab = " Predicted", xlab = "Actual")

abline(0,1)

RMSE.NN = (sum((test_id$Yield - predict_testNN)^2) / nrow(testNN)) ^ 0.5

RMSE.NN

predicted=predict_testNN

predicted

actual=test_id$Yield

comparison=data.frame(predicted,actual)

deviation=((actual-predicted)/actual)

comparison=data.frame(predicted,actual,deviation)

comparison

accuracy=1-abs(mean(deviation))

accuracy

library(Metrics)

mae(actual,predicted)

mape(actual,predicted)

rmse(actual, predicted)
```