Research Article

Journal of Plantation Crops, 2008, 36 (1): 49-52

Application of nonparametric additive model for input-response analysis in arecanut

C.T. Jose[1] and Ravi Bhat

Central Plantation Crops Research Institute, Regional Station, Vittal-574 243, Karnataka, India

(Manuscript Received: 22-02-07, Revised: 19-09-07, Accepted: 07-02-08)

Abstract

The relationships between the inputs and response are very complex in crop production models. Problems like nonlinear relationships between inputs and responses, nonexistence of proper functional form to represent relationship between inputs and response variable and multi-collinearity are very common in crop production data. The traditional multiple linear regression technique may not be adequate in many situations to explain input output relationship. In this paper we have used nonparametric additive regression model to explain input output relationship in arecanut. The comparative study shows that the nonparametric additive modeling technique performed much better than the multiple linear regression technique to explain the input response relationship. The estimated values of the component functions provide the mean response of input variables on the yield of arecanut. The optimum value of the input variables is obtained from the graphical representation of the component functions. The present analysis of data from the two districts show that input response relationship vary depending on the agro-climatic conditions of the locations.

Keywords: Additive model, arecanut, crop production model, input response analysis, nonparametric regression

Introduction

Arecanut is one of the important cash crops of India, which has a prominent place in the social, cultural, religious and economic activities of the society. It is grown under a variety of climatic and soil conditions. Based on the yield performance and other component characters, four high yielding varieties viz., Mangala, Sumangala, Sreemangala and Mohitnagar were released from CPCRI. There are a few other high yielding cultivars in different localities known by the name of the place, where they are grown. The yield of these varieties/cultivars varies between soil and climatic conditions as well as management practices. It is very much essential to have site-specific management practices to get maximum return from unit area. Most of the farmers are generally growing local cultivars and adopting traditional method of cultivation practices.

In this paper we have studied the response of inputs to the yield of arecanut in the major arecanut growing districts of Karnataka viz., Shimoga and Dakshina Kannada (DK) based on the data collected from farmers' garden. Shimoga is about 600 m above the MSL and receive about 2500 mm rainfall every year. The temperature ranges between 10° C to 40° C. The average altitude of the arecanut growing areas in Dakshina Kannada district is about 100 m above the MSL and the temperature ranges between 15° C to 38° C. It receives about 3500 mm rainfall every year. The analysis was carried out separately in each district. We have considered the most popular variety Thirthahalli local in Shimoga district and South Canara Local in Dakshina Kannada district for the study.

The relationships between the inputs (values of various parameters of agro ecosystem and cultivation

[1]Corresponding author: *Email: ctjos@yahoo.com*

practices) and response (yield) are very complex in crop production models. Problems like nonlinear relationships between inputs and responses, nonexistence of proper functional form to represent the relationship and multi-collinearity are very common in crop production data. Therefore, the commonly used multiple linear regression technique may not be adequate in many situations. In this paper we have used nonparametric additive regression model to explain input output relationship, which is more flexible than the multiple linear regression model. In nonparametric additive model we only assume that the relationship is a smooth function.

Materials and Methods

Data from the randomly selected 74 farmers' gardens in Dakshina Kannada district and 60 gardens in Shimoga district was used for the study. Gardens with proper irrigation facility, stabilized yield and more than 250 palms only have been selected. Observations on fertilizer application (N, P and K), Farm Yard Manure (FYM), Green leaf (GL) application and density (number of palms/ha) are used as input variables and average yield (kg/ha) was taken as the response variable for the study. The input variables N, P, K, GL, FYM and density are represented by X_1, X_2, X_3, X_4, X_5 and X_6 respectively. The response variable (yield/ha) is represented by Y.

Nonparametric additive model is used to study the input response relationship in arecanut. The nonparametric additive model extends the notion of a linear model by allowing some or all linear functions of the predictors to be replaced by arbitrary smooth functions of the predictors. Thus, the standard linear model of the form

$$Y = \beta_0 + \sum_{i=1}^{p} \beta_i X_i + \varepsilon$$

is replaced by the nonparametric additive model

$$Y = \alpha + \sum_{i=1}^{p} f_i(X_i) + \varepsilon$$

Where α is the intercept, is independently and identically distributed error term with mean 0 and constant variance. The functions (i = 1,…, p) are assumed to be smooth and $E[f_i(X_i)] = 0$ for i = 1,…p. The backfitting algorithm proposed by Buja *et al.* (1989) and the related fitting procedure in S-PLUS (Chambers and Hastie, 1992) have made the additive model a popular choice for multivariate nonparametric fitting.

The efficiency of the nonparametric additive model against the linear regression model is assessed by comparing the quality of both the fits. Let $\hat{y}_{ml}$ be the fitted values under the linear regression model and let $\hat{y}_{na}$ be the same for the nonparametric additive model. Then the degree to which nonparametric additive model improves upon the linear model can be measured by the difference between

R^2_{na} = square of correlation coefficient between y and $\hat{y}_{na}$

and R^2_{ml} = square of correlation coefficient between y and $\hat{y}_{ml}$

Let $df_{ml(e)}$ and $df_{na(e)}$ be the residual degrees of freedom for the multiple linear and nonparametric additive models, respectively. Under the null hypothesis that the nonparametric additive model is true,

$$F = \frac{R^2_{na} - R^2_{ml}}{(1 - R^2_{na})(df_{ml(e)} - df_{na(e)}) / df_{na(e)}}$$

approximately follows an F distribution with ($df_{ml(e)}$ - $df_{na(e)}$, $df_{na(e)}$) degrees of freedom (Hastie and Tibshirani, 1990).

Results and Discussion

The summary of the input-response data on arecanut collected from the two districts are given in Table 1. Note that there exist sufficient variations in each input variables to study its effect on yield. The functional relationship between inputs and the response variable is studied using both the multiple linear regression

Table 1. Summary of the yield and input data of Dakshina Kannada and Shimoga districts

	Dakshina Kannada			Shimoga		
Variables	Range	Mean	SD	Range	Mean	SD
N(g)	0 - 135	57	36	0 - 220	78	55
P(g)	0 - 165	65	53	0 - 210	69	64
K(g)	0 – 230	89	63	0 - 350	145	97
GL(kg)	0 - 30	13	9	0 - 22	10	7
FYM(kg)	0 - 60	19	13	0 - 63	28	17
Density (palms/ha)	1000-2500	1678	343	1000-3000	1988	438
Yield (kg/ha)	1000-3750	2153	880	1500 -4500	2795	730

technique and the nonparametric additive regression technique. The fitted models using the multiple linear regression technique with all the input variables for both the locations are given in Table 2.

Table 2. The multiple linear regression model fitted to the yield and input variables in arecanut

		Regression Coefficients						
Location	Intercept	X_1	X_2	X_3	X_4	X_5	X_6	R^2
DK	-804.83*	7.00	0.98	3.23*	17.08*	22.4**	0.93**	0.69
Shimoga	1545.0**	0.375	0.55	5.15**	8.40	-1.13	0.47*	0.57

* Significant at p=0.05 ** Significant at p=0.01

The results of the multiple linear regression analysis (Table 2) shows that the fitted model could explain about 69 and 57 % of variations in arecanut yield in Dakshina Kannada and Shimoga districts, respectively. In Dakshina Kannada district, the regression coefficients of X_3 (K), X_4 (GL), X_5 (FYM) and X_6 (density) are significant, whereas, in Shimoga district only X_3 (K) and X_6 (density) are significant. In linear regression we assume that the predictor variables are linearly related to the response variable and the significance test will verify only the linear relationship. In many situations the functional relationship between the response and the predictor variables may not be known in advance. In such cases the more flexible nonparametric regression technique which assumes only a smooth relationship between the response and predictor variables is more useful. The graphical representation of the nonparametric additive model fit provides the mean response for the given range of input values and it is easy to find out the optimum value of the input variables.

The comparison of the linear against the nonparametric fit is given in Table 3. The square of the correlation coefficient of the fitted values against the observed values in the nonparametric regression technique (R^2_{na}) in both the districts (Fig. 1 and Fig. 2) are much higher than that of the multiple linear regression technique (R^2_{ml}). The nonparametric additive model could explain about 88 and 80 % of variations in arecanut yield in Dakshina Kannada and Shimoga districts, respectively. The F-test also shows that the nonparametric fit is significantly better than the linear regression fit (Table 3).

The graphical representation of the nonparametric additive model fit of the yield and input data of arecanut in Dakshina Kannada and Shimoga districts are given in Fig. 3 and Fig. 4, respectively. The smooth function f(N)

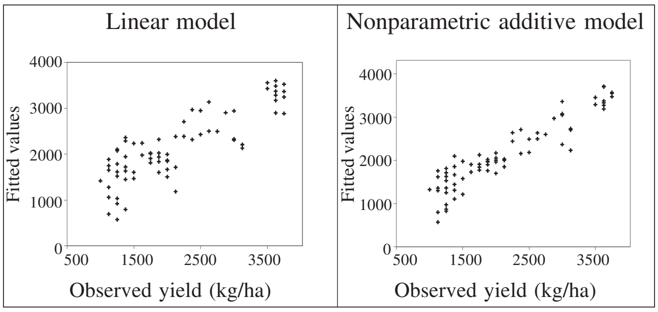

Fig.1. Fitted Vs observed values of the linear and nonparametric additive models of the yield and input data of arecanut in Dakshina Kannada district

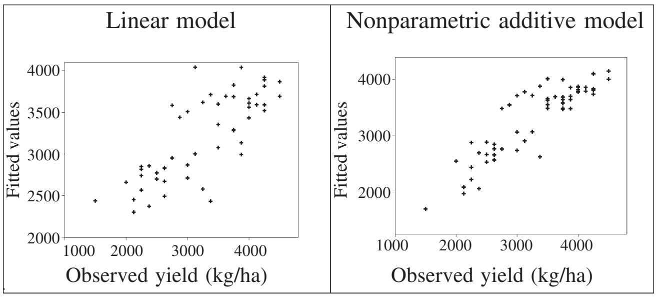

Fig.2. Fitted Vs observed values of the linear and nonparametric additive models of the yield and input data of arecanut in Shimoga district

Table 3. Comparison of Nonparametric Vs Linear fit

Location	R^2_{ml}	R^2_{na}	$df_{ml(e)}$	$df_{na(e)}$	F
DK	0.69	0.88	67	56	8.06**
Shimoga	0.57	0.80	53	42	4.39**

** Significant at p=0.01

corresponding to the response to nitrogen of the nonparametric additive model fit of DK district (Fig. 3) shows that the response of nitrogen to the arecanut yield is maximum at about 100 g which is same as the recommended dose. The analysis shows that the response of phosphorus on arecanut yield is not significant. Note that the small dip in values for low levels of N and P may be due to the random error. The response of potassium is maximum at 100 g and higher doses of potassium did not give any additional advantage. The fitted functions show that the yield is increasing up to 25 kg of green leaf and 40 kg of FYM application. Even though the yield is increasing up to 2500 palms per hectare, the rate of increase is only marginal after 1750 palms per hectare.

The smooth function f(N) corresponding to the response to nitrogen of the nonparametric additive model fit of Shimoga district shows that the mean response of nitrogen to the arecanut yield is maximum at about 50 g and higher doses of nitrogen did not give any advantage (Fig.4). The response of phosphorus on arecanut yield is not significant. The fitted model shows that eventhough the yield is increasing up to 350 g of potassium, the rate of increase after 150 g is only marginal. Note that the yield is not showing any improvement for the application of FYM or the green leaf in Shimoga district. Even though the yield is increasing upto 2250 palms per hectare, the rate of increase is only marginal after 1500 palms per hectare.

values of the component functions provide the mean response of input variables on the yield of arecanut. The optimum value of the input variables can be obtained from the graphical representation of the component functions. The present analysis of data from the two districts shows that input response relationship vary depending on the agro-climatic conditions of the locations. In Dakshina Kannada district, the input

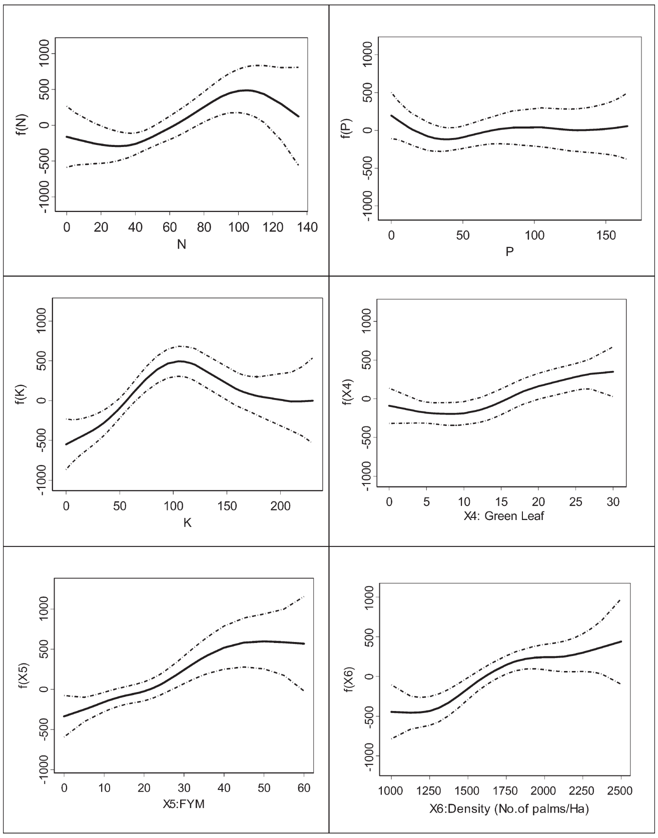

Fig. 3. The expected values (solid line) and the confidence interval (dotted line) of the nonparametric additive model fit of the yield and input data of arecanut in Dakshina Kannada district

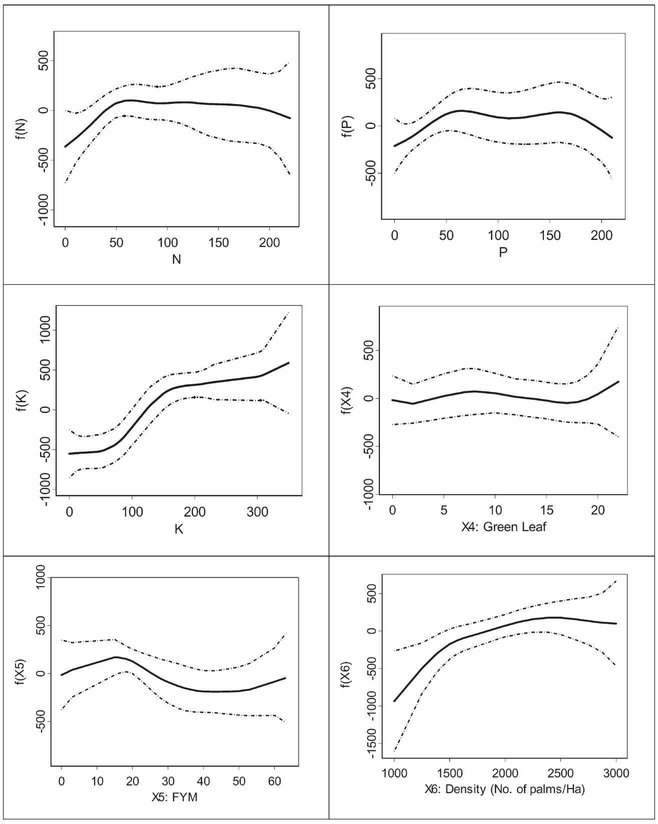

Fig. 4. The expected values (solid line) and the confidence interval (dotted line) of the nonparametric additive model fit of the yield and input data of arecanut in Shimoga district.

The study indicates that the recommendation should be made based on the local conditions. The response will vary from location to location. The application rate can be reduced in fertile soils like Shimoga for getting optimum yield. However, further study combined with soil analysis data is required for giving specific recommendations.

Conclusion

The input response relationship in arecanut is studied using nonparametric additive modeling technique. The comparative study shows that the nonparametric additive modeling technique performed much better than the multiple linear regression technique to explain the input response relationship. The estimated variables such as N, K, Green Leaf and FYM and density have significant effect on yield, whereas, in Shimoga only N, K and density have shown significant effect on yield. The study reveals that the recommendation should be made based on the local conditions. The application rate can be reduced in fertile soils like Shimoga for getting optimum yield.

References

Buja, A. Hastie, T.J and Tibshirani, R.J. 1989. Linear smoothers and additive models. *Ann.Statist.* **17**: 453-555.

Chambers, J.M and Hastie, T.J. 1992. *Statistical models in S*. Wadsworth and Brooks/Cole, Pacific Grove, CA. 624 p.

Hastie, T.J. and Tibshirani, R.J. 1990. *Generalized Additive Models*. Chapman and Hall, Washington, DC. 352 p.