Research Article *Journal of Plantation Crops, 2008, 36 (3): 276-280*

Input-response analysis in arecanut - a semiparametric regression approach

C.T. Jose[1], C. Thamban* and Ravi Bhat

Central Plantation Crops Research Institute, Regional Station, Vittal-574 243, Karnataka
** Central Plantation Crops Research Institute, Kasaragod- 671 121, Kerala*

Abstract

The effect of various input factors on the yield of arecanut has been studied based on the field data collected from farmers' gardens in Kasaragod district. Semiparametric regression technique is used to explain input-response relationship. In this approach, both qualitative (discrete) and quantitative (continuous) input variables can be included. The quantitative variables such as fertilizer application (N, P and K), Farm Yard Manure (FYM), Green leaf (GL) application and density (number of palms/ha) and the qualitative variables such as variety (SK local, Mohitnagar, Mangala), type of irrigation (hose, sprinkler) and cropping models (mono, arecanut+banana, arecanut+cocoa, arecanut+banana+pepper) were considered as input variables. The average yield (kg/ha) was taken as the response variable. The effect of qualitative variables on yield and the optimum input levels of the quantitative variables were obtained. The estimated optimum levels of N, P and K per palm in the study area are 100g, 50g and 120g, respectively, which are very close to the recommended doses. The yield differences in arecanut under monocrop and mixed cropping systems were not significant and therefore, it is better to maintain arecanut gardens under mixed cropping system for better return. The yield variations among three varieties were found to be not significant once the effects of other variables are eliminated. No significant difference between the two methods of irrigation was found, though sprinkler gave comparatively more response than hose irrigation.

Keywords: Arecanut, crop production model, input-response relationship, semiparametric regression

Introduction

Arecanut is a major commercial crop in India and it grows under a variety of climatic and soil conditions. The yield of arecanut is influenced by various factors such as variety, soil and climatic conditions as well as management practices. CPCRI has released five high yielding varieties viz., Mangala, Sumangala, Sreemangala, Mohitnagar and Swarnamangala. Among these, Mangala and Mohitnagar are very popular. There are some other high yielding cultivars in different localities known by the name of the place, where they are grown. 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. Semiparametric regression technique is used to study the input-response relationship in arecanut, based on the field data collected from farmers' gardens in Kasaragod district. Kasaragod is the largest (27%) arecanut producing district in Kerala. The average altitude of the arecanut growing areas in Kasaragod district is about 100 m above the Mean Sea Level (MSL) and the temperature ranges between 15° C to 38° C. It receives about 3500 mm rainfall every year. The major arecanut varieties grown in this district are SK local, Mohitnagar and Mangala. The practice of growing various intercrops such as banana, pepper and cocoa in arecanut garden is very common. Arecanut is mainly grown as irrigated crop. Sprinkler, drip and manual/ hose type of irrigation systems are commonly practiced by the farmers. The yield variation in farmers' garden is very high mainly due to the variations in inputs.

The relationships between the inputs and response are very complex in crop production models. Problems like nonlinear relationships between inputs and response,

[1] *Corresponding author E-mail: ctjos@yahoo.com*

discrete input variables and nonexistence of proper functional form to represent relationship 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. Jose and Bhat (2008) used nonparametric additive regression model, when all the input variables are quantitative (continuous). But in many situations, discrete/qualitative input variables such as variety, soil type, type of irrigation, inter crops etc. are to be considered in crop production models. Comparison of various qualitative variables such as variety, cropping model, irrigation type etc. by taking average yield without eliminating the effect of other input variables will be misleading. In the proposed semiparametric regression model based approach, both qualitative (discrete) and quantitative (continuous) input variables can be included. The quantitative variables such as fertilizer application (N, P and K), FYM, green leaf (GL) application and density (number of palms/ha) and the qualitative variable such as variety (SK local, Mohitnagar, Mangala), type of irrigation (hose, sprinkler) and cropping models (mono, arecanut+banana, arecanut+cocoa, arecanut+banana+pepper) were considered as input variables. The average yield (kg/ha) was taken as the response variable.

Materials and Methods

Data from the randomly selected 84 gardens in Kasaragod 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), FYM, GL application, density (number of palms/ha), variety, type of irrigation and inter crops are taken as input variables and average yield (kg/ha) was taken as the response variable for the study. The quantitative variables N, P, K, GL, FYM and density are represented by X_1, X_2, X_3, X_4, X_5 and X_6. The qualitative variables variety (SK local, Mohitnagar, Mangala), type of irrigation (hose, sprinkler) and inter crops (mono, arecanut+banana, arecanut+cocoa, arecanut+banana+pepper) are represented by dummy variables with values 1 for the presence and 0 for the absence of particular character. The dummy variable $\mathbf{Z}_1 = [Z_{11}\, Z_{12}\, Z_{13}]$ represents the 3 varieties, $\mathbf{Z}_2 = [Z_{21}, Z_{22}, Z_{23}\, Z_{24}]$ represents the 4 cropping models and $\mathbf{Z}_3 = [Z_{31}\, Z_{32}]$ represents the two types of irrigation. The response variable is represented by $\mathbf{Y}$. The semiparametric regression model considered for the study is of the form

$$Y=\alpha+\sum_{i=1}^{p} f_i(X_i)+\sum_{j=1}^{3} Z_{1j}\beta_{1j}+\sum_{k=1}^{4} Z_{2k}\beta_{2k}+\sum_{l=1}^{2} Z_{3l}\beta_{3l}+\varepsilon \qquad (1)$$

where, $\mathbf{Y}$ is the response variable, α is the general mean, p is the number of quantitative variables, X_i's are the quantitative variables, Z_{ij}'s are the dummy variables having values 1 or 0 denoting the presence or absence of qualitative variables, ε is independently and identically distributed error term with mean 0 and constant variance σ^2. The functions $f_i (i = 1,\ldots, p)$ are assumed to be smooth and β_{ij}s are the regression coefficients corresponding to the qualitative variables. The model (1) can be written as

$$Y=\alpha+\sum_{i=1}^{p} f_i(X_i)+Z_1\beta_1+Z_2\beta_2+Z_3\beta_3+\varepsilon$$

The iterative (backfitting) algorithm proposed by Buja *et al.* (1989) is used to fit the model. Let S_i represent centered local linear smoother matrices corresponding to the function f_i, $i=1,\ldots,p$; $Z=[Z_1\, Z_2\, Z_3]$ and $\beta^T = [\beta_1\, \beta_2\, \beta_3]$. The estimators for the above semiparametric additive regression are defined as the solutions of the backfitting algorithm on the following set of equations.

$$\hat{\beta} = [Z^T Z]^{-1} Z^T \left[Y - \sum_{i=1}^{p} \hat{f}_i(X_i)\right]$$

$$\hat{f}_i(X_i) = S_i\left[Y - Z\hat{\beta} - \sum_{j\neq i}^{p} \hat{f}_j(X_j)\right], \qquad i = 1,\ldots,p$$

Let W_i represent the additive model smoother for the additive component function f_i which maps $\mathbf{Y}$ to $\hat{f}_i$, $W = \sum_{i=1}^{p} W_i$ and $\mathbf{f} = \sum_{i=1}^{p} f_i$. Then,

$$\hat{\alpha} = \bar{Y}$$

$$\hat{\beta} = [Z^T(I-W)Z]^{-1} Z^T(I-W)Y$$

$$\hat{f} = W(Y - Z\hat{\beta})$$

$$\hat{\sigma}^2 = \frac{1}{(n-q-1-\mathrm{trace}(W))}[Y - Z\hat{\beta} - \hat{f}]^T[Y - Z\hat{\beta} - \hat{f}]$$

where q is the total number of qualitative variables. The variance of $\hat{\beta}$ is estimated by

$$V(\hat{\beta}) = PP^T\hat{\sigma}^2$$

where, $\mathbf{P} = (Z^T(I-W)Z)^{-1}Z^T(I-W)$. The corrected yield corresponding to the qualitative variables Z_{ij} is given by $\hat{\alpha} + \hat{\beta}_{ij}$ where $\hat{\alpha}\hat{\alpha}$ and $\hat{\beta}_{ij}\hat{\beta}_{ij}$ are the estimated values of α and β_{ij}. The significance of a qualitative variable say, Z_i can be tested by taking $\beta_i = 0$ as the null hypothesis. The significance of the effect of a quantitative variable say X_i can be tested using the lack of fit statistic (Hart, 1997) by taking the null hypothesis as f_i=0. The detailed estimation and testing procedures are given by Hastie and Tibshirani (1990) and Hart (1997).

Results and Discussions

The range, mean and standard errors of the yield and the input (quantitative variables) values collected from the 84 gardens in Kasaragod district are given in Table 1. There exist sufficient variations in each input variables to study the input-response relationship. The average input values and mean yield under different varieties/ irrigation types/ inter crops are given in Table 2. The average yield of Mohitnagar (2320 kg/ha) and Mangala (2243 kg/ha) varieties are higher than the SK local variety (1914 kg/ha). Note that the average input values for Mohitnagar and Mangala varieties are much higher than that of SK local and therefore, the comparison of the performance of different varieties by taking simple averages is not correct. Semiparametric regression model based approach is used to compare performance of different varieties after eliminating the effect of other input variables.

Table 1. Summary of observed yield and input values

Variables	Range	Mean	SE
N (g/palm)	0-200	37.24	4.89
P (g/palm)	0-120	33.27	3.59
K (g/palm)	0-245	77.32	8.25
GL (Kg/palm)	0-40	11.69	0.82
FYM (Kg/palm)	0-40	15.31	1.03
Density(No. of palms/Ha)	556-3000	1430.58	49.63
Yield (Kg/Ha)	375-3667	1986.12	95.12

The graphical representation of the fitted functions with respect to the quantitative variables, N, P, K, GL, FYM and density are shown in Fig.1. The smooth function f(N) indicated that the response of nitrogen to the arecanut yield is maximum at about 100 g which is same as the recommended value. The optimum value of P is 50 g/palm which is slightly higher than the recommended value of 40 g/palm. The fitted function f(K) shows that the yield is increasing up to 120 g of potassium and it is slightly less than the recommended value of 140 g/palm. The analysis shows that the effect of green leaf is not significant and the optimum value of FYM is about 25 kg/palm. Even though the yield is increasing up to 2500 palms per hectare, the rate of increase is only marginal after 1500 palms per hectare. The recommended and the estimated optimum level of quantitative input variables are given in Table 3.

The observed and corrected average yield with standard errors (SE) with respect to the qualitative variables is given in Table 4. The corrected yield corresponding to each qualitative variable is obtained by removing the effect of other input factors from the yield. The comparison of yield based on the observed value shows that the per hectare yield of Mohitnagar is 406 kg higher than that of SK local and Mangala is 329 kg higher than SK local. It can be noted that the average input values of N, P, K, GL and FYM are much higher in Mohitnagr and Mangala gardens compared to SK local (Table 2). Therefore, comparison of average observed yield under different varieties without adjusting for the variations in input values is not correct. The comparison of corrected yield shows that there is no significant difference between the varieties and the difference shown in the observed yield is due to the variations in the input values.

The comparison of observed yield of arecanut under different cropping model shows that the yield under

Table 2. The average input values and mean yield under different qualitative variables

Type	n	N (g/palm)	P (g/palm)	K (g/palm)	GL (kg/palm)	FYM (kg/palm)	Yield (kg/ha)
Variety							
SK Local	67	29.84	28.10	71.64	11.52	14.49	1914
Mohitnagar	6	76.67	45.83	117.17	13.33	16.67	2320
Mangala	11	61.09	57.91	90.18	11.82	19.55	2243
Inter crops/crop models							
Arecanut (Mono)	21	24.05	23.29	77.52	13.71	12.90	1839
Areanut+banana	36	44.72	33.75	86.75	10.86	14.72	2056
Arecanut+banana+pepper	21	36.48	38.62	67.71	10.62	19.14	2026
Arecanut + cocoa	6	41.67	46.67	53.67	13.33	13.83	1945
Irrigation							
Hose/manual	54	39.91	34.78	83.00	12.56	13.69	1984
Sprinkler	30	32.53	30.57	67.10	10.13	18.23	1990

n: number of gardens

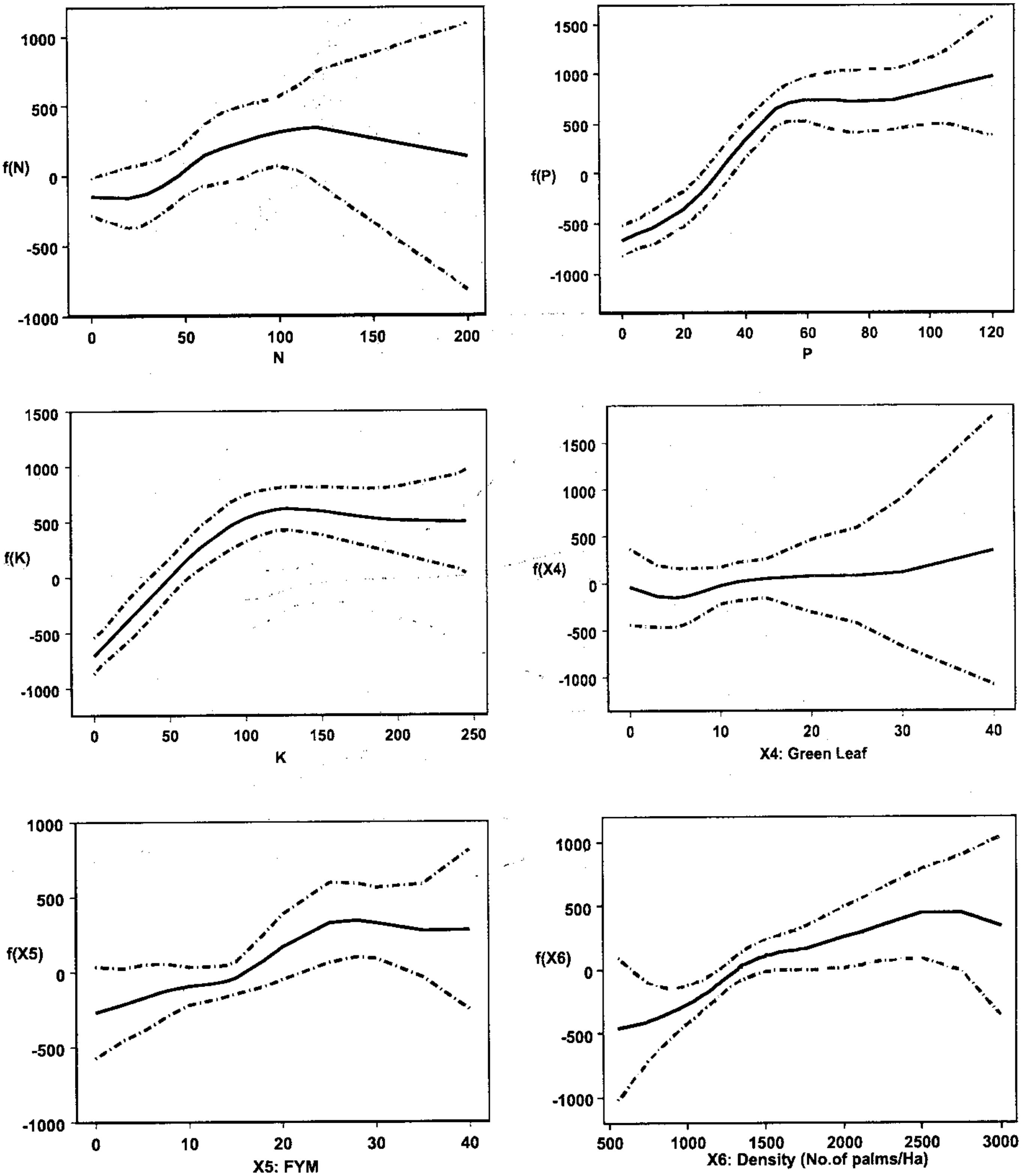

Fig. 1 Fitted functions (solid line) and confidents interval (dotted line) corresponding to the quantitative variables

Table 3. Recommended and estimated optimum level of input values

Item	N(g/ palm)	P(g/ palm)	K(g/ palm)	GL(kg/ palm)	FYM(kg/ palm)	Density(palms/ha)
Recommended	100	40	140	12	12	1370
Estimated	100	50	120	NS	25	1500

Table 4. Observed and corrected yield data with standard error

Type	n	Arecanut yield (kg/ha)		
		Observed	Corrected	SE
Varieties				
SK Local	67	1914	1985	22
Mohitnagar	6	2320	1997	75
Mangala	11	2243	1987	55
Cropping models				
Arecanut (Mono)	21	1839	2003	40
Areanut+banana	36	2056	1966	30
Arecanut+banana+pepper	21	2026	1995	40
Arecanut + cocoa	6	1945	2017	75
Irrigation types				
Hose	54	1984	1970	25
Sprinkler	30	1990	2015	33

mono cropping system is less than the other cropping systems but the corrected yield shows that there is no significant difference between different cropping models (Table 4). Therefore, the variation in observed yield under different cropping model is due to the variations in other input variables and the effect of different cropping models on the arecanut yield is insignificant. Note that there is no significant difference in yield under hose and sprinkler types of irrigation.

Conclusion

Semiparametric regression technique is proposed to study the input-response relationship in arecanut based on the field survey data collected from Kasaragod district. In this approach, both qualitative (discrete) and quantitative (continuous) input variables can be included in the crop production model. Comparison of various qualitative variables such as variety, cropping model, irrigation type etc. by taking average yield without eliminating the effect of other input variables will be misleading. The analysis of quantitative variables indicated that the estimated optimum levels of N, P and K per palm in the study area are 100 g, 50 g and 120 g respectively, which are very close to the recommended doses. The yield differences in arecanut under monocrop and mixed cropping systems were not significant and therefore, it is better to maintain arecanut gardens under mixed cropping system for better return. The yield variations among three varieties were found to be not significant once the effects of other variables are eliminated. No significant difference between the two methods of irrigation was found, though sprinkler gave comparatively more response than hose irrigation.

References

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

Hart, J.D. 1997. *Nonparametric smoothing and lack-of-fit tests*, Springler Verlag, New York, 287p.

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

Jose, C.T. and Baht, Ravi. 2008. Application of nonparametric additive model for input response analysis in arecanut. *J. Plantn. Crops* **36**(1): 44-52.