Available online at www.isas.org.in

JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 63(2) 2009 151-157

Spatial Smoothing Technique in Field Experiments

C.T. Jose[1*], Ravi Bhat[1], B. Ismail[1] and S. Jayasekhar[2]

[1]*Central Plantation Crops Research Institute, Regional Station, Vittal, Karnataka*

[2]*Mangalore University, Mangalagangotri, Karnataka*

(Received: June 2007, Revised: March 2009, Accepted: July 2009)

SUMMARY

We generally use block designs in field experiments to control the experimental error due to positional variations. The underlying assumption in classical block designs that the homogeneity of experimental area within the block may not satisfy always, particularly when the block size is large. Also we may not know in advance the soil fertility gradient and other factors influencing the response variable to divide the experimental area into homogeneous blocks. We propose spatial smoothing technique to estimate/eliminate positional effect in field experiments. We have considered a semiparametric regression model with treatment effect as the parametric component and the positional effect as the nonparametric spatial function. The only assumption about the positional effect is that it is a smooth spatial function. The proposed method is also extended for the analysis of data in the presence of sudden shifts in the spatial function (positional effect). The method is illustrated through both simulated as well as field experimental data.

Keywords: Nonparametric regression, Design of experiments, Positional effect, Semiparametric regression, Jump regression surface.

1. INTRODUCTION

Experimental error or the unexplained variation is the main concern in field experimentation technique. We generally use block designs in field experiments to control the experimental error due to positional variations. The underlying assumption in classical block designs regarding the homogeneity of experimental area within the block may not satisfy always, particularly when the block size is large. Field experiments with perennial tree crops require large experimental area, and it is grown mainly in hilly areas where getting large homogeneous area is difficult. Also we may not know in advance the soil fertility gradient and other factors influencing the response variable to divide the experimental area into homogeneous blocks. Gilmour *et al.* (1997) suggested the covariance modeling technique to tackle this problem. In the present study, nonparametric spatial modeling technique has been used to estimate/eliminate the positional effect in agricultural field experiments. The treatment effect is taken as the parametric component and the positional effect (covariate) is taken as a spatial (bivariate) nonparametric function. The only assumption about the positional effect is that it is a smooth spatial function. The field experiments with perennial or tree crops require large experimental area and it is difficult to get large homogeneous blocks to conduct experiments particularly in farmer's field. In many situations, the soil characters or the environmental variables have some sudden changes in the field or in other words, the spatial function representing the positional effect may have some jumps or discontinuities. The method is extended for the analysis of data in the presence of

Corresponding author : C.T. Jose

E-mail address : ctjos@yahoo.com

sudden jumps in the spatial function. The proposed method is applied to both the simulated as well as field experimental data to see its performance.

2. MODEL SETTINGS AND ESTIMATORS

The semiparametric regression model considered for the study is given by

$$\boldsymbol{Y} = \mu + \boldsymbol{X}\beta + f(U, V) + \varepsilon \quad (1)$$

where $Y = [y_1\ y_2 \ldots y_n]^T$ is the observation vector, μ is the general mean, $X = [x_1\ x_2 \ldots x_n]^T$ is the design matrix, $\beta = [\beta_1 \beta_2 \ldots \beta_p]^T$ is the treatment effect vector, $f(U,V)=[f(u_1,v_1)\ \ldots\ f(u_n,v_n)]^T$ is the nonparametric spatial function representing the positional effect and ε is the independently and identically distributed (iid) random error vector with mean zero. It is assumed that $f(U,V)$ is a smooth function. The backfitting algorithm of Hastie and Tibshirani (1990) is used to compute the estimates for the semiparametric regression model. The backfitting estimators for β and f are equivalent to

$$\hat{\mu} = \bar{Y}\ , \quad \hat{\beta} = (\boldsymbol{X}^T(I - S)\ \boldsymbol{X})^{-1}\ \boldsymbol{X}^T(I - S)\ (\boldsymbol{Y} - \hat{\mu})$$

$$\text{and } \hat{f} = \boldsymbol{S}(\boldsymbol{Y} - \boldsymbol{X}\hat{\beta} - \hat{\mu})$$

where, S is the smoothing matrix derived using local linear regression (Ruppert and Wand 1994). Let S_{uv} be the row of the smoother matrix correspond to the smoother vector S_{uv}^T evaluated at the observation point $(u,v) = (u_1,v_1), (u_2,v_2), \ldots, (u_n,v_n)$. Then

$$\boldsymbol{S} = [S_{u_1 v_1} \ldots S_{v_n v_n}]^T$$

where, $\boldsymbol{S}_{uv}^T = e_1^T (Z_{uv}^T W_{uv} Z_{uv})^{-1} Z_{uv}^T W_{uv}$

with $\mathbf{Z}_{\text{uv}} = \begin{bmatrix} 1 & (u_1 - u) & (v_1 - v) \\ . & . & . \\ . & . & . \\ . & . & . \\ 1 & (u_n - u) & (v_n - v) \end{bmatrix}$, $e_1^T = [1\ 0\ 0]$

and $\boldsymbol{W}_{uv} = diag\left\{ K\left[\left(\frac{u_1 - u}{h_1}\right), \left(\frac{v_1 - v}{h_2}\right)\right], \ldots, K\left[\left(\frac{u_n - u}{h_1}\right), \left(\frac{v_n - v}{h_2}\right)\right]\right\}$ for some bivariate kernel functions K and bandwidths h_1 and h_2.

Under the assumption that the treatments are allotted at random to the spatial locations, it can be shown that $\hat{\beta}$ is asymptotically unbiased and its asymptotic variance is $\sigma^2(\boldsymbol{X}^T \boldsymbol{X})^{-1}$ which is same as when the model is fully parametric (Opsomer and Ruppert 1999). An estimate of σ^2 is given by

$$\hat{\sigma}_1^2 = \frac{1}{(n - p - 1 - trace(S))}\left[\boldsymbol{Y} - \hat{\mu} - \boldsymbol{X}\hat{\beta} - \hat{f}\right]^T \times \left[\boldsymbol{Y} - \hat{\mu} - \boldsymbol{X}\hat{\beta} - \hat{f}\right]$$

The variance of $\hat{\beta}$ is estimated by

$$\hat{V}(\hat{\beta}) = PP^T \hat{\sigma}_1^2$$

where, $P = (\boldsymbol{X}^T (I - \boldsymbol{S})\ \boldsymbol{X})^{-1}\ \boldsymbol{X}^T(I - \boldsymbol{S})$. The significance of the positional effect f is tested using the lack of fit statistic by comparing parametric and nonparametric models (Hart 1997).

Under the null hypothesis that the positional effect $f(U, V) = 0$, the mean residual sum of squares obtained by fitting model (1) is given by

$$\hat{\sigma}_0^2 = (Y - \hat{\mu})^T\ [(\mathrm{I} - \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{X})^{-1}\ \boldsymbol{X}]^T \times [(\mathrm{I} - \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{X})^{-1}\ \boldsymbol{X}]\ (\boldsymbol{Y} - \hat{\mu}) / (n - p - 1)$$

The lack of fit test statistic is given by

$$R_1 = \frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}$$

The statistic R_1 asymptotically follows an F distribution with $(n - p - 1)$, $[n - p - 1 - \text{trace}(S)]$ degrees of freedom and it can be used for testing the significance of the positional effect.

Additive model for positional effect: In many situations, the number of experimental units is comparatively small and estimating the spatial function using the bivariate smoother will be inadequate. In such situations, bivariate additive model can be fitted instead of the two dimensional spatial function used in model (1). By using bivariate additive function, model (1) becomes

$$\boldsymbol{Y} = \mu + \boldsymbol{X}\beta + f_1(U) + f_2(V) + \varepsilon \quad (2)$$

where, f_1 and f_2 are the univariate nonparametric functions representing the positional effect in the U and V directions and it is assumed that $\sum f_1(u_i) = \sum f_2(v_i) = 0$.

Let $\boldsymbol{M_1}$ and $\boldsymbol{M_2}$ are the centered smoother matrices corresponding to $\boldsymbol{U}$ and $\boldsymbol{V}$. The backfitting algorithm will provide an explicit solution to the above semiparametric regression model and the estimates are given by

$$\hat{\mu} = \bar{Y}, \ \hat{\beta} = (\boldsymbol{X}^T(\boldsymbol{I}-\boldsymbol{Q})\boldsymbol{X})^{-1}\boldsymbol{X}^T(\boldsymbol{I}-\boldsymbol{Q})(\boldsymbol{Y}-\hat{\mu})$$

and

$$\hat{f} = \hat{f}_1 + \hat{f}_2 = \boldsymbol{Q}(\boldsymbol{Y}-\hat{\mu}-\boldsymbol{X}\hat{\beta})$$

The matrix $\boldsymbol{Q}$ and the estimates $\hat{f}_1$ and $\hat{f}_2$ are obtained by solving the set of equations

$$\begin{bmatrix} I & M_1 \\ M_2 & I \end{bmatrix}\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}\left(Y-\hat{\mu}-\boldsymbol{X}\hat{\beta}\right)$$

$$\hat{f}_1 = \{\boldsymbol{I}-(M_1M_2)^{-1}(1-M_1)\}\ (Y-\hat{\mu}-\boldsymbol{X}\hat{\beta})$$

$$= Q_1\left(Y-\hat{\mu}-\boldsymbol{X}\hat{\beta}\right)$$

$$\hat{f}_2 = \{\boldsymbol{I}-(M_2M_1)^{-1}(1-M_2)\}(\boldsymbol{Y}-\hat{\mu}-\boldsymbol{X}\hat{\beta})$$

$$= Q_2\left(\boldsymbol{Y}-\hat{\mu}-\boldsymbol{X}\hat{\beta}\right)$$

and $\boldsymbol{Q} = \boldsymbol{Q}_1 + \boldsymbol{Q}_2$

An estimate of σ^2 under model (2) is given by

$$\hat{\sigma}_2^2 = \frac{1}{(n-p-trace(Q))}\left[\boldsymbol{Y}-\hat{\mu}-\boldsymbol{X}\hat{\beta}-\hat{f}\right]^T \times \left[\boldsymbol{Y}-\hat{\mu}-\boldsymbol{X}\hat{\beta}-\hat{f}\right]$$

The significance of the positional effect f is tested using the lack-of-fit test statistic

$$R_2 = \frac{\hat{\sigma}_0^2}{\hat{\sigma}_2^2}$$

The test statistic R_2 asymptotically follows an F distribution with (n – p – 1), [n – p – trace (**Q**)] degrees of freedom and it can be used for testing the significance of the positional effect. An approximate α-level point wise confidence band around the estimated function f is given by $\hat{f}(u_i, v_i) \pm z_{\alpha/2}\hat{\sigma}_2\sqrt{[\boldsymbol{QQ}^T]_{ii}}$ for $i = 1,\ldots,n$, where, $[\boldsymbol{QQ}^T]_{ii}$ represents the element in the ii^{th} position of the matrix $[\boldsymbol{QQ}^T]$.

The variance of $\hat{\beta}$ is estimated by

$$V(\hat{\beta}) = \boldsymbol{PP}^T\hat{\sigma}_2^2$$

where, $\boldsymbol{P} = (\boldsymbol{X}^T(I-\boldsymbol{Q})\boldsymbol{X})^{-1}\boldsymbol{X}^T(I-\boldsymbol{Q})$

Choice of bandwidth: The procedure described above involves two smoothing parameters h_1 and h_2. The choice of bandwidth parameters is very crucial in smoothing technique. We have used the cross-validation technique (Hardle 1990) to obtain the optimum bandwidths. Let y_i, $i = 1,\ldots,n$ are the observations and $\hat{y}_{(i)h_1h_2}$ be its leave-one-out estimate (estimate without using the i^{th} observation) with h_1 and h_2 as bandwidths. Then the cross-validation score is defined by

$$CV(h_1, h_2) = \frac{1}{n}\left[\sum_{i=1}^{n}(y_i - \hat{y}_{(i)h_1, h_2})^2\right]$$

The values of h_1 and h_2 which minimize $CV(h_1, h_2)$ can be used as the bandwidths for estimating the regression model.

3. JUMPS IN THE SPATIAL FUNCTION

Sometimes the soil characters or environmental variables have some sudden changes in the field or in other words, the spatial function representing the positional effect has some jumps or discontinuities. In such situations the procedure given in Section 2 needs to be modified. Let us first define a jump in the spatial function f as follows:

$$f(u, v) = g(u, v) + \Delta(u)\, I_{v>c(u)},\ (u, v) \in [0, 1]^2 \quad (3)$$

where, $g(u, v)$ is the continuous part, $c(u)$ denotes the jump location curve and $\Delta(u)$, is the jump magnitude function. The functions g and Δ are assumed to be smooth. The jump location curve $c(u)$ is assumed to have first order derivative. Note that the jump location curve $c(u)$ divides the entire experimental area into two parts. Under the assumption that the treatments are randomly distributed to the entire experimental area, initial estimates for μ and β are given by

$$\hat{\mu} = \bar{Y} \qquad \hat{\beta} = \left[\boldsymbol{X}^T(I-\boldsymbol{S})\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T(1-\boldsymbol{S})\left(\boldsymbol{Y}-\hat{\mu}\right)$$

Let $\boldsymbol{Y}^* = \boldsymbol{Y} - \hat{\mu} - \boldsymbol{X}\hat{\beta}$

We have used the method of Jose and Ismail (2001) to estimate the jump location curve. Define the set $Q_i(u, v)$, $i = 1, ..., 4$ as the set of points in the i^{th} quadrant with respect to the point (u, v). At any point (u, v), consider the following two kernel weighted least squares (minimization) problem:

Minimize

$$\sum_{i=1}^{n}\left\{y_i^{\#} - b_0 - b_1(u-u_i) - b_2(v-v_i) - a_0(u,v)I\left[(u_i,v_i)\in Q_1(u,v)\right]\right\}^2 I\left[(u_i,v_i)\in Q_1(u,v)\cup Q_3(u,v)\right]K_i \quad (4)$$

Minimize

$$\sum_{i=1}^{n}\left\{y_i^{\#} - b_0 - b_1(u-u_i) - b_2(v-v_i) - a_0(u,v)I\left[(u_i,v_i)\in Q_2(u,v)\right]\right\}^2 I\left[(u_i,v_i)\in Q_2(u,v)\cup Q_4(u,v)\right]K_i \quad (5)$$

where, $K_l = K\left[\left(\frac{u-u_l}{h_1}\right), \left(\frac{v-v_l}{h_2}\right)\right]$ is some bivariate kernel function.

If the slope of the jump location curve at any $(u, v) \in c$ is negative, then for small bandwidths h_1 and h_2, the points in $Q_1(u, v)$ and $Q_3(u, v)$ will be in the opposite sides of c. Similarly, if the slope of c at (u, v) is positive, the points in $Q_2(u, v)$ and $Q_4(u, v)$ will be in the opposite sides of $c(u)$. The estimates of $a_0(u, v)$ obtained by solving the least squares problems (4) and (5) corresponding to the point (u, v) are denoted by $\hat{a}_{01}(u, v)$ and $\hat{a}_{02}(u, v)$ respectively. Among these two estimates, the estimate with maximum absolute value is denoted by $\hat{a}_0(u, v)$. Then the estimate of the jump location curve is given by

$$\hat{c}(u) = \underset{v\in[h_2,\,1-h_2]}{\arg\max}\left|\hat{a}_0(u,v)\right|$$

and $\hat{a}_0(u,\hat{c}(u))$ is the estimate of the jump size function $\Delta(u)$ which divides the experimental area into two disjoint sets, say A and B. The estimate $\hat{f}$ of the spatial function f on both sides of $\hat{c}(u)$ can be estimated separately based on the observations on either sides of $\hat{c}(u)$ by the method of kernel weighted local linear regression (Ruppert and Wand 1994).

Let Y_A, Y_B; X_A, X_B; f_A, f_B and S_A, S_B are the observation vectors, design matrices, positional effect vectors and the smoother matrices correspond to the sets A and B respectively. The final estimate of the treatment vector and the rearranged spatial function f^* are given by

$$\hat{\beta}^* = [\boldsymbol{X}^{*T}(I - \boldsymbol{S}^*)\boldsymbol{X}^*]^{-1}\boldsymbol{X}^{*T}(I - \boldsymbol{S}^*)(\boldsymbol{Y}^* - \mu^*)$$

$$\hat{\boldsymbol{f}}^* = \boldsymbol{S}^*\left[\boldsymbol{Y}^* - \mu^* - \boldsymbol{X}^*\hat{\beta}^*\right]$$

where $\boldsymbol{S}^* = \begin{bmatrix} S_A & 0 \\ 0 & S_B \end{bmatrix}$, $\boldsymbol{Y}^* = \begin{bmatrix} Y_A \\ Y_B \end{bmatrix}$, $\mu^* = \begin{bmatrix} \bar{Y}_A \\ \bar{Y}_B \end{bmatrix}$,

$$\boldsymbol{X}^* = \begin{bmatrix} X_A \\ X_B \end{bmatrix}, \boldsymbol{f}^* = \begin{bmatrix} f_A \\ f_B \end{bmatrix}$$

An estimate of the error variance σ^2 is given by

$$\hat{\sigma}^{*2} = \frac{1}{(n-p-1-trace(S^*))}[\boldsymbol{Y}^* - \mu^* - \boldsymbol{X}^*\hat{\beta} - \hat{f}^*]^T \times [\boldsymbol{Y}^* - \mu^* - \boldsymbol{X}^*\hat{\beta} - \hat{f}^*]$$

The variance of $\hat{\beta}$ is estimated by

$$V(\hat{\beta}^*) = \boldsymbol{P}^*\boldsymbol{P}^{*T}\hat{\sigma}^{*2}$$

where $P^* = [\boldsymbol{X}^{*T}(I - \boldsymbol{S}^*)\boldsymbol{X}^*]^{-1}\boldsymbol{X}^{*T}(\boldsymbol{I} - \boldsymbol{S}^*)$

The above method can be extended to a more general case that the jump location curve does not have the explicit functional form given in (3). Assume that the jump location curve $c(\,)$ induces a partition of the field into disjoint subsets A and B. Then the spatial function f can be defined as

$$f(u, v) = g(u, v) + \Delta(u, v)\, I_B(u, v),\ (u, v) \in [0,1]^2$$

where, g and Δ are smooth functions and $\inf|\Delta(u, v)| > 0$ for all $(u, v) \in c$. As discussed above obtain $\hat{a}_0(u, v)$ for all $(u, v) \in (h_1, 1 - h_1) \times (h_2, 1 - h_2)$. Note that $|\hat{a}_0(u, v)|$ near c are significantly larger than the others. An estimate of c can be constructed by the maximin method suggested by Muller and Song (1994). Find the curve that maximizes the minimum of $|\hat{a}_0(u, v)|$ along curves in Γ; that is,

$$\hat{c} = \arg\max_{\phi \in \Gamma} \left[\min_{(u,v)\in\phi} \left| \hat{a}_0(u,v) \right| \right]$$

where, Γ is a sufficiently rich class of candidate boundaries, containing c. Once the jump location curve is estimated, the positional effect on both sides of the estimated jump location curve can be obtained separately.

4. SIMULATION STUDY

A simulation study is carried out to see the performance of the proposed method. We considered the following model for the simulation study

$$\boldsymbol{Y} = \boldsymbol{X}\beta + f(\boldsymbol{U}, \boldsymbol{V}) + \varepsilon \qquad (6)$$

where Y is the n × 1 observation vector, X is the n × n design matrix, β is the k × 1 treatment effect vector which is taken as $\beta' = [-2\ -2\ 0\ 4]$, $f(u,v)=2\{2+sin[2(u+v)]\}$ and the random error vector ε follows $N(0, \sigma^2)$. The spatial locations of the n observations are obtained by dividing the region [0, 1] × [0, 1] equally and each treatment is allotted randomly to n/k spatial locations. Based on the above, 100 sets of data are simulated for different values of n (100, 400, 900) and σ (0.5, 1.0). The bivariate kernel function considered is $K(u, v)=0.75^2(1 - u^2)(1 - v^2)$ which is the product of two Epanechnikov kernels. The treatment effect vector $\beta^T = [\beta_1\ \beta_2\ \beta_3\ \beta_4]$, the bivariate function f and the error variance σ^2 are estimated using the method given in Section 2. The Average Mean Squared Errors (AMSE) of the estimated values of σ, β and f with the true values of 100 sets of simulated data for different values of n (100, 400, 900) and σ (0.5, 1.0) are given in Table 1. The AMSE of the estimated parameters are calculated as follows:

$$\text{AMSE of, } \hat{\sigma} = \frac{1}{100}\sum_{i=1}^{100}(\sigma - \hat{\sigma}_{(1)})^2$$

$$\text{AMSE of } \hat{\beta}_j = \frac{1}{100}\sum_{i=1}^{100}(\beta_j - \hat{\beta}_{j(i)})^2 \text{, j =1,…, 4}$$

$$\text{AMSE of } \hat{f} = \frac{1}{100}\sum_{i=1}^{100}\frac{1}{n}$$

$$\sum_{j=1}^{n}[f(u_j, v_j) - f_{(i)}(u_j, v_j)]^2$$

where, $\hat{\sigma}_{(i)}, \hat{\beta}_{j(i)}$ and $\hat{f}_{(i)}(u_j, v_j)$ are the estimated values of σ, β_j and $f(u_j, v_j)$ corresponding to the i^{th} simulated data set. It can be observed that the AMSE of the estimates are converging to zero as n increases or in other words, the estimated values are converging to the true values as n increases. This indicates the consistency of the estimates. The MSE varies with change in the choice of bandwidths. The optimum bandwidth (bandwidth corresponds to the minimum MSE) will depend on the curvature of the function. The optimum bandwidth for estimating the regression model is obtained based on the cross validation technique given in Section 2.

Table 1. Average Mean Squared Errors (AMSE) of the estimated values with the true values of the simulated data (Model 6)

σ	n	MSE of the estimates multiplied by 100						
		$X\hat{\beta}+\hat{f}$	$\hat{\sigma}$	$\hat{f}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
0.5	100	3.90	0.18	3.51	0.99	0.95	0.65	0.93
	400	1.42	0.03	1.27	0.21	0.15	0.12	0.14
	900	0.82	0.01	0.73	0.08	0.07	0.08	0.08
1.0	100	9.57	0.28	6.28	1.55	1.50	1.34	1.42
	400	3.37	0.12	2.95	0.45	0.41	0.38	0.43
	900	1.34	0.05	1.16	0.17	0.20	0.22	0.19

The performance of the proposed method in the case of sudden shift or jump in the spatial function is illustrated through a simulation study. For this, the regression model (6) is modified as

$$\boldsymbol{Y} = \boldsymbol{X}\beta + f_1(\boldsymbol{U}, \boldsymbol{V}) + \varepsilon \qquad (7)$$

where Y, X, β and ε are as defined in model (6). The bivariate regression function $f_1(u, v)$ is taken as a jump regression surface of the following form

$$f_1(u, v) = 2\{2 + \sin[2(u + v)]\} + [1 + 2\sin(1 + 2u)]I_{ve\ 0.6\ \sin(1+2u)} \quad (u, v) \in [0,1]^2$$

Based on the above, one set of data is simulated for $n = 900$ and $\sigma = 0.40$, the treatment vector $\beta' = [-2.0\ -2.0\ 0.0\ 4.0]$, the spatial function $f_1(u, v)$ the jump location function $c(u) = 0.6\sin(1 + 2u)$ and the jump magnitude function $\Delta(u) = 1 + 2\sin(1 + 2u)$. The treatment effect vector β, the error variance σ^2, the jump location curve $c(u)$ and jump magnitude function $\Delta(u)$ are estimated using the method given in Section 3. The estimated values of β and σ are respectively $\hat{\beta} = [-1.95\ -1.97\ -0.06\ 3.99]$ and $\hat{\sigma} = 0.44$ which are very close to the true values. The jump location function and jump magnitude function are obtained by smoothing the point wise estimates of the jump location curve and jump size function. The estimated and true values of the jump location curve and jump magnitude function are shown in Fig. 1 and 2 respectively. It can be noted that the estimated and true values are very close.

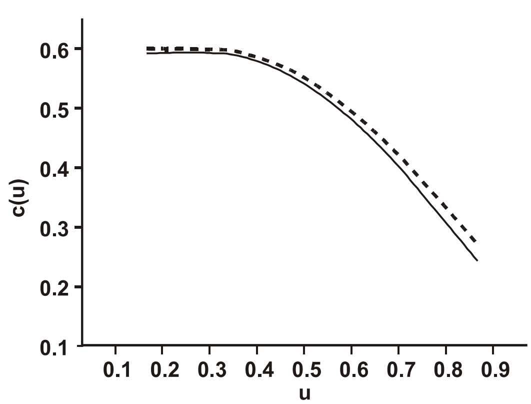

Fig. 1. Estimated (dotted line) and true values (solid line) of the jump location function

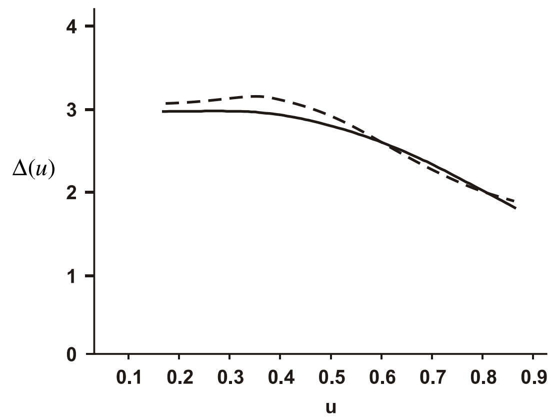

Fig. 2. Estimated (dotted line) and true values (solid line) of the jump magnitude function

5. FIELD APPLICATION

The proposed spatial technique is applied to the data of irrigation cum fertilizer trial of cocoa + areca mixed cropping system at CPCRI Regional Station, Vittal and it has been compared with the traditional method of eliminating the positional effect by blocking the experimental area. The experiment was laid out in randomized block design with 12 treatment combinations, 4 replications and 6 trees per plot. The

Table 2. Estimated parameters with standard errors of the field experiment

	Proposed Method		Method of blocking	
Treatments	μ+β	SE	μ + β	SE
1	7.57	1.15	6.18	1.33
2	11.80	1.14	11.47	1.33
3	7.26	1.14	6.38	1.33
4	8.82	1.15	7.94	1.33
5	11.97	1.12	11.91	1.33
6	9.45	1.13	8.73	1.33
7	13.79	1.12	14.45	1.33
8	12.12	1.13	12.16	1.33
9	11.38	1.15	11.84	1.33
10	14.08	1.14	14.84	1.33
11	14.10	1.14	15.38	1.33
12	14.97	1.14	16.03	1.33
MSE	32.02		42.22	

Note: μ+β is the sum of the estimated values of general mean and treatment effect after eliminating the positional/block effect main objective of the experiment is to compare the effect of different treatments on the yield of cocoa. Four years cumulative yield data has been taken as the study variable. A total of 288 experimental cocoa trees were planted at a spacing of 2.7m × 5.4m. Estimated parameters (general mean + treatment effect) with standard errors and the mean squared errors (MSE) of cumulative yield data of cocoa after eliminating the positional/block effect through both the methods are given in Table 2. There is a significant reduction in the MSE of the proposed method than the traditional method for comparing the treatment effect. We have used MATLAB package to develop programmes for the simulation study and the data analysis.

6. CONCLUSION

We generally use block designs to eliminate positional effect in field experiments. In many situations, the underlying assumption of homogeneity with in the block may not be true. In the present study, a method is proposed to eliminate the positional effect nonparametrically and the only assumption about the positional effect is that it is a smooth spatial (bivariate) function. The method is also extended for the analysis of data in the presence of sudden shifts in the spatial function (positional effect). The proposed method is useful when there is no advance information about the field conditions to divide the experimental area into homogeneous blocks.

REFERENCES

Gilmour, A.R., Cullis, B.R. and Verbyla, A.P. (1997). Accounting for natural and extraneous variation in analysis of field experiments. *J. Ag. Biol. Environ. Stat.,* **2**, 269-273.

Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.

Hart, J.D. (1997). *Nonparametric Smoothing and Lack-of-fit Tests*. Springer Verlag, New York.

Hastie,T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models*. Chapman and Hall, London.

Jose, C.T. and Ismail, B. (2001). Nonparametric inference on jump regression surface. *J. Nonparametric Stat.*, **13**, 791-813.

Muller, H.G. and Song, K. (1994). Maximin estimation of multidimensional boundaries. *J. Mult. Anal.*, **50**, 265-281.

Opsomer, J.D. and Ruppert, D. (1999). A root-n consistent estimator for semiparametric additive modeling. *J. Comput. Graph. Stat.*, **8**, 715-732.

Ruppert, D. and Wand, M.P. (1994). Multivariate locally weighted least squares regression. *Ann. Statist.*, **22**, 1346-1370.