Methodology Article

doi:10.25081/jpc.2021.v49.i3.7455
Journal of Plantation Crops, 2021, 49(3): 214-221

Robust non-parametric spatial regression and its application in field data analysis

C.T. Jose*, K.P. Chandran[1], K. Muralidharan[1], D. Jaganathan[2] and S. Jayasekhar[1]

ICAR-Central Plantation Crops Research Institute, Regional Station, Vittal, Karnataka, India
[1]ICAR-Central Plantation Crops Research Institute, Kasaragod-671 121, Kerala, India
[2]ICAR- Central Tuber Crops Research Institute, Thiruvananthapuram-695 017, Kerala, India

(Manuscript Received: 23-01-2021, Revised: 16-10-2021, Accepted: 21-10-2021)

Abstract

Outlier detection and robust estimation are an integral part of data mining and has attracted much attention recently. Generally, the data contain abnormal or extreme values either due to the characteristics of the individual or due to errors in tabulation/data entry. The presence of outliers will severely affect the data modelling and analysis. A robust nonparametric method is proposed to fit the spatial/surface regression that is not influenced by the presence of outliers in the data. Robust M-kernel weighted local linear regression smoother was used to fit the spatial regression function. The proposed method is useful to estimate/eliminate the spatial effect and identify the high potential trees in an orchard, which is useful for breeding programs. The method is illustrated through simulated data. The comparison of AMSE corresponding to the optimum bandwidth shows that the non-robust Kernel Weighted Local Regression Estimator (KWLRE) performs very badly in the presence of outliers. Among the robust estimators, the robust spatial smoother with biweight robustness weight function performed better than the Huber and Hampel weight functions. Comparison of AMSE corresponding to the optimum bandwidth showed that there is not much difference between different types of robustness weight function in the absence of outliers. In the case of robust spatial smoother with biweight robustness weight function, the AMSE for 0 per cent, 4 per cent and 8 per cent outliers are almost the same, indicating that the method is robust against the outliers. The method was also applied to the annual yield data of 225 coconut palms in a field to eliminate spatial effect and to identify the high potential trees. It was found that by removing spatial effects and outliers, the MSE has reduced more than 50 per cent.

Keywords: Nonparametric regression, M-estimates, outliers, robust inference, spatial techniques

Introduction

The fundamental objective of statistical data analysis is to systematically obtain data and make inferences or appropriate decisions based on the data. The presence of outliers or extreme values in the experimental data is a major concern for data analysis. Outlier is an observation that appears to be inconsistent with the remainder of the observations in the data set. Experimental data in agriculture may contain abnormal or extreme values due to various reasons such as genetic variations (super trees/very low yielders), loss of yield due to pest/disease infestation, errors in tabulation, data entry *etc*. These extreme values or outliers will usually increase the experimental error in data analysis. Detection of outliers and the possible remedies are important in data analysis. These outliers are a nuisance for the data analysts, but when the observation is genuine and represents high potential/super trees, it can be used as mother trees/ palms in breeding programs. Identification of extreme values/high yielders is very important in agriculture. In addition to the genetic variations, the data may also be influenced by the spatial or environmental effect. To identify the high potential/ super trees, it is important to eliminate the spatial or environmental effect from the data. In the present study, we compared the performance of estimating the spatial function and identifying outliers using non-robust Kernel Weighted Local Regression

*Corresponding Author: ctjose@gmail.com; ctjos@yahoo.com

Estimator (KWLRE) and the robust M-Kernel Weighted Local Regression Estimator using biweight, Huber and Hampel weight functions. A robust method is proposed to estimate the spatial effect and identify outliers or the high potential trees from an orchard.

Materials and methods

This paper considers spatially distributed data (y) with spatial coordinates (u, v).

Model settings and estimators

Let $(u_1, v_1, y_1),..., (u_n, v_n, y_n)$ be a set of n spatial data with (u_i, v_i) as the i^{th} spatial location and y_i is the corresponding value of the response variable Y. The nonparametric spatial regression model considered for the study is of the form

$$y_i = m(u_i, v_i) + \varepsilon_i \qquad (1)$$

where, $\boldsymbol{Y}=[y_1\ y_2 ... y_n]^T$ is the observation vector, $\boldsymbol{m}=[m(u_1, v_1)...\boldsymbol{m}=[m(u_n, v_n)$ is the nonparametric spatial regression function and $\boldsymbol{e}$ is the independently and identically distributed (iid) random error vector with mean zero. It is assumed that $\boldsymbol{m}$ is a smooth function. The regression estimators considered are based on the local least-squares fitting of kernel weighted linear regression function (Ruppert and Wand, 1994). The KWLRE of *m(u,v)*is the solution of α_0 to the following weighted least squares problem

$$\text{Minimize} \sum_{i=1}^{n} [y_i - \alpha_0 - \alpha_1(u_i - u) - \alpha_2(v_i - v)]^2 K_{Hi}(u,v)$$

where $K_{Hi}(u,\ v)=K\left[\frac{(u_i - u)}{h_1}, \frac{(v_i - v)}{h_2}\right]$ is some bivariate kernel function with h_1 and h_2 are the bandwidths in $\underline{u}$ and $\underline{v}$ directions.

$$\hat{m}(u,v) = \hat{\alpha}_0$$

The estimate of the spatial regression function **m** is given by

$$\hat{\mathbf{m}} = \boldsymbol{SY}$$

Where, $\boldsymbol{S}$ is the smoothing matrix derived using local linear regression and S_{uv} be the row of the smoother matrix correspond to the smoother vector S^T_{uv} evaluated at the observation point $(\underline{u}, \underline{v})=(u_1, v_1)\ (u_2, v_2)..., (u_n, v_n)$ Then,

$$\boldsymbol{S} = \left[S_{u_1 v_1} \cdots S_{u_n v_n}\right]^T$$

where, $S^T_{uv} = e_1^T (Z^T_{uv} W_{uv} Z_{uv})^{-1} Z^T_{uv} W_{uv}$

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

and $W_{uv} = \text{dig}\left\{K\left[\left(\frac{u_1 - u}{h_1}\right), \left(\frac{v_1 - v}{h_2}\right)\right], ..., K\left[\left(\frac{u_n - u}{h_1}\right), \left(\frac{v_n - v}{h_2}\right)\right]\right\}$

The properties of the estimator are provided in (Ruppert and Wand, 1994; Jose and Ismail, 2001). The cross-validation (leave-one-out) technique is generally used to estimate the optimum bandwidths h_1 and h_2. The cross-validation score is given by

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

Where, $\hat{m}_{h_1, h_2(-i)}(u_i, v_i)$ is the nonparametric estimate of $m(u_i, v_i)$ without using the i^{th} observation and with bandwidths h_1 and h_2. The value of h_1 and h_2, which minimizes the cross-validation score $CV(h_1,h_2)$, will be the optimum bandwidths. The estimate of the regression function *m(u,v)* using the optimum bandwidth is denoted as $\hat{m}(u,v)$.

Cleveland and Devlin (1988) and Hastie and Tibshirani (1990) discussed the estimation of error variance in linear regression smoothers. An approximate estimate of the error variance is given by

$$\hat{\sigma}^2 = \frac{1}{\left[n - 2trace(S) + trace(S^T S)\right]} \sum_{i=1}^{n} \left[y_i - \hat{m}(u_i, v_i)\right]^2$$

The nonparametric regression estimates and cross-validation technique can behave very badly in the presence of outliers in the data or when the

errors are heavy-tailed (Leung D, 2005). One remedy is to remove the influential observations from the data. Another approach is to use robust smoother, which is not as vulnerable as the usual smoothing technique. A robust M-type estimate $\hat{m}$ of the regression function $m()$ can be obtained by minimizing the objective function

$$\sum_{i=1}^{n} \rho\left[\frac{y_i - \hat{m}_{h_1,h_2}(u_i, v_i)}{s}\right] \qquad (2)$$

where, $\rho(.)$ is an even function with bounded first derivative and a unique minimum at zero. The derivative $\psi(x) = \frac{d\rho(x)}{dx}$ is called the influence function and $w(x) = \frac{\psi(x)}{x}$ is the corresponding weight function. Several M-type estimators have been discussed in the literature using diûerent types of inûuence functions (Huber, 1981; Rey, 1983; Hampel *et al*., 1986; Tukey, 1977). The three most used M-type estimators, along with their influence and weight functions, are presented in Table 1.

The iterated reweighted least-squares technique is used to solve the minimization problem (2) to obtain the robust estimate of the regression function m. The estimate of the regression function in the k^{th} iteration is denoted by $\hat{m}_{(k)}(u,v)$, which is the solution of $\alpha_{0(k)}$ to the following least-squares problem.

Minimize

$$\sum_{i=1}^{n}\left[y_i - \alpha_{0(k)} - \alpha_{1(k)}(u_i - u) - \alpha_{2(k)}(v_i - v)\right]^2 K_{Hi}(u,v) w(r_{(k-1)i})$$

Where, $r_{(k-1)i} = \frac{\left[y_i - \hat{m}_{(k-1)}(u_i, v_i)\right]}{s_{(k-1)}}$ is the standardized residual of the i^{th} datum in the $(k\text{-}1)^{th}$ iteration. Median of absolute deviation from the median (MAD) is used as a robust estimate for the scale factor s and $r_{(0)i}=0$ for i=1,…,n.

$$s_{(k-1)} = \frac{median_i\left|e_{(k-1)i} - median_j(e_{(k-1)j})\right|}{0.6745}$$

where, $e_{(k-1)i} = y_i - \hat{m}_{(k-1)}(u_i, v_i)$

The estimate of the spatial regression function m in the k^{th} iteration can be written as

$$\hat{m}_{(K)} = S_{(k)}Y$$

Where, $S_{(k)}$ is the smoothing matrix derived from the robust locally weighted linear regression and $S_{(k)uv}$ be the row of the smoother matrix correspond to the smoother vector $S^T_{(k)uv}$ evaluated at the observation point $(u,v)=(u_1,v_1), (u_2,v_2),\ldots, (u_n,v_n)$ in the k^{th} iteration.

where, $\boldsymbol{S}_{(k)} = \left[S_{(k)u_1v_1} \ldots S_{(k)u_nv_n}\right]^T$

$$S^T_{(k)uv} = e_1^T\left(Z^T_{uv}W^*_{(k)uv}Z_{uv}\right)^{-1}Z^T_{uv}W^*_{(k)uv}$$

$$W^*_{(k)uv} = diag\left\{w^*\left(r_{(k-1)1}\right),\ldots,w^*\left(r_{(k-1)n}\right)\right\}$$

$$w^*(r_{(k-1)i}) = \frac{K\left[\left(\frac{u_i - u}{h_1}\right),\left(\frac{v_i - v}{h_2}\right)\right]w\left(r_{(k-1)i}\right)}{\sum_{i=1}^{n} K\left[\left(\frac{u_i - u}{h_1}\right),\left(\frac{v_i - v}{h_2}\right)\right]w\left(r_{(k-1)i}\right)}, \quad i = 1,\ldots,n$$

Continue the iteration till there is no significant improvement in the estimated values. Here $w(r_{(k-1)i})$ is the robustness weight function corresponding to y_i in the k^{th} iteration. Let w_1 be the final robustness weight assigned to y_i and $\hat{m}_{h_1h_2}(u_i, v_i)$ be the estimated value of $m(u_i, v_i)$ with bandwidth (h_1,h_2). The mean squared errors (*MSE*) of the estimated value with the true values corresponding to the bandwidth (h_1, h_2) is given by

$$MSE(h_1,h_2) = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{m}_{h_1h_2}(u_i,v_i) - m(u_i,v_i)\right]^2$$

The cross-validation score $CV(h_1, h_2)$ does not work well for the robust smoothers because the *CV* function itself will be strongly influenced by the outliers (Wang and Scott, 1994). The cross-validation score is the sum of squares of the prediction errors of the smoother at each of the design points. When there are outliers, the prediction errors corresponding to the outliers will be uncharacteristically extreme, and these extreme prediction errors will affect the performance of *CV* (h_1, h_2). Therefore, the extreme prediction errors should be discounted in a similar way to the robust

smoothing technique. A robust cross-validation score $RCV(h_1, h_2)$ is defined as

$$RCV(h_1,h_2)=\frac{\sum_{i=1}^{n} w_i\left(y_i-\hat{m}_{h_1,h_2(-i)}\right)^2}{\sum_{i=1}^{n} w_i}$$

Where, W_i is the final robustness weight defined earlier, $\hat{m}_{h_1h_2(-i)}$ is the robust estimate of $m(u_i, v_i)$ with h_1, h_2 as band widths and without using the i^{th} observation y_i. The value of h_1 and h_2 which minimizes the robust cross-validation score $RCV(h_1, h_2)$, will be the optimum bandwidths. The robust estimate of $m(u_i, v_i)$ corresponding to the optimum bandwidth is denoted as $\hat{m}(u_i, v_i)$. A robust estimate of the error variance σ^2 is given as

$$\hat{\sigma}=\frac{median_i\left|e_i-median_j(e_j)\right|}{0.6745}$$

Where, $e_i = y_i - \hat{m}(u_i, v_i)$ The standardized residuals se_i (i=1,…,n) are obtained by $se_i=\frac{e_i}{\hat{\sigma}}$

The observations with large se_i (>4) is considered as outliers.

Table 1. Robust functions

Type	Range	$\rho(x)$	$\psi(x)$	W(x)
Huber M	$\lvert x\rvert < c$	$\frac{x^2}{2}$	x	1
	$\lvert x\rvert \geq c$	$c\lvert u\rvert-\frac{c^2}{2}$	$c[sign(x)]$	$\frac{c}{\lvert x\rvert}$
Biweight	$\lvert x\rvert < c$	$\frac{c^2}{2}\left(1-\left[1-\left(\frac{x}{c}\right)^2\right]^3\right)$	$x\left[1-\left(\frac{x}{c}\right)^2\right]^2$	$\left[1-\left(\frac{x}{c}\right)^2\right]^2$
	$\lvert x\rvert \geq c$	$\frac{c^2}{6}$	0	0
Hampel	$\lvert x\rvert \leq a$	$\frac{x^2}{2}$	x	1
	$a < \lvert x\rvert \leq b$	$a\lvert x\rvert-\frac{a^2}{2}$	a[sign(x)]	$\frac{a}{\lvert x\rvert}$
	$b < \lvert x\rvert \leq c$	$ab-\frac{a^2}{2}+(c-b)\frac{a}{2}\left[1-\left(\frac{c-\lvert x\rvert}{c-b}\right)\right]$	$\left[a\left(\frac{c-\lvert x\rvert}{c-b}\right)\right]sign(x)$	$a\left(\frac{c-\lvert x\rvert}{c-b}\right)$
	$\lvert x\rvert > c$	$ab-\frac{a^2}{2}+(c-b)\frac{a}{2}$	0	0

Simulation study

A simulation study on the finite sample performance of the proposed method was performed. The following spatial regression model was used in the simulation study

$$y_i = m(u_i, v_i)+\varepsilon_i,\ i=1,....n$$

Where, y_i, i=1,…,n are the observations, the regression function m is taken as $m(u_i, v_i)$=3 sin $[\pi(u_i,v_i)]$, the spatial locations (u_i,v_i), i=1,…,n are obtained by dividing the region [0,1]x[0,1] equally, the random error ε follows $N(0, \sigma^2)$ and the outliers are taken from N(6 σ, σ^2). Based on the above, 100 sets of data are simulated with n=225, 400 and σ=1.0, 2.0. 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 bandwidths in the u and v directions are taken as the same ($h_1=h_2=h$). Robustness in smoothing was achieved by setting different ρ functions (Table 1.). One set of simulated and estimated data (biweight method) along with the true regression function m for n=400 and σ =2.0 are shown in Figure 1.

The *MSE* of the estimated values with the true values of one set of simulated data is obtained by:

$$AMSE(m) = \frac{1}{100}\sum_{i=1}^{100}\frac{1}{n}\sum_{i=1}^{n}\left[m(u_j,v_j) - \hat{m}_{(i)}(u_j,v_j)\right]^2$$

Where, $\hat{m}_{(i)}\left(u_j, v_j\right)$ is the estimated value of corresponding to the i^{th} data set.

The average cross-validation (*CV*) score in the case of *KWLRE* and the average robust cross-validation (*RCV*) score in the case of robust estimates along with the average mean squared errors (*AMSE*) of the 100 sets of simulated data with the true values for different bandwidths are given

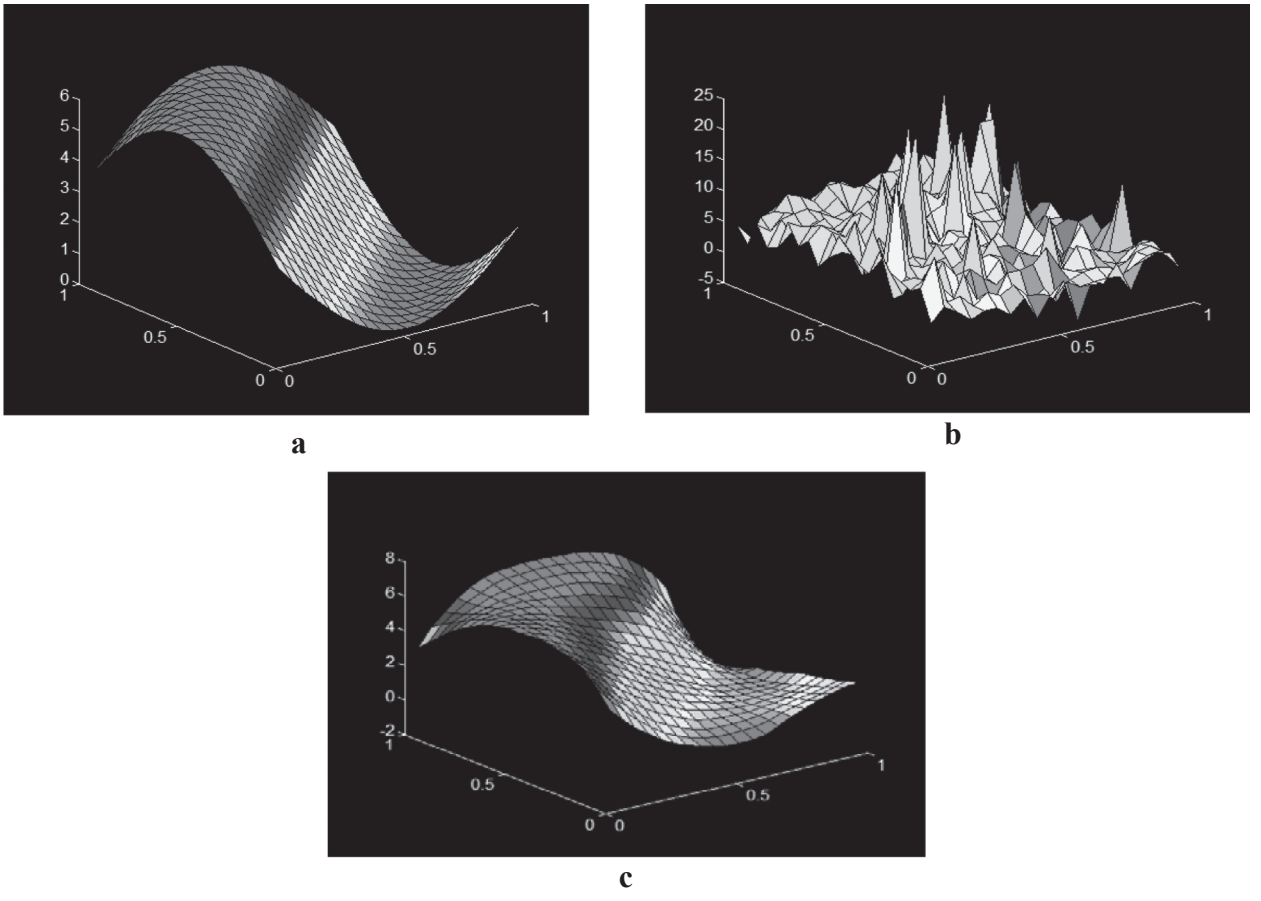


Fig. 1 (a). True spatial function of the simulated data; (b). Spatial representation of one set of simulated data; (c). Estimated spatial function based on robust spatial smoothing technique

Table 2. The average *CV/RCV* score and the corresponding *AMSE* of the 100 set of simulated data with *n*=400 and σ =2

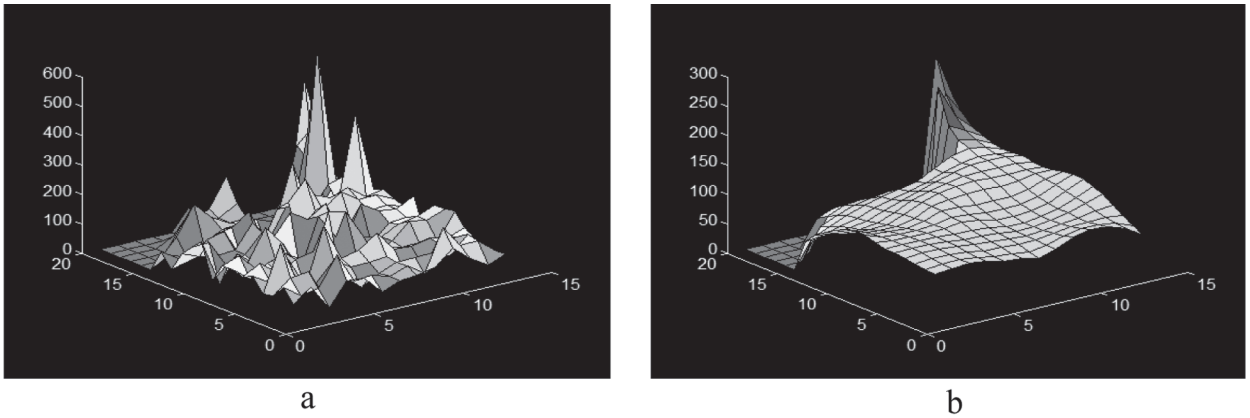
Type	*h*	0% outliers		4% outliers		8% outliers	
		CV/RCV	*AMSE*	*CV/RCV*	*AMSE*	*CV/RCV*	*AMSE*
KWLRE	0.16	4.2361	0.2616	10.2519	0.8585	16.7943	1.8039
	0.24	**4.1973**	**0.1753**	10.1263	0.5745	**16.5387**	1.5109
	0.32	4.2065	0.1867	**10.0495**	0.5647	16.8041	1.4385
	0.40	4.2922	0.2097	10.2994	**0.5582**	16.7438	**1.3812**
Robust Spatial (Huber)	0.16	3.5621	0.2593	5.0669	0.3234	6.7754	0.4309
c=1.345	0.24	3.6230	0.2613	**4.9626**	**0.2245**	**6.6241**	**0.3249**
	0.32	**3.5504**	**0.1931**	5.0754	0.2362	6.6759	0.3298
	0.40	3.6219	0.2466	5.1067	0.2842	6.7397	0.3790
Robust spatial (Biweight)	0.16	3.4366	0.2518	3.5056	0.2628	3.7295	0.2890
c=4.6851	0.24	3.4733	0.1960	**3.4848**	**0.1979**	**3.6206**	**0.2084**
	0.32	**3.3723**	**0.1938**	3.5811	0.2185	3.6292	0.2290
	0.40	3.4139	0.2790	3.5065	0.2525	3.7610	0.2829
Robust spatial (Hampel)	0.16	3.6545	0.2628	4.7700	0.3134	5.8691	0.3712
a=1.70, b=3.40 c=8.0	0.24	**3.6379**	**0.1804**	**4.4993**	**0.2062**	5.6959	0.3049
	0.32	3.6985	0.1934	4.5528	0.2280	**5.6944**	**0.2540**
	0.40	3.7036	0.2226	4.7161	0.2426	5.7606	0.2609

in Table 2. The optimum bandwidth (bandwidth corresponds to the minimum *MSE*) will depend on the curvature of the function. The optimum bandwidth for estimating the regression function is obtained based on the cross-validation technique given in Section 2. The optimum bandwidth *h* and corresponding *AMSE* of the 100 sets of simulated data with σ=2 and σ=1 are given in Table 3 and Table 4, respectively.

Results and discussion

It can be seen that the *AMSE* corresponding to the optimum bandwidth (bandwidth corresponds to the minimum value of *CV* or *RCV)* is the lowest in the case of zero per cent outliers for all types of estimators (Table 3 and Table 4). In the presence of outliers, the *AMSE* corresponds to the optimum bandwidth is not the minimum in the case of non-robust *KWLRE*. The estimated optimum *RCV* of the robust spatial smoothing techniques (Huber, Biweight or Hampel) provided the minimum *AMSE* in the presence or absence of outliers.

The comparison of *AMSE* corresponding to the optimum bandwidth shows that the non-robust spatial technique (*KWLRE*) performs very badly in



a b

Fig. 2 (a). Spatial representation of annual coconut yield data; (b). Estimated spatial function

Table 3. The optimum bandwidth (*h*) and corresponding *AMSE* of the 100 set of simulated data with σ = 2

Type	*n*	0% outliers		4% outliers		8% outliers	
		h	*AMSE*(m)	*h*	*AMSE*(m)	*h*	*AMSE*
KWLRE	225	0.32	0.2600	0.40	0.6595	0.32	1.5606
	400	0.24	0.1753	0.32	0.5647	0.24	1.5109
Robust spatial (Huber)	225	0.40	0.2969	0.32	0.3067	0.32	0.3819
	400	0.32	0.1931	0.24	0.2245	0.24	0.3249
Robust spatial (Biweight)	225	0.40	0.2844	0.40	0.2857	0.32	0.2851
	400	0.32	0.1938	0.24	0.1979	0.24	0.2084
Robust spatial (Hampel)	225	0.32	0.2661	0.32	0.2847	0.40	0.3658
	400	0.24	0.1804	0.24	0.2062	0.32	0.2540

Table 4. The optimum bandwidth (h) and corresponding *AMSE* of the 100 set of simulated data with σ = 1

Type	*n*	0% outliers		4% outliers		8% outliers	
		h	*AMSE*(m)	*h*	*AMSE*(m)	*h*	*AMSE*
KWLRE	225	0.24	0.0990	0.24	0.2874	0.32	0.4466
	400	0.16	0.0726	0.16	0.1809	0.24	0.3913
Robust spatial (Huber)	225	0.24	0.1068	0.24	0.1278	0.24	0.1516
	400	0.16	0.0789	0.16	0.0833	0.16	0.1231
Robust spatial (Biweight)	225	0.24	0.1118	0.16	0.1198	0.24	0.1154
	400	0.16	0.0734	0.16	0.0778	0.16	0.0815
Robust spatial (Hampel)	225	0.24	0.0998	0.24	0.1249	0.24	0.1403
	400	0.16	0.0733	0.16	0.0783	0.16	0.1121

the presence of outliers. Among the robust estimators, the robust spatial smoother with biweight robustness weight function performed better than the Huber and Hampel weight functions. Comparison of *AMSE* corresponding to the optimum bandwidth showed that there was not much difference between different types of robustness weight function in the absence of outliers. In the case of robust spatial smoother with biweight robustness weight function, the *AMSE* for 0 per cent, 4 per cent and 8 per cent outliers are almost the same, and it indicated that the method was robust against the outliers.

The method was applied to the annual yield data of 225 coconut palms (Philipines Ordinary variety) in a field to estimate/eliminate spatial effect and to identify the extreme observations or outliers present in the data. The coconut palms were planted with a spacing of 8m x 8m in a plot with 13 rows and 15 to 20 palms in each row. The spatial representation of annual yield data of 2013 is given in Figure 2(a). The robust nonparametric spatial smoothing technique with the biweight robust function described in Section 2 was used to estimate the spatial response function. The estimated spatial function is shown in Figure 2(b). The outliers were identified using the standardized residuals with MAD (from the median)/0.6745 as the scale factor. Mean was taken as the estimate in the case of without removing the spatial effect. An observation was taken as an outlier if the absolute value of its standardized residual was greater than 4. The number of outliers identified and the *MSE* after removing the outliers based on spatial, robust spatial and without removing the spatial effect (mean) are

Table 5. MSE of the observed value with the estimated value after removing the outliers

Estimate	No. of outliers	*MSE*
Mean	0	4482
KWLRE	3	2031
Robust Spatial (biweight)	5	1933

given in Table 5. By removing spatial effects and outliers, the *MSE* has reduced more than 50 per cent.

Conclusion

In the presence of outliers or extreme observations in the data, the smoothing technique or the nonparametric regression technique performs very badly, particularly estimates near the outliers. The proposed method based on the robust M-kernel weighted local linear regression smoother to fit the spatial regression function performed well in the presence of outliers. The outliers in the data are identified by analyzing the residuals. The proposed method is useful to estimate/eliminate the spatial effect and identify the high potential trees in an orchard, which is useful for the breeding programs.

References

Cleveland, W.S. and Devlin. S.J. 1988. Locally-Weighted Regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association* **83**(403): 596-610.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. 1986. *Robust Statistics-The Approach Based on Influence Functions*. New York: John Wiley and Sons.

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

Huber, P. J. 1981. *Robust Statistics*. New York: John Wiley and Sons.

Jose, C.T. and Ismail, B. 2001. Nonparametric inference on jump regression surface. *Journal of Nonparametric Statistics* **13**: 791-813.

Leung, D. 2005. Cross-validation in nonparametric regression with outliers. *The Annals of Statistics* **33**: 2291-2310.

Rey, W.J.J. 1983. *Introduction to Robust and Quasi-robust Statistical Methods.* Heidelberg: Springer-Verlag.

Ruppert, D. and Wand, M.P. 1994. Multivariate locally weighted least squares regression. *Annals of Statistics* **22:** 1346-1370.

Tukey, J. W. 1977. *Exploratory Data Analysis*. Reading: Addison-Wesley.

Wang, F.T. and Scott, D.W. 1994. The *L*1 method for robust nonparametric regression. *Journal of the American Statistical Association* **89:** 65-76.