# NEAREST NEIGHBOUR PATTERN OF SPATIAL VARIATION IN EXPERIMENTAL FIELDS.

Dauda, T.O.

Institute of Agricultural Research and Training, Obafemi Awolowo University, PMB 5029, Moor Plantation, Ibadan, Nigeria

taofik biomet@fastmail.net

Abstract: Evaluations of the nearest neighbour analysis in the study of spatial variation in experimental plot have been attempted for kenaf using a split plot experiment distributed in a complete randomized design. The experiment was carried out between June and September 2006 at Ilora and Ikenne outstation of the Institute of Agricultural Research and Training, Ibadan to evaluate nearest neighbourhood in experimental plots. The results of the cluster analyses of the stem girth at Ilora showed that 80% of the pairing plots were isotropic in nature while all other plot pairs are anisotropic in nature because their euclidean distances are not equal. For stem girth at Ikenne, isotropic property was exhibited between only  $x_{4,1}$  vs  $x_{3,1}$  and  $x_{5,3}$  vs  $x_{2,3}$  (0.032). All other plots pairs are anisotropic in nature. For plant height at both Ilora and Ikenne, none of the pairing plots exhibited isotropic property but anisotropic property. Also, the nearest neighbour indices are 0.00197 (for stem girth at Ilora), 0.00734 (for stem girth at Ikenne), 0.1831 (for plant height at Ilora) and 0.2456 (for plant height at Ikenne). From the study, the variogram is found to be related to the variance covariance using the model,  $\gamma(h) = C(0) - C(h)$  (where  $\gamma(h) = v$  variogram, V variogram, V variogram, V variogram, V varione at the plot V varione at the plot V varione at the neighbour index obtained in this work implied that the neighbourhood pattern falls between cluster and randomness thereby reflecting patchiness of neighbourhood pattern. [Nature and Science 2010;8(4):44-53]. (ISSN: 1545-0740)]

**Keywords**: Nearest neighbour, Euclidean distance, Clusters, initial soil nutrient deposition ((ISND).

### Introduction

Treatment responses are expected to be cumulative of the initial soil nutrients deposition (ISND). These treatments' responses on any experimental plots are usually subjected to varying degree of statistical tools with the assumption that that the pattern of distribution of the ISND are random. Hence, most of the statistical tools presume randomness of these ISND consequent to the nature of the preliminary investigation carried out randomly on the experimental site. The studies of spatial statistics have included the employment of several statistical tools to analyze treatment responses relative to spatial variability in experimental plots. methods include the variance covariance analysis, well autocorrelation as as autocovariance, correlelograme, the similarity matrices and the global and local indicator of spatial autocorrelations, (Hardy and Sonke, 2004, Christakos and Hirstopulos, 1998, Richard and Dean, 2002 and Schabenberger and Pierce, 2000). These various methods have been employed by different authors to assess spatial variability at different level of spatial statistical study. Nearest neighbour analysis is an important spatial variability measurement tool because it incorporates the similarities and the diversity of such experimental plots into the analysis of experimental plots' variation pattern. Nearest neighbour index have been defined as a measure of the amount of spatial dispersion in a set of point features based on the distance (linear) of any points to its nearest neighbour, (Benwell, et.,al, 2002). It explore the amount of spatial dispersion based on the ratio of average inter – point distance between nearest neighbour (Ad) to the expected value of the average inter – point distance between randomly dispersed nearest neighbour (Ed). That is if X, D are a metric space  $x \in X$  and P is a positive real number then, the D – p- neighbourhood of x is defined to be the set of all points y of X such that D(x,y) < P.

Works on nearest neighbour include Weigelt and Jolliffe (2003)'s indices of plant competition, Purves et., al (2003)'s Nearest neighbour for avalanche forecasting in Scotland and Singh and Ganju (2004)'s Supplement to nearest neighbour method for avalanche forecasting. These works notwithstanding, nearest neighbour analysis of spatial variation in experimental plots is crucial because of its uses in the preliminary assessment of experimental plots. It has been less focused because activities at agricultural field are usually artificially induced unlike in other field (such as geography, ecology and forestry) where natural phenomenon takes its course. Also, nearest neighbour analysis is one of the required regular and periodic updating of spatial statistics. It is one of the tools that can boost the estimation of

spatial variability in experimental plots. Nearest neighbour analysis seeks to answer question relating to the distribution of the treatment responses which are said to be cumulative. The objective of this study was therefore to evaluate spatial variation pattern in experimental plots using nearest neighbour analyses.

### Materials and Methods.

Data sets for this project were from "evaluation of the effect of fertilizer and insecticides on Kenaf" set up each at Ikenne and Ilora out stations of the Institute of Agricultural Research and Training, Ibadan between June and September 2006. Ikenne falls within the forest zone (27°.48!N and 3°.52!) of the country while Ilora is located in the intermediate guinea savanna (126°.52!N and 3°.41!). Each of the experiment was carried out using split plot design. The main plot was the spraying regime ( $S_1 = 300 \text{kg}$ NPK+100kg Furadan + 2 pre flowering insecticide sprays and  $S_2 = 600g$  NPK + 200kg Furadan + 4 pre flowering insecticide sprays) while the sub plot was the varieties ( $V_1$  = Cuba 108,  $V_2$  = Ifeken 400 and  $V_3$ = local cultivar). Data on stem girth and plant height were collected at interval of 2 weeks commencing at 4 weeks after planting and relative to their spatial positions.

The data obtained from the experiment were subjected to descriptive statistics (such as means and variances) as well as variogram of the original data and spatial variability effects data. Also, spatial proximity Matrices were constructed to establish the

measure of nearest between the plots. This is done through Matrix W (n x n) where  $W_{ij}$  represents the measure of nearest between  $O_i$  and  $O_j$ . It should be noted that n = 36.

$$W_{ij} = \begin{cases} 1, & (o_i, o_j) < h \\ 0 \end{cases} o.w$$

Where h is the average distance of all  $O_i$  and  $O_j$ .  $O_i$  and  $O_j$  are respectively, reference and neighbouring plots. Lastly, nearest neighbour index was computed

using 
$$NNI = \frac{Ad}{Ed}$$

where  $Ad = \frac{\left(\sum_{i} d_{i}\right)}{n}$  and  $Ed = \frac{1}{2}\sqrt{\frac{A}{n}}$ 

and n = number of points, A = map area. NNI according to Benwell *et.al.*, 2002 and Scharbenberger and Pierce, 2000 can range between zero (all points are at the same location) and 2.1419.

## Results.

For variogram measurements, N (h) is the number of pairs of values  $Z(x_i)$  and  $Z(x_i + h)$  separated by vector h which suppose to be the separating distance.  $W_{ij}$  as contained in table 1 have been converted to their proportions.

Table 1. Means and Variances of Different Variables at Different Plots irrespective of the weeks.

Treatment	Plot	Stem	Vari-	Stem	Vari-	Height	Vari-	Height	_
	Address	girth — (cm)	ance	girth- (cm)	ance	(cm ) (Ilora)	ance	(cm) (Ikenne)	Vari-
		(IIora)		(Ikenne)		(,		(	ance
V1S1	1,1	0.582	0.039	0.825	0.023	78.604	727.327	100.968	- 852.406
V1S1	1,2	0.660	0.063	0.824	0.042	92.388	928.969	73.893	797.328
V1S1	1,3	0.642	0.043	0.998	0.079	67.452	672.832	83.448	1448.335
V1S1	1,4	0.737	0.043	1.178	0.192	108.376	761.429	131.812	2236.439
V1S1	1,5	0.816	0.043	1.166	0.132	116.092	791.860	141.060	3577.834
V1S1	1,6	0.641	0.033	0.941	0.033	59.192	454.969	75.736	558.897
V2S1	2,1	0.518	0.043	0.758	0.041	43.640	177.289	64.028	851.239
V2S1	2,2	0.627	0.034	0.754	0.062	54.184	301.752	46.984	672.676
V2S1	2,3	0.708	0.031	1.256	0.162	96.872	911.496	126.168	2917.620
V2S1	2,4	0.744	0.032	1.232	0.176	99.304	545.690	136.256	2121.984
V2S1	2,5	0.987	0.057	1.118	0.125	133.704	770.340	139.672	2504.302
V2S1	2,6	0.762	0.025	1.026	0.054	98.316	535.321	118.052	2577.796
V3S1	3,1	0.534	0.031	0.844	0.034	74.060	549.246	89.316	723.825
V3S1	3,2	0.611	0.023	0.829	0.064	82.460	370.370	71.108	744.144
V3S1	3,3	0.791	0.057	1.225	0.292	113.108	656.282	122.820	2174.450

1 tataic a	ma betenee	2010,0(1)							
V3S1	3,4	0.809	0.103	1.033	0.117	72.168	880.843	99.068	2256.535
V3S1	3,5	0.627	0.019	0.965	0.097	44.920	302.203	97.464	1991.748
V3S1	3,6	0.638	0.017	0.963	0.047	54.336	454.670	125.520	2932.981
V2S2	4,1	0.695	0.036	0.856	0.047	95.556	437.521	84.032	556.016
V2S2	4,2	0.639	0.044	0.860	0.067	58.044	294.328	86.812	1380.568
V2S2	4,3	0.665	0.038	0.986	0.101	64.672	488.808	98.992	2479.741
V2S2	4,4	0.851	0.044	1.266	0.135	123.160	1510.566	155.940	3644.054
V2S2	4,5	0.581	0.056	0.896	0.023	44.920	302.203	74.508	642.724
V2S2	4,6	0.623	0.035	0.958	0.049	54.300	456.036	68.760	568.775
V3S2	5,1	0.676	0.035	0.766	0.046	52.260	302.273	66.172	854.861
V3S2	5,2	0.725	0.034	1.096	0.105	109.504	1138.860	139.328	3084.581
V3S2	5,3	0.777	0.030	1.256	0.170	118.012	435.807	134.156	2563.701
V3S2	5,4	0.797	0.060	1.194	0.109	108.952	702.939	148.384	2746.284
V3S2	5,5	0.688	0.042	1.032	0.062	102.164	1021.825	128.220	2735.813
V3S2	5,6	0.784	0.029	0.988	0.166	108.184	510.308	112.768	3516.609
V1S2	6,1	0.690	0.028	0.776	0.062	94.804	477.117	75.252	734.898
V1S2	6,2	0.798	0.031	1.187	0.134	117.920	1089.066	130.636	2046.241
V1S2	6,3	0.844	0.052	1.206	0.177	120.408	953.019	125.096	2373.249
V1S2	6,4	0.774	0.099	1.058	0.075	79.740	414.530	106.732	2325.458
V1S2	6,5	0.762	0.028	0.912	0.023	105.672	890.485	124.648	3102.044
V1S2	6,6	0.753	0.015	0.975	0.064	103.504	290.515	124.908	4438.358

This was because of the need to factorize out the large values of the plant height returned for plots at both sites and for the fact that  $-1\angle \gamma(h)\angle 1$ . Z  $(x_i)$  now becomes  $Z(x_i)$  $Z(xi + h) / \sum_{i=1}^{n} Z(x_i)$  were 0.002260833

Nature and Science

2010;8(4)

(Stem girth, Ilora), 0.004194 (Stem girth, Ikenne), 0.003316145 (Plant height, Ilora) and 0.003611258 (Plant height, Ikenne) while the variogram { \gamma (h)} for stem girth and plant height at Ilora and Ikenne were respectively, 0.0000314, 0.0000583, 0.0000461 and 0.0000502. For the different variables including the SVE, nugget effect was observed to be negligible because semivariogram at  $\gamma(0) = 0$  or approximately zero. The semivariogram thus were said not to present the nugget effects. Also, the semivariogram of the variables showed that the practical range for the stem girth at Ilora and Ikenne were 16 and 14 while real sill were not obtained for the variables at Ilora and The semivariogram of the stem girth Ikenne. continually increased monotonically up till lag 16 (for Ilora) and lag 12 (for Ikenne). The reverse were the case for lag greater than these two lags implying the presence of "hole effects" (Table 2). For plant height however, the presence of hole effects is entire because the semivariogram lack monotonical increments.

The variogram for the SVE of the different variable showed that the data does not fulfill the assumption of stationarity. This is because the fitted line does not follow a linear sill semi variogram pattern (Figure 1). That is none of the fitted line presented real range, sill and a very low nugget effect. The nugget effect of stem girth at Ilora is -0.02 while stem girth at Ikenne has practically no nugget effects. The nugget effects for the plant height at Ilora are 0.06 while that of Ikenne is -0.05. It is noteworthy that the variogram patterns of the raw data differ from that of SVE. Since the variogram of the SVE presented nugget effects (that is  $\gamma(0) \neq 0$ ), hence physical phenomenon at smaller scale which must not have been well resolved existed.

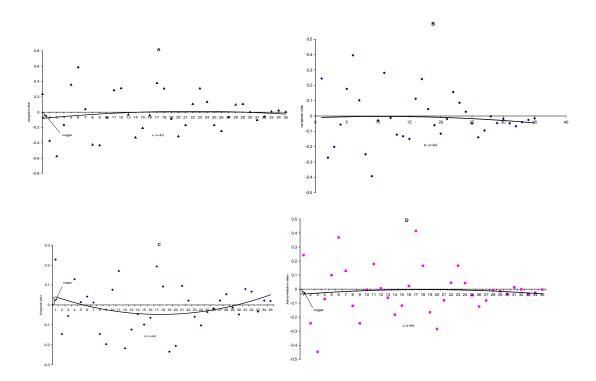


Figure 1. Variogram of the spatial variability effects for the Stem girth at Ilora (A) and Ikenne (B) as well as Plant height at Ilora (C) and Ikenne (D).

Table 2. Semivariogram, Z(h<sub>i</sub>) Values of stem girth and Plant height at Ilora and Ikenne.

Stem girth (Ilora)	Stem girth (Ikenne)	Plant height (Ilora)	Plant height (Ikenne)
5.55E-05	2.45E-05	1.61E-05	3.88E-07
3.48E-05	5.37E-05	2.69E-05	4.06E-05
1.97E-08	0.000223	3.57E-06	3.59E-05
1.97E-08	4.96E-05	3.88E-07	6.61E-05
1.44E-05	0.000297	6.04E-05	0.000551
1.33E-05	3.27E-06	8.37E-05	1.29E-05
9.56E-06	1.92E-05	2.14E-05	4.42E-06
1.97E-06	0.000215	0.000214	0.000332
1.78E-07	2.62E-06	6.08E-05	2.66E-05
7.11E-05	3.82E-05	2.52E-05	6.67E-06
0.000123	0.000133	2.78E-05	2.24E-07
6.4E-06	2.09E-05	1.16E-07	0.00024
1.23E-05	4.06E-05	2.29E-05	5.93E-08
0.000149	0.000751	5.29E-05	0.000158
0.000133	0.000355	2.15E-05	3.2E-07
0.000661	8.15E-06	0.000197	3.48E-06
9.67E-07	8.17E-05	2.02E-05	3.83E-05
6.64E-05	0	2.15E-07	0.000394
7.9E-06	1.6E-05	1.84E-05	7.77E-05
3.34E-06	3.15E-05	3.2E-05	6.73E-05
3.34E-06	2.17E-05	0.000366	4.71E-05
1.44E-05	0.000426	0.000601	0.000517
4.55E-05	4.56E-05	2.05E-05	9.52E-07

7.9E-08	4.44E-07	2.05E-05	1.22E-05
1.97E-08	0.00011	0.000349	0.000292
3.34E-06	7.01E-05	0.000216	1.01E-05
0.000102	6.09E-05	4.14E-05	1.32E-06
3E-05	5.87E-05	3.87E-05	4.21E-09
2.28E-05	0.000227	0.000115	2.06E-05
4.94E-07	0.000229	7.27E-07	0.000437
1.97E-06	0.000124	0.000162	0.000138
5.34E-05	2.74E-05	5.91E-06	5.11E-06
0.000149	0.000196	0.000144	1.02E-07
0.000432	0.000139	0.000117	2.35E-05
3.65E-05	9.6E-05	0.000213	5.03E-05

# Spatial proximity matrices and nearest neighbour index

The spatial proximity matrix of the stem girth at Ilora station showed that magnificent percentage (31.5%) of the pairing neighbours were distinct from each other while the remaining 68.5% were similar to each other, (Appendix I).  $x_{1.6}$ ,  $x_{5.2}$  and  $x_{5.6}$  were distinct from their neighbours because their proximity matrix returned zero values. However, these were not the case with the stem girth at Ikenne with 11.7% non similar plots. Also, there were some other non similar plots which include  $x_{4,5}$  vs  $x_{3,2}$ ,  $x_{4,5}$ vs  $x_{3,6}$ ,  $x_{4,5}$  vs  $x_{4,1}$ ,  $x_{5,5}$  vs  $x_{5,2}$ ,  $x_{5,6}$  vs  $x_{4,5}$  and  $x_{6,4}$  vs  $x_{4,5}$ . Also not similar were  $x_{6,5}$  vs  $x_{2,2}$  and  $x_{6,5}$  vs  $x_{2,5}$ (Appendix II). Different trends were obtained for the similarity matrices of the plant height at Ilora where 17.4% of the pairing neighbours were not similar with each others, (Appendix III). There was no plot that was completely different from the others. More than 50% plots pairing with both plots  $x_{1,2}$  and  $x_{2,3}$  were dissimilar. All neighbouring plots of plant height at Ikenne station were similar and none was distinct from the others, (Appendix IV). The cluster analyses of the stem girth at Ilora showed that 0.8 of the pairing plots were isotropic in nature. This isotropic property was exhibited between the following pairs of plots;

Plots  $x_{4,3}$  vs  $x_{1,3}$  and  $x_{6,2}$  vs  $x_{5,3} = 0.044$ ;  $x_{1,4}$  vs  $x_{4,3}$  and  $x_{1,4}$  vs  $x_{4,2} = 0.46$ ;  $x_{4,5}$  vs  $x_{1,1}$  and  $x_{5,6}$  vs  $x_{6,2} = 0.051$ ;  $x_{6,6}$  vs  $x_{5,6}$  and  $x_{2,4}$  vs  $x_{5,6} = 0.052$ ;  $x_{2,4}$  vs  $x_{3,3}$  and  $x_{3,2}$  vs  $x_{2,4} = 0.053$ ;  $x_{3,2}$  vs  $x_{5,1}$  and  $x_{3,1}$  vs  $x_{2,1} = 0.058$ ;  $x_{4,3}$  vs 14 and  $x_{1,5}$  vs  $x_{5,4} = 0.061$ . All other plot pairs are anisotropic in nature because their euclidean distances are not equal. For stem girth at Ikenne, isotropic property was exhibited between only  $x_{4,1}$  vs  $x_{3,1}$  and  $x_{5,3}$  vs  $x_{2,3}$  (0.032). All other plots pairs are anisotropic in nature. For plant height at both Ilora and Ikenne, none of the pairing plots exhibited isotropic property but anisotropic

property. This is because none of the pairing plots cluster at the same Euclidean distances.

The nearest neighbour indices for the different parameters at both sites (Ilora and Ikenne) were calculated using the average interpoint distance-Ad between nearest neighbour, (Benwell, 2002). These inter point distances are 0.00411 (for stem girth at Ilora), 0.0153 (stem girth at Ikenne), 0.381 (plant height at Ilora) and 0.512 (plant height at Ikenne). The expected average inter point distance between nearest neighbour on the other hand is 2.083 for both sites and parameters.

The nearest neighbour indices are therefore; Stem girth = 0.00197 (Ilora) and 0.00734 (Ikenne) Plant height = 0.1831 (Ilora) and 0.2456 (Ikenne).

# Appendix I. Similarity Matrix of Stem girth at Ilora Station.

Appendix II. Similarity Matrix of Stem girth at Ikenne Station.

2 0 1
3 1 0 1
4 1 0 1 1
5 0 0 0 0 1

**6** 1 0 1 1 0 1

# Appendix III. Similarity Matrix of Plant height at Ikenne Station.

# Appendix IV. Similarity Matrix of Plant height at Ikenne Station.

# Discussion and conclusion.

It is obvious that the variogram does not follow a definite pattern hence, could be classified as zero nugget effects. Also, the variogram is noticeably not a function of the size of the values of the variable under consideration but may be function of both the number of the lag period returning negative as well as positive signs in addition to the size of the data. The practical range as obtained from the semivariogram implied that the spatially correlated data exist at almost around the same place regardless of the zones of the experimentation (that is Ilora or Ikenne). The variogram according to Doncker *et.al.*, (2006) can be related to the variance covariance using the model,  $\gamma(h) = C(0) - C(h)$  (where  $\gamma(h)$  = variogram, C(0) is the variance at the plot  $x_i$  and C(h) is the covariance at both plot  $x_i$  and  $x_j$ ). In addition, pattern

of neighbourhood differs across sites. This might be a reflection of the spatial distribution of the soil or other site condition. Pattern of nearest neighbour from the nearest neighbour indices differs across the different growth variables. This may be hinged on the fact that the different growth variables are enhanced as well as inhibited by different nutrients and factor. The use of nearest neighbour analysis in the spatial variability measurements are often and widely discussed in geography and ecology due to dynamics of natural process. The original and formal idea of neighbourhood is restricted to object sharing common boundaries. This nearest neighbour has thus been incorporated into experimental agriculture where experimental activities are also based partly on natural phenomenon. Conceptually, the clear message is that randomness of spatial pattern does not imply lack of neighbourhood. Meanwhile, the three types of spatial pattern as identified by nearest neighbour index depicted different types of neighbours. These are: when similar plots share the same boundary; similar plots does not share the same boundary and when similar plots does not exist in the experimental plot. The low but positive nearest neighbour index as in this work implied that the neighbourhood pattern falls between cluster and randomness thereby reflecting patchiness.

Based on these therefore, nearest neighbour analysis is recommended to be carried out along with the preliminary investigation of the experimental plots. This would guide in the distribution of sampling points for preliminary investigation. The effects of plant type/plant species on the spatial variation in experimental plots is hereby recommended for further study.

# Acknowledgement

The author hereby acknowledges the contributions of Professor O.E.Asiribo, Professor L.O.Ojo and Dr K.S. Adekeye for the supervision and Dr G.A. Agbaje for the collaboration.

### References

- Benwell, G.L. McLennan, Grasberger, T. and Fryer, J. Spatial Data Analysis for Aboriginal Rock Extraction Sites at Brewarrina, NSW, Australia. Presented at the 14<sup>th</sup> Annual Colloquium of the Spatial Information Research Centre held at Victoria University of Wellinghton, Wellinghton, New Zealand between December 3<sup>rd</sup> 5<sup>th</sup> 2002,pp 11.
- Christakos, G., and Hristopulos, D.T. Spatio Temporal Environmental Health modeling: A Tractus Stochasticus. Dordrecht, the Netherland: Kluwer Academic Publisher, 1998; Pp 451.
- Critoru, C.A. and Doytsher, D.Y. Random Field and (Marked) Point process: A Pratical Comparison of Two Stochastics Models Proceedings of Joint Workshop on Multiscale Representations of spatial Data held in Ottawa, Canada, July, 7<sup>th</sup> -8<sup>th</sup> 2002;1-9.
- Doncker, P.H., Dricot, J.M., Mey, R., Helier, M. and Tabbara, W. Electromagnetic Fields Estimation Using Spatial Statistics. Electromagnetes, 2006; 26. 111 122.
- Fagroud, M. and Meirvenne, M. V.. Accounting for Soil spatial Autocorrelation in the Design of Experimental Trials. Soil Science Society of Americal Journal. 2002; 66: 1134 1142.
- Hardy, O.J. and Sonke, B.. Spatial Pattern analysis of Tree Species distribution in a Tropical Rainforest of Cameroon: Assessing the role of limited dispersal and niche differentiation in Forest Ecology and Management, 2004; 197; 191 202.
- Purves, R.S., Morrison, K.W., Moss, G., and Wright, D.S.B. Nearest neighbours for avalanche forecasting in Scotland development, verification and optimization of a model. In cold Region Science and Technology, 2003; 37 (3) 343-355.
- Schanbenberger, O. and Pierce, F.J.. Contemporary Statistical Models for the Plant and Soil Science. Washington, D.C; CRC Press, 2002; Pp 738.
- Singh, A., and Ganju, A. A Supplement to nearest-neighbour method for avalanche forecasting. In cold Region Science and Technology, 2004; 39, 2-3 pp 105-113.
- Weigelt, A., and Jolliffe P., Indices of Plant Competition. In Journal of Ecology, 2003; Vol 91 Pp 707-720.
- Xin X, Shaw S. and Lin H. Developing a Spatio Temporal Data Model for an exploratory Study on the Land use-Transportation interaction A temporal GIS Attemp- Being a report of a Project Sponsored by Florida Department of Transportation's University Research Programme (www. document URL <a href="http://www.amc">http://www.amc</a> on 23/02/2005. Pp 16.

12/21/2009