Removing Unplanned Spatial Variability in RCB Experiments
(A Poster Presentation)

Prakash Basnyat, Brian G. McConkey, and F. Selles
Semiarid Prairie Agricultural Research Centre, 1030 Airport Road, Swift Current, Saskatchewan S9H 3X2

Introduction:

The classical assumption of independence between plots is violated in the presence of significant spatially correlated trends. In this situation, classical statistical analysis may fail to detect real treatment differences. The classical approach is sufficient if measured soil and crop properties exhibit random variability with little or no spatial correlation. In such case, geo-statistical methods provides algorithms to identify and isolate unplanned spatial variability. The semivariogram is potentially a useful tool for quantifying spatial correlation between treated plots and indicating when classical methods of analysis may fail. One of the difficulties from using semivariograms of plots yield to evaluate spatial correlation is that yield is affected by the presence of trends as well as by the pattern in treatment randomization and replication. There are two general approaches taken toward modeling the trend and error structure in spatial processes. These are:

1. First remove the trend and perform any subsequent analysis on the resulting residuals.

2. Model the trend and error structure simultaneously.

We have taken the first approach looked at the median polish as a means of trend removal. Median polish is the name of an iterative algorithm for removing any trend present by computing medians for various coordinates on the spatial domain.

Objectives:

A. To evaluate the effectiveness of semivariogram in the identification of spatial variability in RCB experiments, and

B. To test the power of median polish in removing trend.

Methods:

The median polish algorithm assumes the following decomposition of the random process Z:

$$Z(t) = \mu(t) + \varepsilon(t)$$

and then proceeds by estimating the “grand”- effect, the row - effects (one for each row), and the column effects (one for each column) (Cressie 1993).

The steps of the algorithm are given below:

1. Take the median of each row and record the value to the side of the row. Subtract the row median from each point in that particular row.

2. Compute the median of the row medians, and record the value as the grand - effect. Subtract the grand effect from each row medians.
3. Take the median of each column and record the value beneath the column. Subtract the column from each point in that particular column.

4. Compute the median of the column medians, and add the value to the current grand effect. Subtract this addition to the grand effect from each of the column medians.

5. Repeat steps 1-4 until no changes occur with the row or column medians. It should be emphasized that through the course of this algorithm, the relationship given above between Z(t) and the various median polish effects is preserved at every step.

6. Create a trend surface from the results of a median polish residuals using kriging.

7. Extract the residuals of the points of interest from the kriged surface.

8. Add row effects, column effects, alleffects and extracted residual to create adjusted sample values.

To test the median polish algorithm the organic carbon data from the 1995 field experiment at the Experimental Farm of Agriculture and Agri-Food Canada, at Scott, Saskatchewan were used. The study area is located at Scott, Saskatchewan ($52^0$22’ N, $108^0$50’ W) covering 18 hectares. The study consisted of a combination of three different levels of farm input use and three crop rotations of varying diversity (Figure 1). Previous study by Selles and others (1999) in the same experiment looked into the relationship between biological and chemical measures of N supplying power and total soil N at field scale using geostatistics.

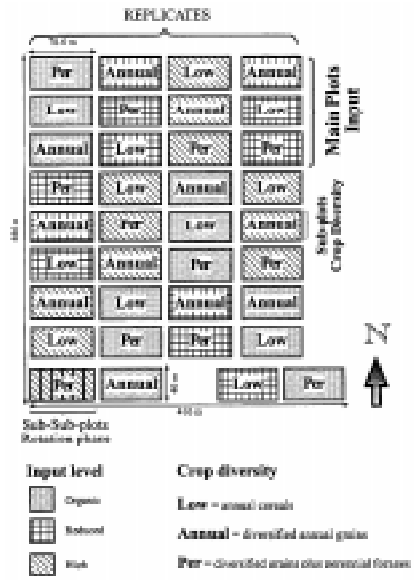

Figure 1. Layout of experiment in the field (Selles et al. 1999)

Results and Discussion:

Using easting and northing cocordinates, the value for the organic carbon was plotted and result is given in Figure 2.

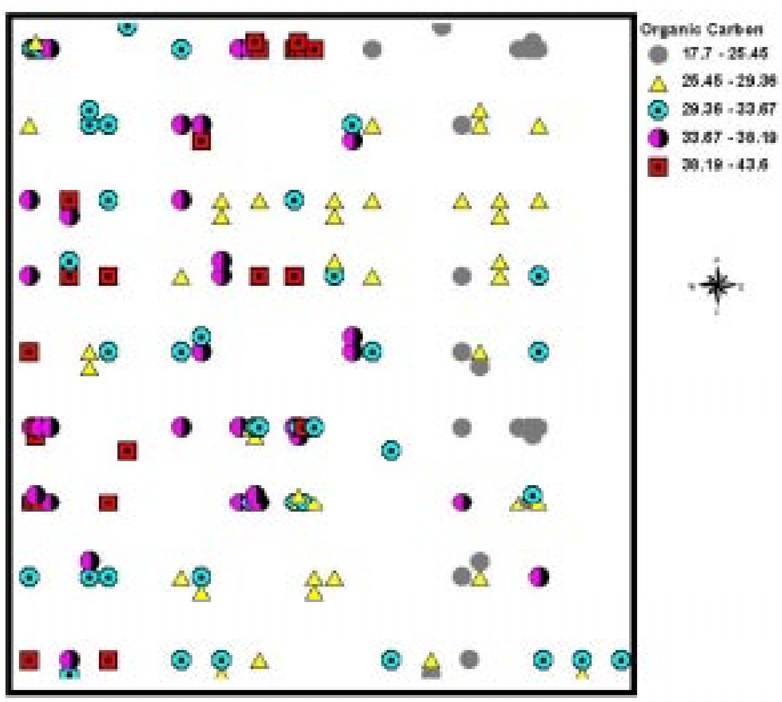

Figure 2. Sample Point Distribution

From this initial posting, we noticed that there are some unusually high values in places, and that there seemed to be a trend of larger to smaller soil organic carbon moving roughly from the SW to NE corner of the plot. seeded. The trend was also examined by computing the means and medians for each row and column and displaying them as shown in the Figure 3.

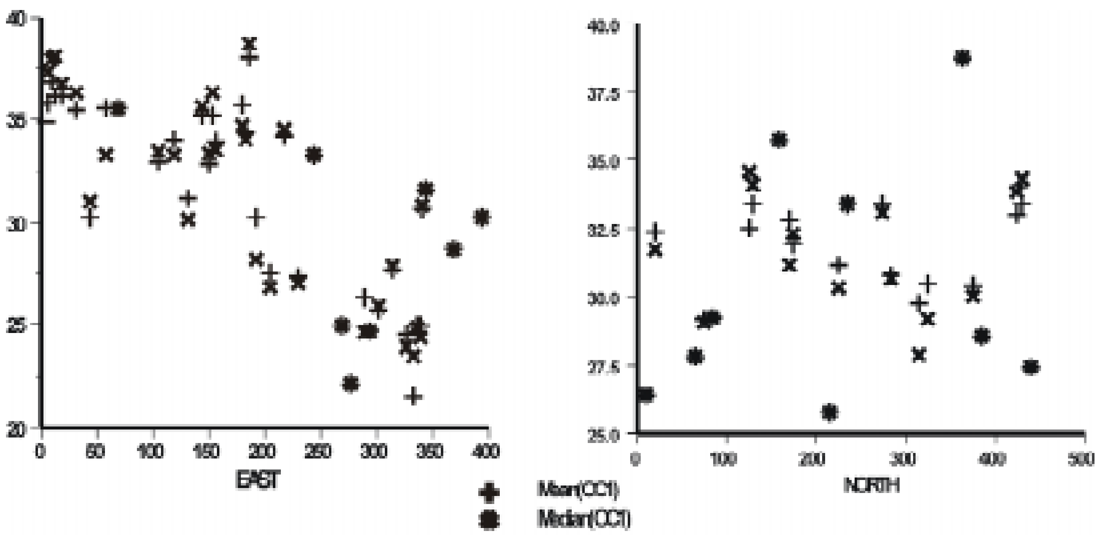

Figure 3. Means and Medians across rows and columns

Although there is clearly spatial trend across this area, it is not clear that the trend would be easily modeled by some function of the row and column coordinates. Hence to isolate large and small scale variation, we used the median polish algorithm on these data. The median polish algorithm returned the overall effects, row effects, column effects, and residuals using iterative process. The final table of this process is given in Table 1.

Table 1. Median Polish Results.

Sample #	EAST	NORTH	OC1	ROWEFFECTS	COLEFFECTS	ALLEFFECT	RESIDUALS
1.000	32.000	10.000	29.650	-4.283	4.620	31.106	-1.793
2.000	132.000	10.000	27.280	-4.283	0.073	31.106	0.384
3.000	270.000	10.000	23.170	-4.283	-4.532	31.106	0.879
10.000	157.000	20.000	27.350	1.175	-0.073	31.106	-4.858
11.000	244.000	20.000	33.480	1.175	0.406	31.106	0.794
.	.	.	.	.	.	.	.
150.000	231.000	425.000	24.160	2.235	-2.621	31.106	-6.560
158.000	337.000	430.000	22.680	4.000	-8.673	31.106	-3.753

After the separating data into large effects and small effects, kriging was done on the residuals first looking at autocorrelation using semivariogram, then based on the best fit model (i.e. linear in our case) kriging was performed to create a kriged surface. Information related to each sample points were extracted based on the kriged surface of the residuals. Then all effects, column effects, row effects, and residual information from the kriged surface for the sample points were added back to create adjusted data. 3D surface were generated to observe and compare the improvement (Figure 4 and Figure 5)

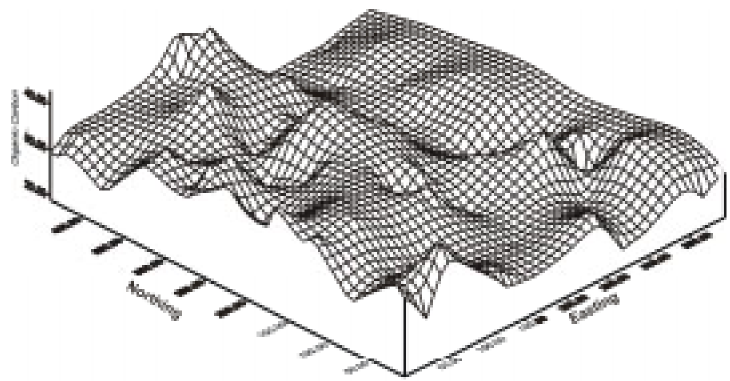

Figure 4. Original Data

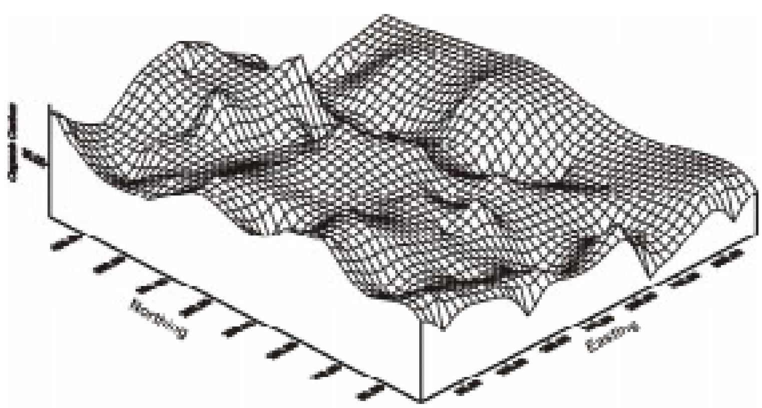

Figure 5. Adjusted Data

To assess whether or not the median polish successfully removed any unplanned spatial variability surface plots of the regions before and after median polish. The technique of median polish is generally used prior to computing the variogram to remove any spatial trend present which is not well-described by some parsimonious function. As estimation of the variogram presumes a stationary mean and variance, it is generally reasonable to assume that the residuals resulting from a median polish satisfy this property. Alternative to median polish to remove unplanned variability is to model the trend and covariance structure simultaneously, as is done with universal kriging. The difficulty here is that it is often very hard to separate which effects are due to large-scale and small-scale variation.

Conclusions:

Though there are numerous other types of polishing, medians is the most common choice due to its general robustness when there are erratic data values. The median polish is not guaranteed to converge. For this reason, the following stopping rule is often used in the algorithm: Whenever none of the entries in the table change by more than a given level at a given iteration, the algorithm is terminated. Removal of the trend prior to modeling the covariance structure receives mixed reviews from users of spatial techniques. However, failure to remove any underlying unplanned variation negates the validity of the variogram due to the stationarity assumption.

A practical application of the algorithm presented over would be analyzing yield data from GPS and a combine yield monitor and help the farmer to determine the effect of applied strips treatments (i.e. fertilizers, pesticides, herbicides etc.).

Reference:

Cressie, N. A. C. 1993. Statistics for Spatial Data. John Wiley & Sons, Inc., New York.pp690.

Selles, F., C.A. Campbell, B.G. McConkey, S.A. Brandt, and D. Messer 1999. Relationships between biological and chemical measures of N supplying power and total soil N at field scale. Can. J. Soil Sci. 79:353-366.