Locating Soil Monitoring Sites Using Spatial Analysis of Multilayer Data

Viacheslav I. Adamchuk[A], Luan Pan[A], David B. Marx[B], and Derrel L. Martin[A]

[A] Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska, USA.
[B] Department of Statistics, University of Nebraska-Lincoln, Lincoln, Nebraska, USA.
E-mail: vadamchuk@unl.edu

Abstract
Wireless technology is increasingly used to monitor temporal changes in soil conditions. Because of the limited number of such monitoring units that can be economically installed throughout an agricultural field, it is critical to identify proper locations for these units. On-the-go soil sensing technology provides the opportunity to rapidly obtain high-resolution data on soil spatial variability at a relatively low cost. Prescribing representative monitoring sites based on multiple sensor-based data layers is an important process, yet in practice, this selection is conducted in a very subjective manner. This research provides an analytical methodology for assessing the quality of selection of a set of strategic locations in an agricultural field. The methodology is shown using a combination of apparent soil electrical conductivity (EC_a) and field elevation maps to identify sites for water content monitoring.

Key Words
On-the-go soil sensing, targeted sampling, electrical conductivity, elevation, irrigation.

Introduction
The efficiency of a centre pivot irrigation system depends on its ability to meet the water demands of the growing crop (Sadler *et al.* 2005). While limited water supplies can reduce crop yield due to water stress, excessive irrigation can result in wasted resources and, if extreme, may also reduce yields. The optimum quantity of irrigation water changes temporally and spatially. Various methods have been used to focus on either level of variability. Thus, soil water monitoring and crop modelling often facilitate improved irrigation scheduling. On the other hand, dense-resolution proximal soil sensing allows a producer to identify the spatial variability of topsoil water storage capacity, which also affects the need for irrigation water. A combination of data layers with high spatial resolution can be used to define targeted field locations for soil water monitoring. The most appropriate strategy might be placement of a few sets of soil water sensors with wireless communication capability in these locations. Since water storage capacity depends on the properties of the soil profile and the potential for surface water runoff, high density measurement of apparent soil electrical conductivity (EC_a) and field elevation can be used to define field locations with different levels of water available to the crop during the growing season.

Selecting the appropriate number of strategic locations within a field is critical, but this process is typically subjective. Practitioners who use high-resolution data layers rely on the following general rules: 1) selected locations must cover the entire range of data from each source; 2) selected locations must avoid field boundaries and other transition zones; and 3) practitioners must spread locations over the entire field. While these "guidelines" are useful, they do not translate into an operational algorithm and, therefore, can produce numerous solutions that result in differing degrees of satisfaction. In principle, this process is similar to prescribing the targeted sampling locations needed to calibrate high-resolution data or to quantify the agronomic soil attributes of established management zones (Lesch 2005; Minasny and McBratney 2006; Brus and Heuvelink 2007; de Gruijter *et al.* 2008). The objective of this paper is to define a set of criteria that can be used to compare alternative schemes of soil sensor telemetry placement, basing these criteria on the three general rules used to analyze soil EC_a and field elevation maps.

Methods
Optimization Criteria
As previously mentioned, the definition of the optimum guided sampling scheme is quite vague. There are many parameters that can quantify 1) the spatial separation; 2) the spread across both sets of measurements; and 3) the local homogeneity within each set of measurements. Furthermore, there are several alternatives to deriving the overall objective function as a combination of these parameters. The goodness of spatial separation among selected sampling locations can be assessed by comparing the field areas represented by each sample; by comparing the variance of the error for an interpolated surface; or by calculating the

horizontal distance between each pair of locations. In this study, the S-optimality criterion (SAS 2008) was selected. It seeks to maximize the harmonic mean distance from each sampling location to all other sampling locations. Data spread across each of the two sets of measurements can be accessed using either the degree of variability within the selected layer of data or an information-based criterion, such as D-optimality (SAS 2008), which increases with greater coverage of the range of measurements. The D-optimality based on the assumption of a linear model (the proportional relationship between sensor outputs and agronomic parameters of interest) was selected. Finally, local homogeneity can be derived from the slope of a non-smooth surface constructed from all the measurements, or based on neighbourhood statistics. It was decided this project would consider only the immediate neighbours. Because of the potential for the anisotropic behaviour of sensor data, the selection of neighbours involved searching for the two nearest measurements from the same pass and the two from both neighbouring passes:

$$H_{cr} = 1 - \frac{\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left(z_i - z_j\right)^2}{\sum_{i=1}^{N} n_i \cdot H_{\max}} \quad (1)$$

where n_i is the number of existing nearest neighbours for the i^{th} location (n_i = 2 to 4); and H_{max} is the maximum value of $1\text{-}H_{cr}$ for the given dataset, obtained using N points with the greatest mean squared difference with neighbours.

To set the overall objective function (OF), the geometric mean of all criteria was selected. The D-optimality and local homogeneity criteria were calculated twice (for EC_a and for field elevation). Prior to multiplication, each layer was normalized with respect to the median of a large number (e.g., 100,000) of randomly selected sets of monitoring site locations. Because of cost constraints, the number of such locations was limited to nine, which nonetheless allowed the three locations to correspond to three levels (low, medium, and high) of EC_a and field elevation (even with poor correlation between them).

Sensor Data

To obtain high resolution maps of apparent electrical conductivity and elevation, a Veris® 3150 unit (Mobile Sensor Platform, Veris Technologies, Inc., Salina, Kansas) equipped with an RTK-level AgGPS® 442 GNSS receiver (Trimble Navigation Limited, Sunnyvale, California) was used to map a 37-ha field located at the University of Nebraska-Lincoln Agricultural Research and Development Centre near Mead, Nebraska, USA (Figure 1). Although the EC_a was mapped with two depths of investigation (specifically, 0-30 cm and 0-90 cm), only the shallow measurements were used in this research. Both EC_a and elevation data were collected with 1 Hz mapping frequency while moving at approximately 1.5 m/s travel speed with a 13.7 m swath width, which resulted in about 30 thousand data points.

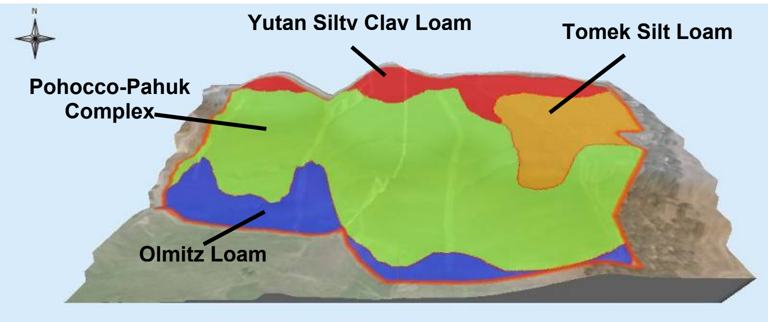

Figure 1. Field 1.14 at the University of Nebraska-Lincoln Agricultural Research and Development Centre (Mead, Nebraska).

A script written in Matlab® R2007a (The MathWorks, Inc., Natick, Massachusetts) allowed calculations of the OF for any subset of data points dedicated as potential monitoring site locations. However, to run a large number (100,000) of random combinations in reasonable time, the original dataset had to be reduced. This was accomplished by averaging a 20-m square grid. As a result, 859 locations with corresponding EC_a and elevation measurements were considered as potential sites for the installation of soil matric potential and temperature sensors, as well as wireless communication equipment (Figure 2). After initial runs of the script, the nine locations that provided the highest value of the OF were selected and manually adjusted to avoid field and waterway boundaries, pivot tracks, etc. The resulting locations were used to install soil sensors connected to the telemetry nodes for further research on optimized irrigation management. To evaluate the quality of our selection, each criterion and the OF were compared with corresponding values for 100,000 random selections of nine locations.

Results

Figure 3 illustrates maps of apparent soil electrical conductivity and field elevation, along with the set of locations selected to install temporal monitoring equipment. Visually, it appeared that these locations had decent field coverage that represented the entire range of both data layers (Figure 4).

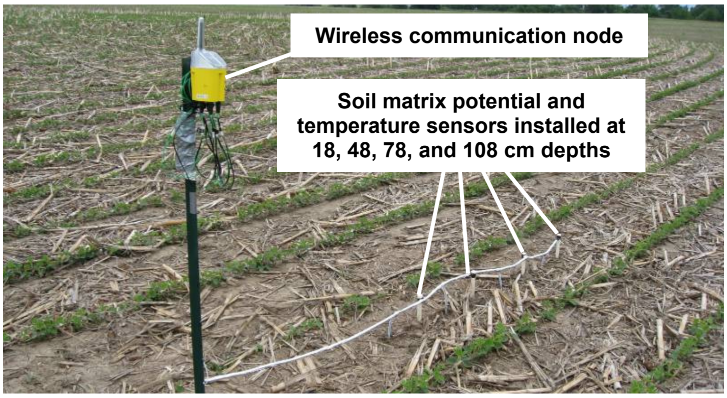

Figure 2. A telemetry soil water potential and temperature profile monitoring system installed in one of nine selected field locations.

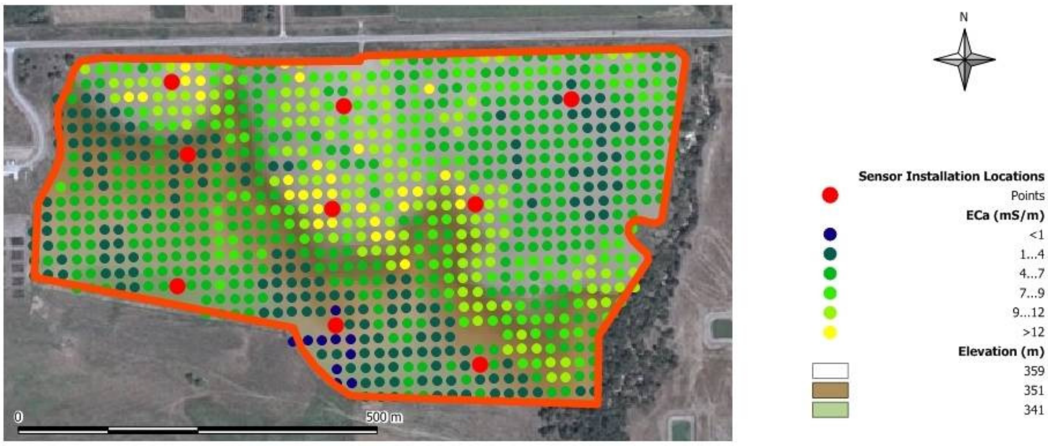

Figure 3. Map showing shallow (0-30 cm) apparent electrical conductivity, field elevations, and a set of targeted soil water monitoring locations.

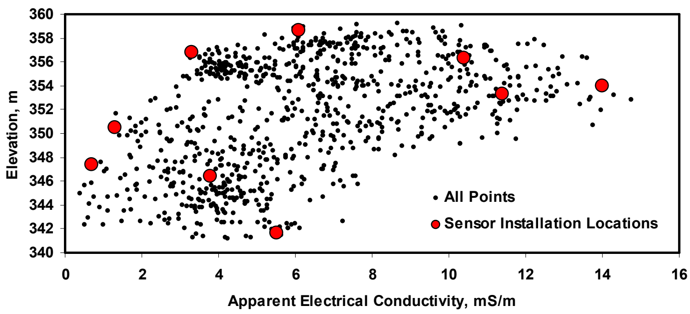

Figure 4. Relationship between apparent electrical conductivity and field elevation.

As shown in Figure 5, D-optimality (D_{opt}) was the criterion with the highest range and, therefore, the factor given the highest weight in formulating the OF. The homogeneity criterion (H_{cr}), on the other hand, had the least influence. From a practical point of view, this is a positive result, since the main objective of sensor placement is to cover the entire range of high-density measurements. Local homogeneity and proximity to field boundaries are more useful criteria for restricting rather than ranking potential placements.

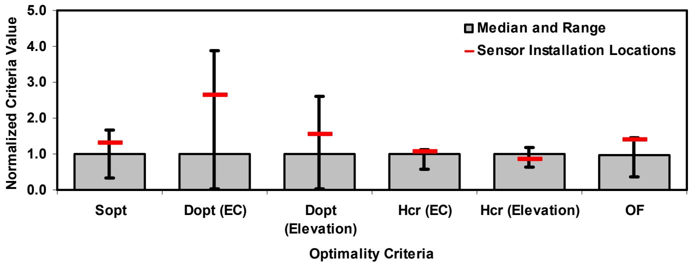

Figure 5. Median and range of individual normalized criteria and OF values for 100,000 random soil water monitoring selections, with actual installation and optimal sets of targeted points.

Although the set of monitoring locations with the highest value of the OF was one of 100,000 random selections, it would be impractical to locate monitors based on this set, because some locations were in restricted field areas (close to boundaries and pivot tracks). Avoiding such areas through manual intrusion naturally decreased the values of the criteria as well as the overall OF. Although this reduction was not significant, eliminating restricted parts of the field prior to the selection process might be appropriate in the future.

Although S_{opt} and D_{opt} criteria are highly popular in spatial statistics, our definition of H_{cr} is quite vague. The established criterion is well-defined for weighting power since only highly variable locations would reduce otherwise high OF values. However, this does not consider the overall neighbourhood statistics. A more involved homogeneity criterion will be considered in the future to restrict the consideration of relatively small areas with apparent measurement stability, and to allow the selection of field areas with a directional gradient of a given measurement.

Conclusion

In this research, we proposed a method for comprehensive evaluation of soil monitoring site locations with respect to multiple high-resolution sensor-based data layers. An objective function was developed representing the entire range of sensor data spread across the field and local homogeneity. Partially subjective selection of nine locations to install soil water potential and temperature monitoring equipment was shown to provide a relatively high OF value that was, however, less than the maximum value for 100,000 random selections. A more involved process is needed that would automatically select the most appropriate set of targeted sampling/monitoring locations.

Acknowledgement

This publication is a contribution of the University of Nebraska Agricultural Research Division, supported in part by funds provided through Nebraska Centre for Energy Science Research through the Water, Energy and Agriculture Initiative (WEAI).

References

Brus DJ, Heuvelink GBM (2007) Optimization of sample patterns for universal kriging of environmental variables. *Geoderma* **138**, 86-95.

de Gruijter JJ, McBratney AB, Taylor J (2008) Sampling for high resolution soil mapping. In 'Proceedings of the First Global Workshop on High Resolution Digital Soil Sensing and Mapping' (The University of Sydney: Australia, Sydney).

Lesch SM. 2005. Sensor-directed spatial response surface sampling designs for characterizing spatial variation in soil properties. *Computers and Electronics in Agriculture* **46**, 153-180.

Minasny B, McBratney AB (2006) A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences* **32**, 1378-1388.

SAS (2008) Optimality criteria. In 'SAS/QC User's Guide' (SAS Institute, Inc: Cary, North Carolina).

Sadler EJ, Evans RG, Stone KC, Camp CR (2005) Opportunities for conservation with precision irrigation. *Journal of Soil and Water Conservation* **60**, 371-379.