INTERPOLATION AND DATA COLLECTION ERROR SOURCES FOR ELECTROMAGNETIC INDUCTION–SOIL ELECTRICAL CONDUCTIVITY MAPPING

R. C. Wilson, R. S. Freeland, J. B. Wilkerson, W. E. Hart

ABSTRACT. Interpolation inaccuracies can occur when spatially mapping bulk soil conductivity (ECa). These inaccuracies are compounded when surveying large acreages using mobile equipment. As a mobile survey is typically more time extensive and covers more acreage than a pedestrian survey, allowances for temporal effects and instrument orientation are necessary.

To evaluate these inaccuracies, data were gathered to determine the influences of short–term conductivity shifts occurring over a data collection period, as well as from travel route patterns that determine instrument orientation. Using root–mean–squared error (RMSE) to quantify the transformation accuracy of interpolated ECa maps, a data collection method and an appropriate geostatistical model were determined for the loessial soils of West Tennessee.

Analysis showed that parallel travel paths, rather than crossing travel paths, helped to insure improved map quality; and interpolation inconsistencies were reduced when measurements were obtained in a manner by which all adjacent measurements were temporally contiguous as possible. For our specific application, ordinary kriging (OK) gave acceptable map transformation predictions, as did the inverse distance weighted (IDW) and radial basis function (RBF) interpolation models.

Keywords. Electrical conductivity, Interpolation, Inverse distance weighted (IDW), Kriging, Offset data tracking, Radial basis function (RBF).

Our research involves mapping subsurface moisture pathways within the loessial–over–alluvium soils of southwest Tennessee. Identifying these pathways help target possible agrochemical migration patterns beneath the surface. Towards this goal, we employ spatial maps of bulk soil electrical conductivity (ECa) to nonintrusively highlight major shifts in soil morphology and soil moisture regime patterns. Repeated short–interval surveys reveal that positional ECa measurements are dynamic and are multidimensional in regard to instrument orientation and time (Geonics, 1995; Wilson et al., 2003). Mapping patterns can vary depending upon the interpolation method and/or model employed. The data itself is influenced from the manner of data collection, primarily from the many possible travel routes that can be taken when traversing the field while gathering data.

In our research, we employ an EM31–MK2 (Geonics Ltd., Mississauga, Ontario, Canada) conductivity meter to nonintrusively measure ECa, which is rated by the manufacturers to operate to an optimal depth of 6 m. Actual measurement depth varies depending upon soil parameters. The measurement is a depth–weighted bulk average, whereby features as they occur farther beneath the surface have a lessening influence on the overall measurement. To probe shallower, the instrument can be rotated 90° along its longitudinal axis in order to sense at approximately half its optimal depth (Geonics, 1995). Subsurface transitions are detected by observing a relative change in ECa while traveling. Also, pivoting the instrument horizontally about a point near a subsurface morphology transition zone and observing maximum to minimum readings gives a sense of the boundary orientation (fig. 1). McNeill (1980) gives a complete description of the function and theory behind the EM31–MK2.

Figure 1. Rotating the EM–31 above any given point may generate a wide range of conductivity values, depending upon the subsurface morphological boundaries.
Measurements of ECa across large acreages are typically obtained with some form of mobile transport, such as a towed, cart-mounted conductivity meter described by Freeland et al. (2002) (fig. 2a). Data are sampled at closely spaced, geo-referenced points within the field. From this data set, Geographical Information Systems (GIS) software is used to create continuous-surface ECa maps using one of several interpolation models. Considerations of temporal effects and instrument orientation are necessary, as this mobile protocol is typically more time extensive and covers more acreage than a pedestrian survey whereby the instrument is handheld (fig. 2b).

Spatially interpolating discrete ECa data using different interpolation models and/or parameters can produce pattern inconsistencies, in part, due to a discrete point conceivably yielding a full range of ECa values due to instrument orientation (fig. 1). Varying readings at any given spatial point may also occur due to the instrument being raised or lowered above the point, instrument calibration drift, and from time-varying soil moisture content (Sudduth et al., 2001).

These observations of mapping inconsistencies have led to our examination of the available geostatistical interpolation models and survey methods. Thus, the scope of this project is to develop a protocol to determine an acceptable transformation from the discrete measured data to the interpolated surface. The standard for comparison is the most predictable agreement between the discrete field measurements (whether true or erroneous) and the resulting spatially interpolated surface. This required our field test to focus on a variety of driving pattern scenarios, as this affected both sample time and instrument orientation.

**LITERATURE REVIEW**

Englund (1990) performed a study where identical spatial data sets were given to 12 geostatistical investigators. Each investigator was asked to independently analyze the same data set and create spatial interpretations. Results varied considerably, in part due to the vast number of models, methods, and options that are available for producing spatial maps.

Goovaerts (2001) noted that no one interpolation model works well in all cases, but rather there exists a “toolbox of algorithms” from which to select appropriate methods. The selection of a particular interpolation model depends on characteristics of the data set as well as the study objectives. Individual models can contain numerous user-defined settings, methods, and variables that influence mapping transformation accuracy. With larger data sets, some are more processing-intensive and require more user-selected model parameters than other simpler, less robust models.

A review of the literature showed many interpolation model studies for the natural sciences. For example, researchers have studied the relationships between ordinary kriging (OK) and inverse distance weighted (IDW) for mapping soil nitrate (NO−3) and organic matter content for variable-rate fertilizer applications in corn production (Zea mays L.) on Midwest soils (Gotway et al., 1996). They found that OK provided reasonably accurate results in all cases. They also found that model accuracy was dependent upon the soil parameter being mapped. Burgess and Webster (1980a, 1980b) used punctual and block kriging to estimate soil properties for small and large blocks of land. They concluded that block kriging was more appropriate than punctual kriging in estimating average values over large areas. They also found kriging to be especially pertinent to physical properties associated with water in the soil. Bishop and McBratney (2001) evaluated the performance of multiple linear regression, OK, and Kriging with External Drift (KED) for mapping Cation Exchange Capacity (CEC) using the secondary variables of yield, ECa, elevation, and satellite images. They suggested that KED, with the use of secondary information such as ECa, could more accurately predict the CEC than could OK. Isaaks and Srivastava (1989) compared OK, IDW, and triangulation on several clustered data sets and found that OK produced the lowest prediction error in their applications. Niemann et al. (2001) compared real surface data to the radial basis function (RBF)-derivative Completely Regularized Spline (CRS) and fractal interpolation methods on topographic data in simulated river networks. The CRS method was viewed as the smooth interpolator, in this case, while the fractal method was a rough interpolator. They tested the ability of each method to estimate unobserved elevations, slopes, and curvatures as well as to simulate their distribution. The CRS interpolation produced better estimates of slope than the fractal method. They concluded that the CRS method was a good compliment to existing interpolation methods used in simulating river networks. Kravchenko and Bullock (1999) conducted a comparative study of various interpolation methods for mapping soil test P and K data. They found OK with the optimal number of neighboring points and an appropriate variogram performed better than IDW.

In reporting the sampling design and subsequent interpolation of soil electromagnetic induction data for precision agriculture applications, Corwin and Lesch (2003) suggested that depending upon the level of detail desired, from a hundred to several thousand spatial measurement of ECa...
should be gathered, typically in evenly-spaced traverses. Eigenberg and Nienaber (2003) used OK, because they employed the maps only for visual interpretation, and they found that OK gave visually appealing plots. Johnson et al. (2003) employed IDW to reveal that $EC_a$-classified within-field variance approximated the plot-scale experimental error found in field-scale research.

As suggested above, each interpolation application is often a function of individual prerogative and analytical skill. In fact, multiple approaches may be taken to explore different aspects of the data set. This study focused on three interpolation models found in ArcMap 8.3 Geostatistical Analyst (ESRI, Inc., Redlands, Calif.): IDW, OK, and RBF. However, the concepts used to evaluate the model results presented herein may be applied to other interpolation models and software platforms.

**Common Interpolation Models**

The IDW model creates a surface from measured points based on their similarity and distance. Weights are assigned to control points during interpolation, such that the influence of one point relative to another decreases with distance from the calculated point. As the power increases, the closer the value of the calculated point is to the nearest observed point (Isaaks and Srivastava, 1989). Inverse distance weighted gives reasonable results for many types of data as well as being easy to use and calculate. When using IDW, the choice of weighting function is difficult if there is a non-uniform spatial distribution of data points. Also, the occurrences of maximum and minimum values occur only within the range of measured data. This model is more likely to produce “bull’s eyes” around data points.

Due to its robustness and effectiveness, kriging has become almost synonymous with spatial interpolation among laypersons. However, the flexibility provided by certain kriging models may require extensive analyst decisions. Kriging derives its weights from variation patterns expressed within a semivariogram, whereby an optimal model (e.g., circular, exponential, logarithmic, Gaussian, etc.) is fitted by the analyst for calculating unmeasured points. The semivariogram illustrates the spatial correlation among measured and unmeasured points as a function of separation distance and directional angle within a search window. One method of kriging is OK, a flexible form of kriging where there are few assumptions, but as such can be less powerful than other kriging methods that require user-input parameter settings.

Radial basis functions are flexible and can handle regularly spaced or scattered data points. There are many forms of RBF, but this study focused on the CRS function. Pollution concentrations, elevation points, water table heights, and other gently varying surfaces are well suited for RBF interpolations (Johnston et al., 2001). The weight of a CRS point is defined by the third derivative in a curve minimization expression. The overall curvature of the surface is reduced and interpolated data are forced through a specified number of data points. Radial basis functions are simple to compute, requiring only the solution of linear equations (Hickernell and Hon, 1999).

**Mapping Transformation Accuracy**

The ultimate goal in generating spatial maps is supplying an accurate surface representation as provided by the measured control points. The resulting surface may not be of the “true or actual” value in reality, as it is only a projected representative surface that was construed solely from the measured data. A concept known as “jack-knifing,” or cross validation, removes each measured point one at a time. Its replacement value is then calculated using an interpolation model. The difference between the removed measured value and its predicted replacement value, repeated for all measured points within the data set, forms a predictor of overall mapping transformation accuracy of the projected surface. Various validation indices can be used as a measure of prediction quality, the most common of which are the root-mean-squared error (RMSE) and mean error (Bishop and McBratney, 2001).

**Objectives**

Our field experience suggested that three sources of mapping inconsistencies or transformation inaccuracies can occur due to (1) temporal soil conductivity shifts and instrument calibration drift over an extended data collection period, (2) instrument orientation due to driving pattern, and (3) from the application of a selected interpolation model itself. The objectives of this case study for our field site were to:

- Determine if $EC_a$ data collection driving patterns have an influence on the mapping consistency of interpolated surface $EC_a$ maps, and
- Evaluate common interpolation models (IDW, OK, RBF) as to the impact of each model on transformation accuracy when mapping $EC_a$.

**PROCEDURES**

**METHODS AND MATERIALS**

Bulk soil electrical conductivity data were gathered using a mobile system (fig. 2a) that allowed automated measurements over large acreages without a pre-established survey grid (Freeland et al., 2002). Using a differential global positioning system (DGPS), geospatial data were merged with $EC_a$ data using synchronized time stamps. Resultant data were imported into ArcView 8.3 (ESRI, Inc., Redlands, Calif.) for spatial mapping and interpolation purposes. A large concrete pad was installed in a shaded area that is adjacent to the site for calibration and monitoring instrument drift. The in-phase reading of the EM31 was calibrated over this pad immediately prior to each survey.

Using this survey method, two studies were conducted to determine the effect of data collection procedures on map transformation accuracy. The first study took place on a small 0.4-ka plot with instrument orientation as the variable. Separated by 24 h, two sets of data were collected (DAY_1 and DAY_2), each in a single driving event. Each full data set was subdivided into three subgroups of instrument orientations (fig. 3). Bidirectional pattern data were a subgroup extracted from the larger perpendicular data set. Unidirectional data were extracted from the bidirectional data set by grouping transect data that were traversed in a similar direction, approximately northeast (NE) and southwest (SW). The analysis did not include the end-of-row turn data.

Vol. 21(2):
The second study was in a nearby field of 6 ha, where both planned and non–planned driving patterns were evaluated (fig. 4). A non–planned driving pattern was driven by an operator using personal choice. The operator had no prior knowledge of the test results from the previous studies. The operator was instructed to canvas the entire field. Using personal choice, a typical encircling mowing pattern was selected. This pattern was time efficient, as it eliminated the need for end–of–row turning.

A second driving pattern, which was pre–planned, was then driven, whereby the operator relied on information gained from the previous study that predicted an optimal driving pattern for data quality. Parallel tracks were driven to attempt as much as possible spatially and temporally contiguous measurements of the same instrument orientation. Traveling this route took approximately 7% longer than traveling the circular route.

Each of the data sets from the two studies was evaluated using a standardized methodology. A single sector spherical search neighborhood that incorporated 10 surrounding data points determined the value at unsampled locations. The OK interpolation parameters included using a spherical semivariogram to estimate the weighting factor. A power of two was used to determine weighting values in the IDW interpolation.

The RBF interpolation used a CRS function to calculate unknown values.

**RESULTS AND DISCUSSION**

The results are presented in terms of “best” transformation accuracy, where accuracy is defined as the predictable agreement between ECa field measurements (whether true or erroneous) and the resulting spatially interpolated surface. The parameter RMSE was selected as the transformation accuracy indicator, as it is calculated for OK, IDW, and RBF. There is no assessment of prediction errors for IDW and RBF by ArcMap 8.3 Geostatistical Analyst (ESRI, Inc., Redlands, Calif.).

Figures 5 and 6 present the RMSE values for the small plot driving pattern tests illustrating influences due to data collection methods as well as interpolation models. Small differences were observed between instrument orientations in the unidirectional data, as they also had the smallest RMSE. The bidirectional data produced a low RMSE in the driving pattern tests, and it is also the most time–efficient driving pattern in that no “deadheading” for reorientation is required. The maps produced in the perpendicular survey as a group had higher RMSEs than the bidirectional survey.
Figure 3 presents each orientation sequence including the resultant interpolation maps using the three models as obtained from applying OK, IDW, and RBF models.

The second study determined that planned surveys produce lower RMSE maps than non–planned surveys. Visual interpretation of the data (fig. 4) indicates little variability between the maps produced by each model, but great visual variability occurs between the driving patterns. Figure 7 depicts lower RMSE values for the planned survey. Differences are apparent between RMSE values as compared between each of the models. Ordinary kriging produced lower RMSE results in the planned survey, while OK and IDW produced lower RMSE results in the non–planned survey.

**SUMMARY AND CONCLUSIONS**

This manuscript discusses the potential sources of errors when mapping discrete EC₄ measurements. The sources of mapping inconsistencies or transformation inaccuracies are from (1) instrument orientation due to driving pattern and

![Graph showing RMSE comparisons of instrument orientation in 0.4-ha plot on Day 1. Means with the same letter are not significantly different (α = 0.05, Duncan's Multiple Range Test).](image)

Vol. 21(2): 5
related EC_a anisotropy, and (2) from the application of a selected interpolation model itself. Mapping transformation accuracy was evaluated by comparing RMSE values for each of the interpolation models as well as the different data collection methods. Selection of an optimal surveying scenario, including post processing, was based upon minimizing RMSE as the goal. Using RMSE to quantify the transformation accuracy of EC_a maps, a data collection method and an appropriate geostatistical model were determined for the loessial soils of West Tennessee. The results are presented in terms of “best” transformation accuracy, where accuracy is defined as the predictable agreement between the discrete EC_a measurements (true or erroneous) and the resulting interpolated surface.

For this study site, the data collection−driving pattern was found to influence the mapping transformation accuracy of continuous surface EC_a maps, with a travel route providing a parallel orientation of the instrument supplying the smaller RMSE. Analysis showed that a bidirectional travel path helped to lower RMSE, as transformation inaccuracies were reduced when measurements were obtained in a manner that limited the directional influence of the EM31 orientation passing over subsurface transitions. The OK and IDW models demonstrated a trend of having lower RMSE values as compared to the RBF model; however, this trend was at times not statistically different (α = 0.05). Resulting interpolated maps from employing the OK, IDW, and RBF models were visually different.

**ACKNOWLEDGEMENTS**

Support for this project was provided by the United States Department of Agriculture−Water Resources Assessment Protection Program of the National Research Initiative, “Predicting Offsite Subsurface Migration of Agrochemicals — Noninvasive Surveying,” contract no. 99−35102−8273, and by the Trustees of the Hobart Ames Foundation, Ames Plantation, Grand Junction, Tennessee. The authors are grateful for the technical assistance of Mark Shiers and Melanie Hammond.

**REFERENCES**


**ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECa</td>
<td>bulk soil electrical conductivity</td>
</tr>
<tr>
<td>IDW</td>
<td>inverse distance weighted</td>
</tr>
<tr>
<td>OK</td>
<td>ordinary kriging</td>
</tr>
<tr>
<td>RBF</td>
<td>radial basis function</td>
</tr>
<tr>
<td>RMSE</td>
<td>root−mean−square error</td>
</tr>
</tbody>
</table>