Geostatistical mapping of soil conductivity and clay content beyond field boundaries

A. Brenning[1], S. Koszinski[1], M. Sommer[1]
1 Institute of Soil Landscape Research, Leibniz-Center for Agricultural Landscape Research (ZALF) e.V., Müncheberg, Germany
Corresponding author: brenning@zalf.de

ABSTRACT: Electromagnetic induction is a non-invasive method that is used to characterize soil properties of agricultural land in terms of the apparent electrical conductivity (ECa). ECa is a relative measure that cannot be directly compared between fields due to field-specific as well as temporal random effects. This restricts its interpretation and makes a fieldwise calibration with soil samples necessary if e.g. the spatial distribution of clay content is to be mapped. To overcome these limitations, we present a geostatistical framework that allows to map soil conductivity and physical soil properties across field boundaries. This approach consists of rescaling fieldwise ECa measurements to a synthetic, regional conductivity map, ECa(syn) that is continuous across field boundaries. Optimal rescaling parameters are estimated by minimizing an objective function that measures the continuity of ECa(syn) across the boundaries of adjacent fields, helping reduce calibration effort and laboratory cost. The proposed method is applied in a case study in Central Germany.

KEYWORDS: Spatial data homogenization, soil conductivity mapping, precision farming, spatial interpolation.

1. Introduction

Electromagnetic induction is a non-invasive mapping technique that is frequently applied in precision farming to map soil characteristics, to direct soil sampling and improve yield prediction (Johnson et al., 2001; Barnes et al., 2003; Sommer et al., 2003; Corwin and Lesch, 2005; Jung et al., 2005; Sudduth et al., 2005; Triantafilis and Lesch, 2005). Most of these studies are performed on the field scale since the absolute values of the apparent electrical conductivity (ECa) measured under different cultivation or soil moisture conditions cannot be directly compared (Sudduth et al., 2001). This is a limitation for ECa-directed soil sampling and mapping across field boundaries, especially in the case of physical soil properties such as clay content that are continuous in space independently of recent crop cultivation patterns.
To overcome the problem of discontinuity of the observed ECa signal, we propose a geostatistical approach that consists of rescaling the field-scale ECa measurements on adjacent fields to a common scale. This will allow to perform ECa-directed sampling and the calibration of empirical models of the distribution of soil properties across several fields, reducing soil sampling and analytical cost.
Conceptionally, our aim is to eliminate field-specific and time-dependent effects that lead to discontinuities between fields. Similar problems may occur in the statistical analysis of, for example, repeated observations in longitudinal or experimental studies. However, when ECa is mapped on adjacent fields, there are no repeated measurements in the strict sense, but highly correlated observations separated by small distances ("pseudo-replications").
We approach the problem from an interpolation point of view. We assume that we wish to map some soil property, for example clay content, that varies continuously in space and can be predicted from ECa. Our aim is to produce a synthetic ECa surface that is continuous

across field boundaries but still honors the spatial distribution pattern of ECa observed on each field. As long as there are discontinuities between fields, an extrapolation of ECa from one field to an adjacent one will result in errors with respect to the ECa values observed there. Therefore our approach consists of minimizing this extrapolation or discontinuity error by rescaling field-scale ECa measurements.
We will demonstrate the applicability of the proposed method in a case study in Central Germany. Since the "true" data-generating distribution is not known in the case of real data, we perform a simulation study to generalize the results and show the validity of our method.

2. Methods

2.1. Introduction

The starting point of our approach can be formulated as the assumption that the ECa surface would be continuous in space if the random transformations that arise from field-specific cultivation history or time-dependent (weather-dependent) variation were absent. Our aim is to reconstruct this "true", undisturbed conductivity surface from the actual ECa field measurements. Equivalently, we could be interested in homogenizing the data from adjacent fields to obtain a "synthetic" ECa map that can be used to model physical soil properties across field boundaries.

2.2. Model formulation

Formally, we consider the synthetic ECa surface as the realization of a real-valued random field $Z = (Z(x))_{x \in D}$ defined on a domain $D \subset R^2$. However, we only observe random fields $Y^{(1)},...,Y^{(p)}$ on p fields $D_1,...,D_p \subset D$. These fields will generally be non-overlapping, but present some degree of adjacency. The random fields are related to Z by a possibly nonlinear transformation f,

$$Y^{(i)}(x) = f_i(Z(x)), \quad x \in D_i, i = 1,..., p,$$

representing field-specific variation. Random field $Y^{(i)}$ is observed at n_i locations $x^{(i)}{}_1,...,x^{(i)}{}_{n_i} \in D_i$. In this work, we use a parametric representation of f, specifically a linear relationship

$$Y^{(i)}(x) = a_i + b_i Z(x), \quad x \in D_i, a_i \in R, b_i \in R^+.$$

By convention, we set f_1:=id (identity), or $a_1 = 0$, $b_1 = 1$, i.e. all transformations are relative to the observed ECa values on one selected field.
The scaling parameters a_i and b_i cannot be estimated directly because there is no overlap between adjacent fields. However, we can predict (extrapolate) the ECa values from a field to its neighbors. We use the mean-squared difference between ECa values extrapolated from one field and the ECa measurements on the other field as a measure of the discontinuity of the ECa surface. We will minimize this *discontinuity error* to estimate optimal scaling parameters. To account for the increase of uncertainty as the extrapolation distance increases, we apply a weighting scheme that is based on the kriging variance.

2.3. Study area and field measurements

The study area near Wulfen (Sachsen-Anhalt, Central Germany) is characterized by late Quaternary deposits including glacial till, solifluidal slope deposits and fluvial deposits of the

former ice marginal valleys. These sources of material control the spatial variation of soil properties within fields and on the landscape scale.
In the case study, we use temperature-corrected ECa measurements conducted with the EM38 in vertical mode on five adjacent fields with a total area of 121 ha. After preprocessing the data (local outlier filtering, median-aggregation of ECa measurements less than 1 m apart), 11090 ECa values were available for the fields, ranging between 12 and 106 mS m^{-1}.
Finger probe estimates of soil texture are available for 86 locations on one of the five fields (Kiesberg field) for soil horizons up to a depth of 100-150 cm. For the statistical analysis, clay contents were averaged across the soil horizons using the signal density of the EM38 in vertical mode as a weight function.

2.4. Simulation study

The simulation study is designed to assess the rescaling approach. We use a set-up representing a squared domain of 1 km², which is divided into four equally-sized squared fields. The baseline log-scale soil electrical conductivity is simulated as a Gaussian random field with mean value 1.6, standard deviation 0.075 and a spherical covariance structure with range 300 m. This baseline log-ECa surface is the target surface that is to be recovered by the proposed homogenization technique. To the baseline ECa surface, we add a short-ranged random field to represent field-specific local noise.
This log-ECa surface of fields 2, 3 and 4 is linearly scaled relative to field 1 using random transformation parameters. We simulate the additive (resp. multiplicative) parameters from independent normal distributions with mean 0 (1) and standard deviation 0.1 (0.1).
We simulate the log-ECa random field on two sets of locations. First, we use a squared grid with 20 m spacing between points as ECa measurement locations. Second, we generate 500 randomly distributed test points that we use to estimate the overall interpolation error. Between the fields, we use a separation distance of 20 m between ECa measurements on adjacent fields.
The rescaling parameters are estimated from the simulated data by numerically minimizing the root-mean-squared discontinuity error between fields using local kriging. We use 200 independent simulations of log-ECa fields to estimate the scaling parameters and assess the kriging error.
In addition to the discontinuity error defined above, we consider a second error type in the simulation study. The *interpolation error* of an interpolator trained on data is estimated by the root-mean-squared error of the interpolator's predictions compared to the baseline ECa data at the test points.
We compute the interpolation error for the interpolator trained on the baseline data itself (i.e. for known "true" scaling parameters), for the interpolator corresponding on the synthetic log-ECa signal rescaled with the estimated scaling parameters, and for the interpolator fitted to the untransformed measurements.

3. Results and Discussion

3.1. Simulation study

Discontinuity and interpolation errors for all simulations are shown in figure 1. The discontinuity errors corresponding to our rescaling approach using the estimated scaling parameters are very similar to those achieved with the true scaling parameters. If no homogenization is applied, the discontinuity error is significantly higher and shows a wide spread. This can also be observed in the case of the interpolation error.

The simulation study shows that the scaling approach is able to eliminate the interpolation errors that are introduced by linear random effects in ECa data of adjacent fields. Our scaling approach outperforms the interpolation of unscaled data in terms of interpolation error.

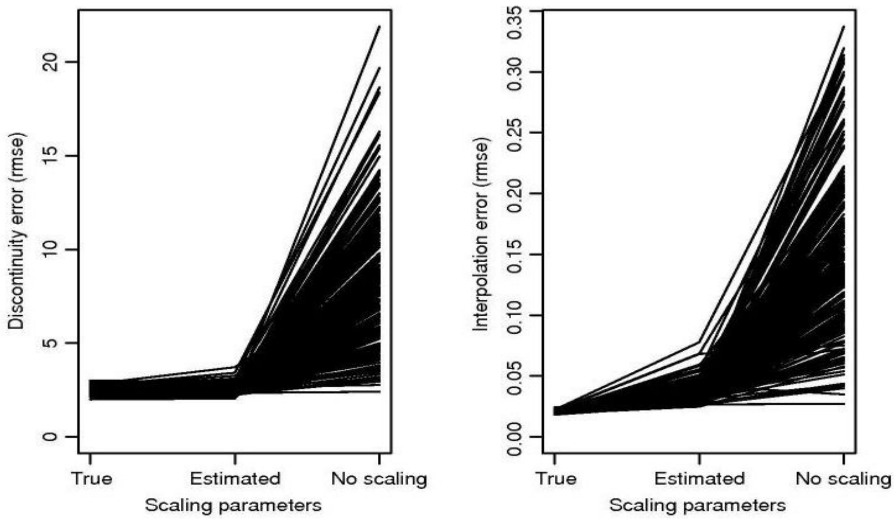

Fig. 1. Parallel coordinates plot of discontinuity (left) and interpolation error (right) using the "true" scaling parameters, the estimated parameters and no scaling.

3.2. Case study

The numerical optimization reduced the discontinuity error from 22.4 without rescaling by 17% to 18.6 after optimal rescaling. Figure 2 visualizes the distribution of the partial discontinuity errors between the fields and the associated weights. The estimated (log-scale) scaling coefficients range between 1.11 and 1.53 for the multiplicative parameter and between -0.13 and 0.15 for the additive parameter.

The magnitude of the estimated parameters reveals that perturbations of measurements on adjacent fields may have significant effects. Our approach removes these perturbations that are unrelated to soil properties that continuously vary across field boundaries.

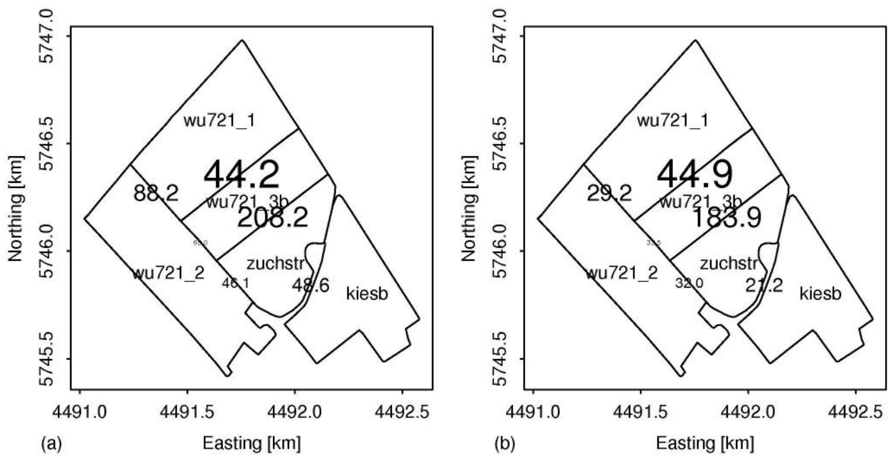

Fig. 2. Partial discontinuity errors of kriging across field boundaries in the case study. (a) Interpolation without rescaling, (b) interpolation after homogenization of ECa values. Type size is proportional to the weight given to each between-field discontinuity error based on the kriging variance.

The clay content of a soil influences soil humidity and hence crop yield. Since it has a strong effect on ECa, we use the latter to map the overall clay content of the soil column. Clay content is a bounded variable taking values between 0 and 100%; therefore we use logistic regression to model the relationship between the ECa(syn) field homogenized across field boundaries, and the weighted average of the clay content across soil horizons. In our case study there is a good agreement between predicted and observed clay content for the 86 locations on the Kiesberg field (Pearson correlation: 0.91). This empirical relationship is used to predict clay content for all five fields based on the homogenized ECa(syn) field.

4. Conclusions

The linear scaling approach presented in this work was able to reduce the discontinuity error of kriging soil conductivity across field boundaries, which is a prerequisite for mapping soil properties such as clay content continuously on adjacent fields. This data homogenization approach may help reduce the cost of ECa-directed sampling and mapping beyond the field scale.

Acknowledgments: This work is part of the research project pre agro II on Precision Agriculture, which is funded by the German Federal Ministry of Education and Research, grant no. 0339740/2. Soil texture analyses and ECa data were provided by the pre agro I project.

REFERENCES

Barnes, E. M., Sudduth, K. A., Hummel, J. W., Lesch, S. M., Corwin, D. L., Yang, C., Daughtry, C. S. T. & Bausch, W. C., 2003. Remote- and ground-based sensor techniques to map soil properties. *Photogrammetric Engineering & Remote Sensing* (69): 619–630.

Corwin, D. L. & Lesch, S. M., 2005. Apparent soil electrical conductivity measurements in agriculture. *Computers and Electronics in Agriculture* (46): 11–43.

Johnson, C. K., Doran, J. W., Duke, H. R., Wienhold, B. J., Eskridge, K. M. & Shanahan, J. F., 2001. Field-scale electrical conductivity mapping for delineating soil condition. *Soil Science Society of America Journal* (65): 1829–1837.

Jung, W. K., Kitchen, N. R., Sudduth, K. A., Kremer, R. J. & Motavalli, P. P., 2005. Relationship of apparent soil electrical conductivity to claypan soil properties. *Soil Science Society of America Journal* (69): 883–892.

Sommer, M., Wehrhan, M., Zipprich, M., Weller, U., Castell, W., Ehrich, S., Tandler, B. & Selige, T., 2003. Hierarchical data fusion for mapping soil units at field scale. *Geoderma* (112): 179–196.

Sudduth, K. A., Drummond, S. R. & Kitchen, N. R., 2001. Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture. *Computers and Electronics in Agriculture* (31): 239–264.

Sudduth, K. A., Kitchen, N. R., Wiebold, W. J., Batchelor, W. D., Bollero, G. A., Bullock, D. G., Clay, D. E., Palm, H. L., Pierce, F. J., Schuler, R. T. & Thelen, K. D., 2005. Relating apparent electrical conductivity to soil properties accross the north-central USA. *Computers and Electronics in Agriculture* (46): 263–283.

Triantafilis, J. & Lesch, S. M., 2005. Mapping clay content variation using electromagnetic induction techniques. *Computers and Electronics in Agriculture* (46): 203–237.