Geostatistical analysis of salinity indicators in a coastal reclamation area of China

She Dongli^{1,2}, Fei Yuanhang¹, Chen Qian¹, Luis Carlos Timm³

¹Key Laboratory of Efficient Irrigation-Drainage and Agricultural Soil-Water Environment in Southern China, Ministry of Education, College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China

²Jiangxi Provincial Key Laboratory of Soil Erosion and Prevention, Jiangxi Institute of Soil and Water Conservation, Nanchang 330029, China

³Faculty of Agronomy, Federal University of Pelotas, Department of Rural Engineering, Pelotas, RS, Brazil

*Corresponding author, email: shedongli@hhu.edu.cn

Introduction

This document presents a geostatistical analysis of data that was used in the paper: “Spatial scaling of soil salinity indices along a temporal coastal reclamation area transect in China using wavelet analysis” authored by She Dongli, Fei Yuanhang, Chen Qian, and Luis Carlos Timm, and was published in the Archives of Agronomy and Soil Science to which it is linked.

Full details of the justification for the study and of the study site and wavelet analysis of three salinity indicators are given in the main paper. In short, coastal reclamation is being used worldwide by numerous coastal nations in order to create valuable land resources. Factors that affect salinity are spatially distributed, not least because the reclamation has to be carried out incrementally, which naturally creates temporal gradients along which the properties of the reclaimed soil vary. In this study, soil pH, electrical conductivity ($EC_{1:5}$) and sodium ion content (SIC) were used as indicators of salinity in order to assess the degree of reclamation along a temporal transect (Figure 1 below and in the manuscript) and to investigate relationships among them and factors that could affect them.

While the main paper used wavelet analysis to identify factors that affected the salinity indicators at different scales, this document presents a simpler geostatistical analysis of the same data, for comparison and/or complementary purposes. It should be noted that such an analysis might not be suitable for data series that are non-stationary, as are many soil data series ([Kravchenko et al. 1999](#); [Seuront et al. 1999](#); [Nielsen & Wendroth 2003](#)). Furthermore, the analysis only determines the distributions of the parameters at one scale, indicated by the range value derived from a fitted semivariogram.

Geostatistical analysis

The analysis was conducted according to the method described by [Nielsen and Wendroth 2003](#). Semivariograms were used to evaluate the spatial dependence of the indicators and to

discover the range values. Based on the regionalized variable theory and intrinsic hypotheses a semivariogram can be expressed as:

$$\gamma(h) = \frac{1}{2N(h)} \sum [z(x) - z(x+h)]^2 \quad (1)$$

where $\gamma(h)$ is the semivariance, h is the lag distance, $N(h)$ is the number of sampling couples of the variable z in the interval h , $z(x)$, and $z(x+h)$ are values of z at positions x and $(x+h)$ (Feng et al. 2004). The obtained empirical semivariograms were fitted to theoretical semivariogram models to produce the geostatistical parameter range. The software, GS⁺ (Gamma Design Software 2004), was used for this purpose.

An autocorrelation analysis was also conducted.

Results

An autocorrelation analysis was used to identify how a variable was spatially correlated with itself. The autocorrelation functions (ACF_s) of the saline-sodic soil indicator properties are shown in Figure 2. A t-test was used to illustrate their spatial autocorrelation ranges at 95% confidence intervals. The significant autocorrelation range for the three salinity indicators was 2 lags, which indicated that the current sampling interval was sufficient to identify the spatial autocorrelation between adjacent observations of each index.

The models and fitted semivariogram parameters are presented in Table 1. Note that Ln(EC_{1:5}) or Ln (SIC) were used because of the lognormal distributions of EC_{1:5} and SIC. In all cases, the model well fitted the data (Figure 3 and Table 1). Soil pH and SIC were both better fitted by Gaussian models ($R^2 > 0.9$) than by the spherical model used to fit the EC_{1:5} data. The nugget-to-sill ratios of 51.3%-61.8% indicated moderate spatial dependence at the transect measurement scale of the three salinity indicators (Nielsen & Wendroth 2003). The nugget contribution to the total variance (nugget + sill) was about 49% in the case of the SIC but was lower in the cases of soil pH (~38%) and EC_{1:5} (~39%). The nugget variance was due

to random variation, which is typically attributed to human activity (Liu et al. 2012) but along the study transect both sources of variation were greatly influenced by the human activity involved in reclamation. The process of reclamation was systematic along the transect and involved making sequential changes to the soil derived from sediments in the land-sea boundary region (Figure 1), which might have been affected by some differences in their original sources and mode of deposition. Similarly, the land use type along the transect came under a greater degree of human management as it evolved from salt meadow in the more recently reclaimed areas to become farmland under intense cultivation in the older areas. Therefore, the process of reclamation likely induces more systematic than random errors.

The range values for all the indicators were higher than the measurement interval (285 m) used for sampling. Therefore, according to the range value derived from the semivariograms, the sampling scheme was sufficiently intense to detect spatial variations at the transect scale.

However, when compared with the large scales derived from the decomposed wavelet by the wavelet analysis presented in the main paper, the range value determined by geostatistical analysis would not have taken in account processes or variability occurring at those scales that were involved in reclamation. The d_4 scale exceeded the range values for pH and SIC determined by the geostatistical analysis indices, while the a_4 scale does so for all three of the salinity indicators. This might be attributed to the non-stationary nature of the soil properties in the reclamation area, and geostatistical analysis is not well suited to such data.

Conclusions

Geostatistics could identify the main scale, but only that scale, at which much of the variability in the salinity indicators could be explained. It could assist in identifying the processes involved with reclamation at different points along the transect. However, the wavelet analysis identified multiple scales at which the salinity indicators were influenced and helped identify the factors that were influencing their spatial variability at those scales. Some

of these factors would operate at scales much greater than the measurement scale or the range identified by the geostatistical analysis. This shortcoming of the geostatistical analysis was attributed to the likely non-stationary nature of the soil data series for which geostatistical analysis is not well suited.

References

- Feng Q, Liu YS, Mikami M. 2004. Geostatistical analysis of soil moisture variability in grassland. *J Arid Environ.* 58:357–372.
- Kravchenko AN, Boast CW, Bullock DG. 1999. Multifractal analysis of soil spatial variability. *Agron J.* 91: 1033-1041.
- Liu ZP, Shao MA, Wang YQ. 2012. Large-scale spatial variability and distribution of soil organic carbon across the entire Loess Plateau, China. *Soil Res.* 50:114-124.
- Nielsen DR, Wendroth O. 2003 *Spatial and temporal statistics: Sampling field soil and their vegetation.* Verlag: Catena.
- Seuront L, Lagadeuc FSY, Schertzer D, Lovejoy S. 1999. Universal multifractal analysis as a tool to characterize multiscale intermittent patterns: Example of phytoplankton distribution in turbulent coastal waters. *J Plankton Res.* 21: 877-922.

Table 1 Geostatistical analysis of saline-sodic soil parameters in the coastal reclamation region

	Nugget	Structured variance	Sill	Nugget/Sill (%)	Range (m)	Class (model)	R ²
pH	0.035	0.056	0.091	61.8	8574	Gaussian	0.92
EC _{1:5} ^a	0.154	0.238	0.392	60.7	10390	Spherical	0.81
SIC ^a	0.132	0.140	0.272	51.3	6720	Gaussian	0.93

EC_{1:5}, electrical conductivity of a 1:5 soil to water extract; SIC, sodium ion concentration. R², squared correlation coefficient between measured and estimated values calculated by leave-one-out cross-validation for ordinary kriging.

*Fitted by Ln(EC_{1:5}) or Ln (SIC) values because of the lognormal distribution of EC_{1:5} and SIC.

Figure1 Location of the study site and of the west-east transect in the coastal reclamation area in China. Sites A to F represents the six farmland areas reclaimed at different times (A: 2007; B: 1981; C: 1960; D: 1950; E: 1940; F: 1916)

Figure 2 Autocorrelation functions (ACF_s) for three saline-sodic soil parameters in the coastal reclamation region. *Calculated by $\ln(EC_{1:5})$ or $\ln(SIC)$ values because of the lognormal distribution of $EC_{1:5}$ and SIC. $EC_{1:5}$, electrical conductivity of a 1:5 soil to water extract; SIC, sodium ion concentration.

Figure 3 Experimental semi-variograms (symbols) and fitted models (lines) for three saline-sodic soil parameters in the coastal reclamation region. * Calculated using $\text{Ln}(EC_{1:5})$ or $\text{Ln}(SIC)$ values because of the lognormal distribution of $EC_{1:5}$ and SIC . $EC_{1:5}$, electrical conductivity of a 1:5 soil to water extract; SIC , sodium ion concentration.