

Chapter 18

Predictive Mapping of Soil Organic Matter at a Regional Scale Using Local Topographic Variables: A Comparison of Different Polynomial Models

Xiao-Dong Song, Gan-Lin Zhang and Feng Liu

Abstract Borrowing the idea of software engineering, this paper aimed to evaluate the mapping accuracy of soil organic matter (SOM) content from the “black box” perspective by combining regression kriging (RK) with local terrain attributes calculated by different polynomial models. When calculating local terrain attributes, we applied two neighborhood shapes (square and circular) and six frequently used algorithms (Evans-Young, Horn, Zevenbergen–Thorne, Shary, Shi, and Florinsky). Overall, 35 combinations of first- and second-order derivatives were produced as secondary information for RK. For comparison, the ordinary kriging (OK), ordinary cokriging (COK), and universal kriging (UK) were also utilized to map the SOM spatial distribution. The results of the study showed that the RK application outperforms OK, COK, and UK in improving the prediction quality of SOM content in a region where the soil properties were strongly influenced by the toposequence and the altitude was with a wide range. The most accurate mapping result was obtained by the combination of the Evans-Young algorithm and Zevenbergen–Thorne algorithm for the calculation of first- and second-order derivatives, respectively. The mapping results from the higher-order approach (Zevenbergen–Thorne and Florinsky) yielded less prediction errors and the circular-neighborhood method could enhance some algorithms for the calculation of local terrain attributes.

Keywords Digital soil mapping · Soil organic matter · Regression kriging · Local terrain attributes · Environmental correlation

X.-D. Song · G.-L. Zhang (✉) · F. Liu
State Key Laboratory of Soil and Sustainable Agriculture,
Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China
e-mail: glzhang@issas.ac.cn

© Springer Science+Business Media Singapore 2016
G.-L. Zhang et al. (eds.), *Digital Soil Mapping Across Paradigms, Scales and Boundaries*, Springer Environmental Science and Engineering,
DOI 10.1007/978-981-10-0415-5_18

18.1 Introduction

In the past thirty years, significant advances have been made in information technology, especially in Geographic Information System (GIS), remote and proximal sensors, and digital elevation models (DEMs), which have significantly boosted the vitality of soil science (McBratney et al. 2003). Taking DEM as an example, a number of fundamental topographic attributes have been proposed to quantitatively identify landform classes and features within geomorphology (Wilson 2012), and thus, diverse algorithms are presented focusing on specific goals and scenarios. Much attention, therefore, has been devoted to predict soil properties by using the terrain attributes. A large number of studies have shown that prediction methods incorporating these pieces of secondary information outperform generic geostatistical models (e.g., ordinary kriging) (Bishop and McBratney 2001).

Some of topographic attributes are distinguished from non-local or regional parameters, and hence are referred to as local terrain attributes, which are derived directly from DEMs without additional inputs and usually calculated by moving a three-by-three window (Behrens et al. 2010; Florinsky 1998; Shary et al. 2002; Wilson 2012), such as slope, curvature, roughness, and elevation percentile. After a traversal across DEM, a new grid with the same dimension will be produced, whose cells are each filled with a calculated value of land surface parameter. For morphometric variables, the terms local and non-local are usually used regardless of the study scale or DEM resolution and associated with the mathematical sense of a particular variable (Florinsky 2011).

Among local terrain parameters, slope and aspect, twelve kinds of curvatures (Shary 1995) are also called first- and second-order derivatives, respectively, as they are defined by the formulae depending on the first- and second-order partial derivatives of altitudes. Multifarious mathematically modeling methods have been developed to calculate these derivatives from a gridded DEM focusing on various landscapes (Evans 1980; Horn 1981; Minár et al. 2013; Shary 1995; Shary et al. 2002; Shi et al. 2007; Zevenbergen and Thorne 1987). As the accuracy of the variables is unavoidably influenced by the DEM data and calculation algorithms, numerous studies have been published to estimate the accuracy of these algorithms (Schmidt et al. 2003; Warren et al. 2004), analyze the relationships between errors of derived parameters with DEM data characteristics (Chang and Tsai 1991; Gao 1997), and compare computed slope gradients with actual field measurements (Bolstad and Stowe 1994; Warren et al. 2004). Nevertheless, none of those studies is within the context of soil mapping and their results are hardly applicable to knowledge-based digital soil mapping (Shi et al. 2012). The selection approaches of terrain attributes also have not received the attention they deserve in soil science literature (Behrens et al. 2010).

The purpose of this research was to evaluate the mapping performance of soil organic matter (SOM) that results from RK technique combined with local terrain attributes based on different polynomial models. Nine terrain attributes were calculated from grid DEMs: elevation, topographic wetness index (TWI), slope,

aspect, plan curvature, profile curvature, tangent curvature, maximal curvature, and minimal curvature. The local terrain attributes were derived from six quadratic and Lagrange polynomials and two types of neighborhood shapes. Among the six algorithms, the Evans-Young algorithm (Evans 1980; Young 1978), the Horn algorithm (Horn 1981), and the Shary algorithm (Shary 1995) are based on a quadratic polynomial, and the Zevenbergen–Thorne algorithm (Zevenbergen and Thorne 1987), the Shi algorithm (Shi et al. 2007), and the Florinsky algorithm (Florinsky 2009) are based on a Lagrange polynomial. At the beginning of interpolation, Pearson correlation and partial correlation analyses were performed to scan the relations between SOM and all variables. We then compared the results of ordinary kriging (OK), ordinary cokriging (COK), universal kriging (UK), and regression kriging (RK). Furthermore, we discussed the combination of local terrain variables for RK which achieved acceptable quality for predicting the spatial variation of SOM contents and potentially other soil properties.

18.2 Materials and Methods

18.2.1 Data

The study area, the upper and middle reaches of the Heihe River Basin, is located along the northeast margin of the Qinghai-Tibetan Plateau in China at the intersection of the Tibetan Plateau, the Inner Mongolia-Xinjiang Plateau, and the Loess Plateau (Fig. 18.1). With the geographical boundary of about $97^{\circ}20'$ – $101^{\circ}51'E$ and $37^{\circ}41'$ – $39^{\circ}59'N$, this area stretches for 340 km from the northwest to the southeast at a width between 115 and 180 km. Soil sampling was conducted in July to August 2012, including regular sampling and purposive sampling (Zhu et al. 2008) based on the concept of soil–environment relationships. A total of 223 topsoil (0–20 cm) samples recorded in above collections were compiled in a digital database. These data points were randomly split into calibration (80 %; $n = 178$) and validation (20 %; $n = 45$) datasets using the subset function of Geostatistical Analyst in ArcGIS (ESRI 2010).

18.2.2 Calculation of Local Terrain Variables and Other Related Terrain Variables

Terrain variables used in this study for the estimation of SOM were slope gradient, slope aspect, profile curvature, maximal curvature, minimal curvature (Table 18.1), elevation, and TWI. SRTM DEM data were employed and geo-referenced from three-arc second resolution to $90\text{ m} \times 90\text{ m}$ resolution. The principal differences among most algorithms for the computing of local terrain variables are the number

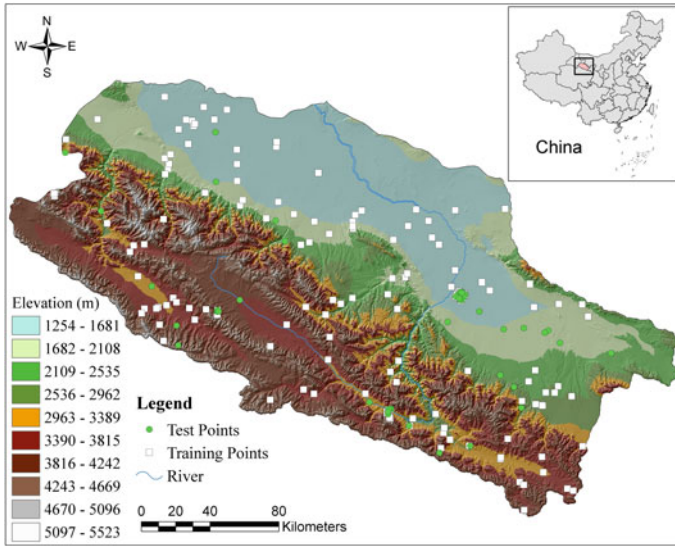


Fig. 18.1 Location of the study area and distribution of soil sampling sites

Table 18.1 Descriptive statistics of measured soil organic matter stock (SOM), log-transformed SOM (LnSOM), of the study area

| Variables | SOM g kg ⁻¹ | LnSOM g kg ⁻¹ |
|------------------------------|---------------------------|-----------------------------|
| Mean | 33.61 | 2.99 |
| Median | 19.44 | 2.97 |
| Minimum | 1.21 | 0.19 |
| Maximum | 269 | 5.96 |
| Standard deviation | 41.49 | 1.02 |
| Coefficient of variation (%) | 140.10 | 34.11 |
| Coefficient of skewness | 3.91 | 0.19 |
| Coefficient of kurtosis | 20.06 | 0.28 |

SOM the soil organic matter stock; *LnSOM* log-transformed soil organic matter stock

of grid cell used and the weight given to each of those cell values. In general, most algorithms utilize some elevation values in a three-by-three window centered on the elevation cell in question, so that one can find all the unknown coefficients for a polynomial. However, a three-order polynomial should be fitted over all points in a 5 × 5 neighborhood for approximation of all the coefficients (Florinsky 2009; Minár et al. 2013).

The approximations for regular grid DEMs used were bivariate second-, third-, and partial fourth-order polynomials. In this paper, the first- and second-order terrain attributes were selectively calculated using the circular and square neighborhood, which resulted in a total of 39 layers (14 first-order derivatives and 25 s-order derivatives). The first-order derivatives, slope and aspect, were computed

by seven algorithms: the Horn, Zevenbergen–Thorne, and Florinsky algorithms using the square neighborhood, the Shi and Evans-Young algorithms using both square and circular neighborhood. Five kinds of curvatures (Table 18.1) were achieved by five algorithms: the Zevenbergen–Throne, Shary, and Florinsky algorithms with square neighborhood, and the Evans-Young algorithm with both square and circular neighborhood. Hence, 35 combinations of first- and second-order derivatives were grouped. All combinations were incorporated into the multiple linear regression of RK, so as to test which group would yield the best performance. The formula of aforementioned variables could be found in literatures (Florinsky 2011; Horn 1981; Shary 1995; Shary et al. 2002; Shi et al. 2007; Zevenbergen and Thorne 1987).

For convenience, in the rest of this paper, a specific terrain variable and all attributes with the same order are abbreviated to “*Variable _ Method _ Neighborhood*” and “*Method*” + “*n*” + “*Neighborhood*,” respectively, where *n* is the order of local topographic attributes. For example, Slp_EY_C denotes the slope gradient using circular neighborhood and the Evans-Young algorithm; FY_2_Q is the second derivatives calculated by the Florinsky algorithm with square neighborhood. Most of the layers were generated by the Terrain Analysis function of ArcSIE[®], and other algorithms were implemented in C++ using GDAL library.

18.2.3 Methods

Four geo-statistical methods were involved in this study, including ordinary kriging (OK), cokriging (COK), universal kriging (UK), and regression kriging (RK). As a most general and widely used method of kriging, OK was employed to characterize the spatial variation of SOM and map overlays. If an interpolation is merely based on sample dataset, OK is commonly applied. OK uses the spatial correlation structure of the dataset to calculate weights for linear prediction from known points. Therefore, this method may require dense sample data for an interpolation with reasonable accuracy. In addition to OK, UK, COK, and RK are hybrid interpolation methods in which the variation of soil properties is quantified by deterministic and stochastic (empirical) models and can incorporate one or more ancillary variables in the estimation.

Cross-validation procedure was conducted to evaluate the accuracy of different models through three statistical measurements of the prediction error. The accuracy of estimates was assessed by the mean absolute error (MAE), the root mean squared errors (RMSE), and mean relative error ratio of performance to deviation (RPD). These indices were derived according to Eqs. (18.1), (18.2), and (18.3), respectively:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n [|Z^*(x_i) - Z(x_i)|] \quad (18.1)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n [Z^*(x_i) - Z(x_i)]^2} \quad (18.2)$$

$$\text{RPD} = \frac{\text{STD}}{\text{RMSE}} \quad (18.3)$$

where $Z(x_i)$ is the observed value of Z at locations x_i , $Z^*(x_i)$ the predicted value at the same location, n the number of samples, and STD the standard deviations of the variable. MAE and RMSE were used to estimate the accuracy of the predictions which should be as low as possible for accurate interpolation. The RPD was employed so as to interpret the prediction ability of each model.

18.3 Results

18.3.1 Exploratory Data Analysis

The summary statistics for SOM and log-transformed SOM (LnSOM) are presented in Table 18.1. The observed SOM content in surface soils varied from 1.21 to 386.00 g kg⁻¹, with a mean value of 34.61 g kg⁻¹. The coefficient of variation (CV) was 140.10 g kg⁻¹, indicating that SOM for all samples had a very large variability. The value of skewness was 3.91 g kg⁻¹, suggesting that samples had a positively skewed distribution (Fig. 18.2a). The Kolmogorov–Smirnov (K–S) test (p -value = 0.000 < 0.05) rejected the null hypothesis of normality for samples. The SOM stock data were transformed by natural logarithm to create an

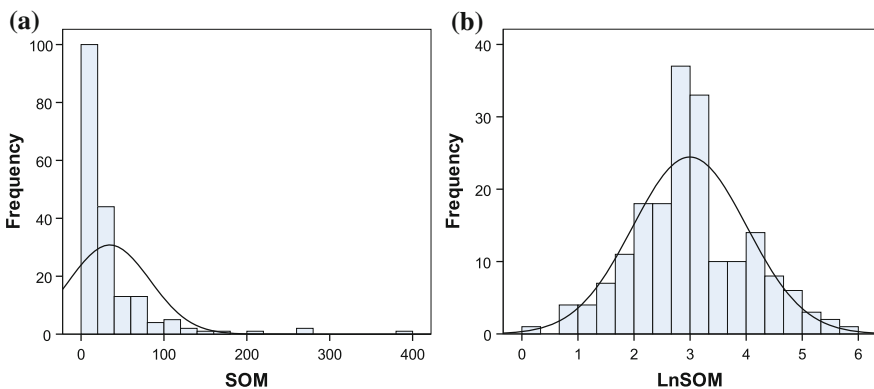


Fig. 18.2 Histogram of raw (a) and processed (b) datasets of SOM

approximately normal distribution, with mean (2.99 g kg^{-1}) and median (2.97 g kg^{-1}). Coefficients of skewness and kurtosis of lognormal SOM stock dropped from original values to 0.19 and 0.28 g kg^{-1} , respectively. Finally, the prediction values of SOM were back-transformed to original units.

Pearson's correlation analysis was carried out to explore the relationship between LnSOM and the terrain attributes based on the Evans-Young algorithm using square neighborhood (Table 18.2). These correlations were significant at the 0.01 level, suggesting that topography has important impacts on the distribution of SOM. The step-wise regression therefore was executed, aiming to derive the best subset of predictor variables and reduce the number of predictors (Table 18.2).

18.3.2 Prediction Accuracy of Different Kriging Methods

The aforementioned 45 validation datasets were used to assess the performance of different kriging methods with elevation, TWI, and local terrain attributes (Table 18.3). The prediction accuracy of SOM in this study was improved using RK with various combinations of local topographic attributes. The smallest and the largest prediction errors were produced by RK(EY1S_ZT2S) and RK(FY1S_EY2S), respectively. Compared with the worst method, the MAE and RMSE produced by RK(EY1S_ZT2S) method decreased by 6.46 g kg^{-1} and 20.23 g kg^{-1} , respectively, and the MRE increased by 0.84. The results of validation indicated that the combination of EY1S_ZT2S for the deriving of the local terrain attributes could remarkably improve the prediction accuracy of SOM prediction in this study area. RK(HN1S_ZT2S) also achieved a considerable accuracy, while the Horn and Zevenbergen–Throne algorithms might be the most widely used to calculate the first- and second-order derivatives due to the integration of main-stream GIS software. In the case of MAE, no values were close to zero, suggesting that there was a biased prediction. The RMSE values were slightly smaller than the standard deviations of the soil sample values (41.49 for SOM), and most of the RPD values were larger than 1.4. The inclusion of more auxiliary information in the RK regression models significantly improved the prediction performance.

Another important finding was that the performances of RK method whose second-order terrain attributes (SI1S_FY2S, ZT1S_FY2S, and SI1C_FY2S) were calculated by the third-order polynomial (Florinsky 2009) method outperformed most of the RK combinations and other kriging methods. Simultaneously, all the RPD values of RK combinations with FY2S, EY2C, and ZT2S were greater than 1.4, whereas the combinations with SA2S and EY2S were smaller than 1.4. Among RK results, RK with EY2S achieved the poorest performance, whereas all the RK with EY2C produced acceptable errors ($\text{RPD} > 1.4$). For all RK combinations, the circular neighborhood did not perform consistently better than the square neighborhood. This confirmed the previous conclusion (Shi et al. 2007) that the circular-neighborhood method may be more advantageous when used together with a specified neighborhood size, especially on a high-resolution DEM.

Table 18.2 The Pearson correlations between the logit-transformed soil organic matter content (LnSOM) and terrain variables calculated by the Evans-Young method

| | Elev | TWI | Asp_EY_S | Slp_EY_S | MaC_EY_S | MiC_EY_S | PiC_EY_S | PrC_EY_S | TaC_EY_S |
|----------|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| LnSOM | 0.625 ^{***} | -0.166 [*] | 0.266 ^{***} | 0.445 ^{***} | 0.205 ^{***} | -0.347 ^{***} | -0.093 | 0.139 | -0.187 [*] |
| Elev | 1 | -0.274 ^{***} | 0.318 ^{***} | 0.563 ^{***} | 0.451 ^{***} | -0.263 ^{***} | -0.281 ^{***} | -0.038 | -0.306 ^{***} |
| TWI | | 1 | -0.205 ^{***} | -0.530 ^{***} | -0.390 ^{***} | 0.017 | 0.412 ^{***} | 0.022 | 0.328 ^{***} |
| Asp_EY_S | | | 1 | 0.307 ^{***} | 0.183 [*] | -0.175 [*] | -0.036 | -0.043 | -0.114 |
| Slp_EY_S | | | | 1 | 0.439 ^{***} | -0.297 ^{***} | -0.421 ^{***} | 0.123 | -0.633 ^{***} |
| MaC_EY_S | | | | | 1 | 0.028 | -0.651 ^{***} | -0.532 ^{***} | -0.548 ^{***} |
| MiC_EY_S | | | | | | 1 | -0.368 ^{***} | -0.536 ^{***} | -0.065 |
| PiC_EY_S | | | | | | | 1 | 0.091 | 0.802 ^{***} |
| PrC_EY_S | | | | | | | | 1 | -0.050 |
| TaC_EY_S | | | | | | | | | 1 |

Elev elevation; *TWI* topographic wetness index; *Asp* slope aspect; *Slp* slope aspect; *Slp* slope gradient; *MaC* maximal curvature; *MiC* minimal curvature; *PiC* plan curvature; *PrC* profile curvature; *TaC* tangent curvature; *EY* Evans-Young algorithm; *S* square neighborhood

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

Table 18.3 Assessment of the various methods for predicting soil organic matter

| Methods | MAE | RMSE | RPD | Methods | MAE | RMSE | RPD |
|---------------|-------|-------|------|---------------|-------|-------|------|
| RK(EY1S_ZT2S) | 15.50 | 23.08 | 1.80 | RK(EY1S_EY2C) | 18.03 | 28.68 | 1.45 |
| RK(SI1S_FY2S) | 16.25 | 24.85 | 1.67 | RK(SI1S_EY2C) | 16.98 | 28.73 | 1.44 |
| RK(ZT1S_FY2S) | 16.97 | 26.36 | 1.57 | RK(SI1C_SA2S) | 19.53 | 37.41 | 1.11 |
| RK(SI1C_FY2S) | 16.98 | 26.39 | 1.57 | RK(ZT1S_SA2S) | 19.55 | 37.45 | 1.11 |
| RK(HN1S_ZT2S) | 17.28 | 27.40 | 1.51 | RK(SI1S_SA2S) | 20.10 | 38.29 | 1.08 |
| RK(EY1C_ZT2S) | 17.26 | 27.47 | 1.51 | RK(EY1C_SA2S) | 20.53 | 38.41 | 1.08 |
| RK(FY1S_ZT2S) | 17.31 | 27.73 | 1.50 | COK | 19.15 | 38.85 | 1.07 |
| RK(FY1S_FY2S) | 17.43 | 27.78 | 1.49 | RK(HN1S_SA2S) | 20.77 | 38.99 | 1.06 |
| RK(EY1C_EY2C) | 17.00 | 27.81 | 1.49 | RK(ZT1S_EY2S) | 21.08 | 39.32 | 1.06 |
| RK(HN1S_EY2C) | 17.16 | 27.86 | 1.49 | RK(SI1C_EY2S) | 21.09 | 39.35 | 1.05 |
| RK(EY1C_FY2S) | 17.50 | 27.92 | 1.49 | RK(FY1S_SA2S) | 20.80 | 39.47 | 1.05 |
| RK(EY1S_FY2S) | 18.40 | 28.06 | 1.48 | OK | 21.29 | 39.52 | 1.05 |
| RK(HN1S_FY2S) | 17.56 | 28.07 | 1.48 | RK(EY1S_SA2S) | 21.23 | 39.73 | 1.04 |
| RK(FY1S_EY2C) | 17.19 | 28.28 | 1.47 | UK | 20.71 | 40.29 | 1.03 |
| RK(SI1S_ZT2S) | 17.19 | 28.35 | 1.46 | RK(EY1C_EY2S) | 21.32 | 40.84 | 1.02 |
| RK(ZT1S_ZT2S) | 16.95 | 28.44 | 1.46 | RK(SI1S_EY2S) | 21.49 | 41.47 | 1.00 |
| RK(ZT1S_EY2C) | 16.69 | 28.48 | 1.46 | RK(HN1S_EY2S) | 21.68 | 41.76 | 0.99 |
| RK(SI1C_ZT2S) | 16.96 | 28.49 | 1.46 | RK(EY1S_EY2S) | 22.47 | 42.60 | 0.97 |
| RK(SI1C_EY2C) | 16.72 | 28.55 | 1.45 | RK(FY1S_EY2S) | 21.96 | 43.31 | 0.96 |

RK regression kriging; *OK* ordinary kriging; *COK* cokriging; *UK* universal kriging; *MAE* mean absolute error; *RMSE* root mean squared error; *RPD* ratio of performance to deviation. *EY* Evans-Young algorithm; *ZT* Zevenbergen–Throne algorithm; *HN*: Horn algorithm; *SA* Shary algorithm; *SI* Shi algorithm; *FY* Florinsky algorithm; *S* square neighborhood; *C* circular neighborhood

It is clearly seen that RK produced the SOM maps with more marked fluctuation than those of OK, COK, and UK (Figs. 18.3 and 18.4), especially when the maps were draped over the DEM they were based on. The obvious differences between the SOM maps generated by RK and other three kriging methods were the predicted SOM values in south part of study area (Qilian Mountain). The maps produced by RK showed more details of SOM content in spatial variation, which convincingly indicated the significant influences of toposequence as only the terrain attributes were used within the multiple linear regression.

One of the overall aims of this study was to compare the accuracies of SOM maps derived from RK with various combinations of local terrain attributes. Different from quantitative surface analysis (Jones 1998; Zhou and Liu 2004), it is an application-specific scenario for the mapping of SOM in regional area where the topography undulates greatly. Different combinations of first- and second-order derivatives provide diverse SOM maps due to their describing abilities of the general geomorphometry of land surface. In common with one of the objectives of geomorphometry, to a certain extent, digital soil mapping aims to quantitatively describe and model the variation of soil properties in terrestrial ecosystem. This

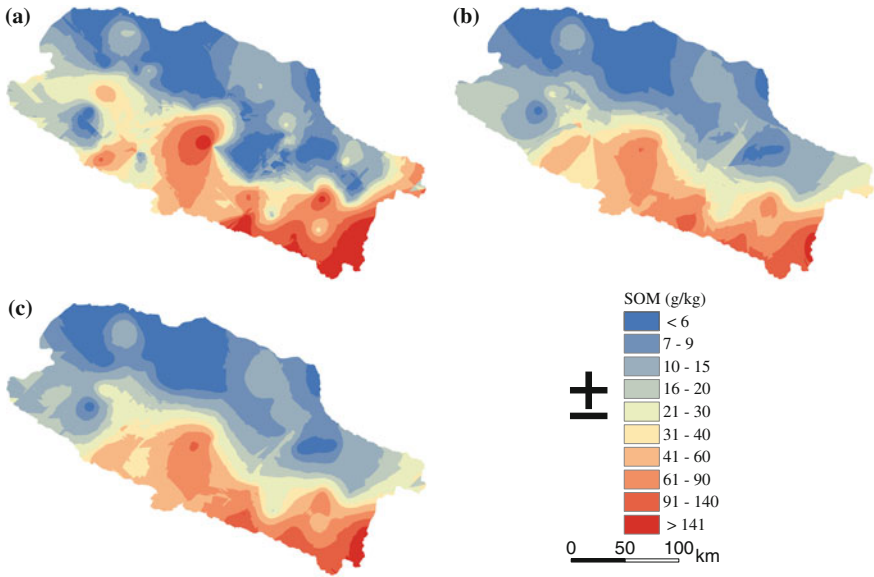


Fig. 18.3 The spatial predictions of soil organic matter content (g kg^{-1}) by universal kriging (a), ordinary kriging (b), and ordinary cokriging (c)

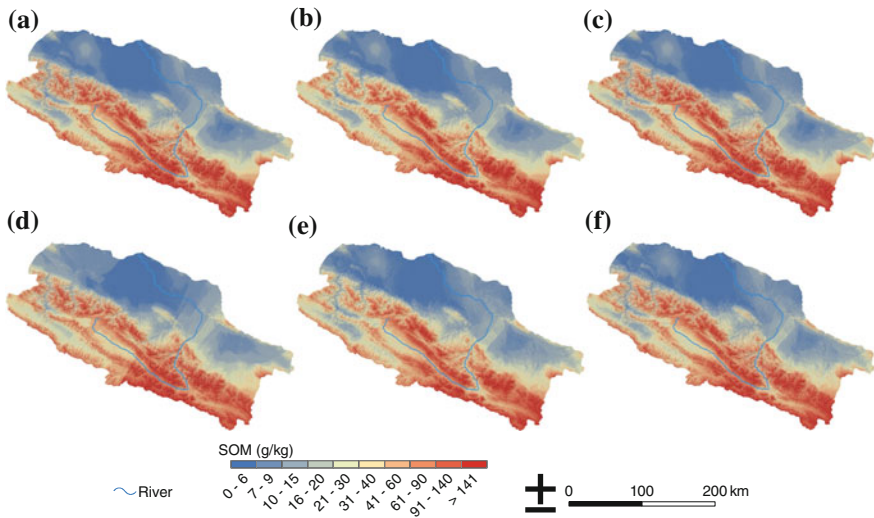


Fig. 18.4 Predicted soil organic matter (SOM) maps using regression kriging. *Note* The prediction values are draped over the DEM they are based on

quantitative description could be seemed as a scientific approach to evaluating the land surface modeling, which is reflected directly by the correlations between topographic variables and soil properties. Generally, it is confirmed especially when

the soil patterns are not affected by the agriculture and other anthropogenic activities.

It is helpful to arrive at a conclusion that we could achieve an optimal combination of first- and second-order derivatives based on disparate algorithms rather than the same algorithm. Other contrastive studies of polynomial models also found that modeling results from higher-order approaches show higher sensitivity to local variations (Florinsky 2009; Schmidt et al. 2003), such as the Zevenbergen–Throne algorithm and the Florinsky algorithm. This was coincided with the results of cross-validation listed above. There were 8 and 14 RK methods whose second-order derivatives were calculated by the Zevenbergen–Throne and Florinsky algorithms in the top 10 and 20 combinations. The main advantage of the Florinsky algorithm is the local denoising by approximating the polynomial to elevation values of the 5×5 window which could enhance the calculation of partial derivatives. Likewise, the modified Zevenbergen–Thorne algorithm with circular-neighborhood method is more sensitive to noise in the DEM, whereas the square-neighborhood method is less sensitive (Shi et al. 2007).

18.4 Conclusions

The contrast results of the current study could be deemed as the benchmark of different algorithms of local topographic variables. Nevertheless, it does not mean that the best method for the calculation of local parameters in this study will outperform others with different spatial resolutions and neighborhood sizes, especially when the DEM datasets are generated variously due to the vital accuracy of DEM. Comparing with traditional application, we can conclude that the performance of predictive methods that can incorporate auxiliary variables might be improved by using the same local terrain variable calculated by different methods. However, although the “black box” approach of digital soil mapping is working in hindsight, a more accuracy soil map of large poorly accessible area or difficult terrain might be achieved, which takes up a little time and energy rather than high sampling costs. In conclusion, our findings are important to select the algorithms of local morphometric variables for the RK technique or other prediction methods especially for the high-relief sites. Our study also provides a promising approach to choose the ancillary variables for mapping the spatial variation of other soil properties.

Acknowledgments The study was supported financially by the National Natural Science Foundation of China (grant No. 41130530, No. 91325301, No. 41401237) and partly by the Jiangsu Province Science Foundation for Youths (No. BK20141053). The authors are grateful to Prof. Ian S. Evans for sharing his published materials and to Dr. Florinsky for the interpretation of his method.

References

- Behrens T, Zhu AX, Schmidt K, Scholten T (2010) Multi-scale digital terrain analysis and feature selection for digital soil mapping. *Geoderma* 155: 175-185.
- Bishop TFA, McBratney AB (2001) A comparison of prediction methods for the creation of field-extent soil property maps. *Geoderma* 103: 149-160.
- Bolstad PV, Stowe T (1994) An evaluation of DEM accuracy: elevation, slope and aspect. *Photogrammetric engineering and remote sensing* 60: 1327-1332.
- Chang KT, Tsai BW (1991) The effect of DEM resolution on slope and aspect. *Cartography and Geographic Information Systems* 18: 69-77.
- ESRI. (2010). ArcView and ArcInfo. Version 10.0. California, Redlands: ESRI.
- Evans IS (1980) An integrated system of terrain analysis and slope mapping. *Zeitschrift für Geomorphologie, N.F., Supplementband* 36: 274-295.
- Florinsky IV (1998) Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science* 12, 47-62. Florinsky, I.V., Eilers, R.G., Manning, G.R., Fuller, L.G., 2002. Prediction of soil properties by digital terrain modelling. *Environmental Modelling & Software* 17: 295-311.
- Florinsky IV (2009) Computation of the third-order partial derivatives from a digital elevation model. *International Journal of Geographical Information Science* 23: 213-231.
- Florinsky IV (2011) *Digital terrain analysis in soil science and geology*. Elsevier Academic Press, Amsterdam, p. 7-16.
- Gao J (1997) Resolution and accuracy of terrain representation by grid DEMs at a micro-scale. *International Journal of Geographical Information Science* 11: 199-212.
- Horn BKP (1981) Hill shading and the reflectance map. *Proceeding of the IEEE* 69: 14-47.
- Jones KH (1998) A comparison of algorithms used to compute hill slope as a property of the DEM. *Computers & Geosciences* 24: 315-323.
- McBratney AB, Mendonça Santos ML, Minasny B (2003) On digital soil mapping. *Geoderma* 117: 3-52.
- Minár J, Jenčo M, Evans IS, Minár J, Kadlec M, Krcho J, Pacina J, Burian L, Benová A (2013) Third-order geomorphometric variables (derivatives): definition, computation and utilization of changes of curvatures. *International Journal of Geographical Information Science* 27: 1381-1402.
- Schmidt J, Evans IS, Brinkmann J (2003) Comparison of polynomial models for land surface curvature calculation. *International Journal of Geographical Information Science* 17: 797-814.
- Shary PA (1995) Land surface in gravity points classification by complete system of curvatures. *Mathematical Geology* 27: 373-390.
- Shary PA, Sharaya LS, Mitusov AV (2002) Fundamental quantitative methods of land surface analysis. *Geoderma* 107: 1-32.
- Shi X, Zhu AX, Burt J, Choi W, Wang RX, Pei T, Li BL, Qin CZ (2007) An experiment using a circular neighborhood to calculate slope gradient from a DEM. *Photogrammetric Engineering and Remote Sensing* 73: 143-157.
- Shi X, Girod L, Long R, DeKett R, Philippe J, Burke T (2012) A comparison of LiDAR-based DEMs and USGS-sourced DEMs in terrain analysis for knowledge-based digital soil mapping. *Geoderma* 170: 217-226.
- Warren SD, Hohmann MG, Auerswald K, Mitasova H (2004) An evaluation of methods to determine slope using digital elevation data. *Catena* 58: 215-233.
- Wilson JP (2012) Digital terrain modeling. *Geomorphology* 137: 107-121.
- Young M (1978) *Terrain analysis: program documentation. Report 5 on Grant DA-ERO-591-73-G0040, 'Statistical characterization of altitude matrices by computer'*. Department of Geography, University of Durham, England. 27 pp.
- Zevenbergen LW, Thorne CR (1987) Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* 12: 47-56.

- Zhou Q, Liu, X (2004) Analysis of errors of derived slope and aspect related to DEM data properties. *Computers & Geosciences* 30: 369-378.
- Zhu AX, Yang L, Li B, Qin C, English E, Burt JE, Zhou CH (2008) Purposive sampling for digital soil mapping for areas with limited data. In: Hartemink AE, McBratney AB, Mendonca Santos ML (Eds.), *Digital Soil Mapping with Limited Data*. Springer-Verlag, New York, pp. 233-245.