Soil & Tillage Research 192 (2019) 1–11

Contents lists available at ScienceDirect

Soil & Tillage Research

journal homepage: www.elsevier.com/locate/still

Spatial analysis of soil aggregate stability in a small catchment of the Loess Plateau, China: II. Spatial prediction

Luping Ye[a,c], Wenfeng Tan[a,b,*], Linchuan Fang[a], Lingling Ji[b]

[a] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling 712100, China

[b] Key Laboratory of Arable Land Conservation (Middle and Lower Reaches of Yangtze River), Ministry of Agriculture, Huazhong Agricultural University, Wuhan 430070, China

[c] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Keywords:
Soil aggregate
Soil erodibility
Co-kriging

ABSTRACT

As indicators of soil degradation vulnerability, soil aggregate stability indices play important roles in representing soil resistance to water erosion, and their spatial variability provides both agriculturally and environmentally important information. The spatial variability of aggregate stability indices is synergistically affected by the soil, topography, vegetation, and human factors. To understand the formation processes of aggregates by a spatial analysis, a prediction model combining soil properties with natural and human factors should be developed to improve the accuracy of the spatial interpolation of soil aggregate stability indices. In this study, the mean weight-diameter (MWD, mm), water-stable aggregates greater than 0.25 mm ($WSA_{>0.25}$, %) and soil erodibility factor (K factor) were satisfactorily predicted by multiple stepwise regression (MSR) and regression kriging (RK) based on soil properties and natural and human factors ($0.436 \leq R^2 \leq 0.578$). In addition, spatial variability and prediction modeling of aggregate stability indices were highly dependent on the quantification of land use type and landscape structure (the spatial structure of landscape elements and the connections between the different ecosystem types or landscape elements). It has received little attention in previous studies. The exclusion of all soil variables did not affect the predictions of K factor, and for MWD and $WSA_{>0.25,}$ even though the performance of the models may appear relatively low, but also significant ($0.183 \leq R^2 \leq 0.312$), indicating that the prediction of the spatial distributions of aggregate stability indices with easily available auxiliary data is practicable and effective. Residual maps showed that high residuals are distributed around built-up land (transportation land and residential land) or farmland, indicating that anthropogenic factors increase the uncertainty of the models. The spatial distribution maps of MWD, $WSA_{>0.25}$ and K factor can be useful in landscape planning and decision making to minimize water erosion risks.

1. Introduction

Soil aggregate stability indices are indicators of soil structure and play a key role in assessing soil vulnerability to degradation. The mean weight-diameter (MWD, mm), water-stable aggregates greater than 0.25 mm ($WSA_{>0.25}$, %) and soil erodibility factor (K factor) are often used to represent the soil aggregate stability (Veihe, 2002). Their spatial variability depends on the interactions of natural ecological processes and intensive human activities, including soil properties, land use type and landscape structure, topography, hydrothermal conditions, and vegetation cover. Landscape structure represents the spatial structure of landscape elements and the connections between the different ecosystem types or landscape elements. Data on land use type and landscape structure, topography, hydrothermal conditions, and vegetation cover can be obtained in a more economical way from remote sensing data or existing databases. Hence, they are widely used as auxiliary data to accurately predict the spatial distribution of soil properties (Ou et al., 2017). It is of great interest to know the spatial distribution of MWD, $WSA_{>0.25}$ and K factor for various purposes such as the evaluation of soil degradation vulnerability. Analysis of the relationship between the spatial heterogeneity of these aggregate stability indices and natural and human factors may facilitate a better understanding of the reasons for the spatial variability of these indices. Our previous research revealed that spatial analysis showed great potential

* Corresponding author at: Key Laboratory of Arable Land Conservation (Middle and Lower Reaches of Yangtze River), Ministry of Agriculture, College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China.
E-mail address: tanwf@mail.hzau.edu.cn (W. Tan).

https://doi.org/10.1016/j.still.2019.03.009
Received 27 June 2018; Received in revised form 13 February 2019; Accepted 17 March 2019
0167-1987/ © 2019 Elsevier B.V. All rights reserved.

to be applied in the analysis of the influencing factors of soil aggregate stability indices. Moreover, the strong spatial variability of soil aggregate stability parameters were confirmed (Ye et al., 2018). Hence, it is necessary to further explore the impacts of potential influencing factors (including soil properties, natural factors, and human factors) on MWD, $WSA_{>0.25}$ and K factor.

Prediction methods have been used to interpolate aggregate stability indices from sampling sites to continuous distributions. These methods include a variety of statistical regression methods (Besalatpour et al., 2013; Jaksik et al., 2015) and geostatistical methods (Annabi et al., 2017; Mummey et al., 2010). Multiple linear regression (MLR), regression kriging (RK), and co-kriging (CK) are typical soil-landscape modeling methods and widely used to map soil properties by combining soil property data with auxiliary environmental variables (Mummey et al., 2010; Ou et al., 2017). A number of researchers used MLR for mapping soil properties due to its simplicity and efficiency (Grunwald, 2009; Zhao et al., 2017). Regression kriging, one of the most extensively accepted and used interpolation methods that combines the auxiliary variables, has received increasing attention in the mapping of soil properties (Hengl et al., 2004). Regression kriging includes a multivariate linear model between a soil variable and environmental variables and the residual kriging of the soil variable. Co-kriging is developed to improve the interpolation of a variable by using the information of other spatially correlated variables in spatial interpolation process. A number of studies have confirmed the superiority of RK and CK in comparison with many other mapping methods (Hengl et al., 2004; Li et al., 2013; Mulder et al., 2011). Besalatpour et al. (2013) used different mathematical methods to predict the MWD based only on six specific variables, which may directly or indirectly affect the aggregate stability. Jaksik et al. (2015) employed multiple linear regressions (MLR) to predict aggregate stability indices based on the topographical and soil data, and Annabi et al. (2017) only used soil property data for predicting aggregate stability indices. Although their results provided satisfactory predictions, they ignored the effects of other important environmental factors, such as landscape structure, temperature difference between seasons, aridity, and vegetation information, which are the key factors in the formation and stabilization of aggregates (Amezketa, 1999).

Soil aggregate stability is largely determined by the land use type and landscape structure, temperature and humidity (Amezketa, 1999). In previous studies, land use has often been regarded as a categorical variable and has been rarely taken into consideration. In the process of modeling, temperature and humidity are also rarely taken into account even though these two variables are quantitative indicators. From the landscape ecology perspective, the landscape metrics can be used as an indicator of landscape structure. Remote sensing data can provide the information of temperature and humidity in a relatively efficient way. The commonly used indices are land surface temperature (LST) and temperature-vegetation dryness index (TVDI) (Ma et al., 2017; Sandholt et al., 2002). Therefore, it is expected that the introduction of landscape metrics, LST and TVDI into the prediction will facilitate a better understanding of the relationships of landscape structure and hydrothermal conditions with soil aggregate stability.

We previously explored the spatial structures of aggregate stability indices by semivariograms, local indicators of spatial association, ordinary kriging, and inverse distance weighted in Zhifanggou catchment, a typical hilly gullied loess landscape of LPC, and manifested the strong spatial variability structure of soil aggregate stability (Ye et al., 2018). To understand the formation processes of aggregates clearly, a prediction methodology that integrates the soil properties and natural and human factors should be developed to explore the spatial correlations between soil aggregate stability and the potential influencing factors. Pearson's correlation analysis and canonical correspondence analysis (CCA) method were employed to explore the relationships between soil aggregate stability indices and soil properties, landscape characteristics, topography, vegetation cover, and hydrothermal conditions. Spatial cross-correlation was further used to analyze their spatial correlations, which can enable us to understand the major controls on the formation processes of aggregates within ecosystems deeply (Jia et al., 2011). Multiple linear regression and RK were used to interpolate the spatial distribution of these indices in detail. This study aims to (i) investigate the correlations of the composition and spatial configurations (landscape metrics) of land cover with the MWD, $WSA_{>0.25}$, and K factor; (ii) examine the relationships of MWD, $WSA_{>0.25}$ and K factor with terrain, vegetation, temperature and aridity indices; and (iii) develop an optimum combination of the auxiliary variables to optimize the spatial prediction of MWD, $WSA_{>0.25}$, and K factor.

2. Materials and methods

2.1. Study area and soil sampling

The representativeness, geographical location, natural conditions, and soil of Zhifanggou catchment were described in detail in the first paper in this study (Ye et al., 2018). Before the implementation of the "Grain-for-Green" policy, Zhifanggou catchment had been suffering from severe soil erosion. After the implementation of afforestation, the soil erosion obviously decreased. As shown in Figure S1a, the distribution of 5, 100 t/(km^2·a) is the most extensive. The erosion is mainly of moderate grade, followed by strong, very strong and acute erosion (Figure S1b). The data of the erosion modulus and erosion grades were developed by the Loess Plateau Data Center, National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China (http://loess.geodata.cn).

A stratified random sampling irregular grid was designed based on the information of terrain condition, land-use type, and accessibility. Its detail refers to our previous study (Ye et al., 2018). Then, total of 70 sampling sites were selected to represent the major landscape units. The soil samples were collected with aluminum containers and cutting rings (undisturbed soil) and with ziplock bags (disturbed soil) from 0 to 10 cm and 10–20 cm soil layers, respectively. Three disturbed soil samples within a radius of 5 m were collected and mixed. Then, approximately 1 kg of soil was stored in ziplock bags by the quartering method for physical and chemical analyses. In a similar procedure, another set of undisturbed samples were also taken. A total of 140 disturbed soil samples were collected from two soil layers in 70 sampling sites (70 sites × 2 layers = 140 samples). Additionally, 700 undisturbed soil samples were taken (70 sites × 2 layers × 5 samples = 700 samples). The soil samples were subsequently air dried at room temperature under well-ventilated conditions.

2.2. Laboratory analysis and calculations

The undisturbed soil samples in the cutting ring (5 cm in diameter; 100 cm^3 in volume) were used to measure the soil bulk density (BD, g/cm^3) by the oven-drying method (Chen et al., 2014). Total porosity (TP, %) was calculated based on the BD and particle density values (Barik et al., 2014). The undisturbed soil samples in the aluminum containers were used to measure the aggregate stability by the wet sieving method (Kemper and Rosenau, 1986). Then, MWD, $WSA_{>0.25}$, geometric mean diameter (GMD), and K factor were calculated as follows (Li et al., 2016; Ye et al., 2018):

$$MWD = \sum_{i=1}^{5} \bar{x}_i w_i \quad (1)$$

$$WSA = 100 \times \sum_{i=1}^{5} w_i \quad (2)$$

$$GMD = \exp\left(\sum_{i=1}^{5} w_i ln\bar{x}_i / \sum_{i=1}^{5} w_i\right) \quad (3)$$

Table 1
Descriptive statistics of soil properties and auxiliary data.

Parameters	Min		Max		Mean		SD		CV (%)	
	0–10 cm	10–20 cm	0–10 cm	10–20 cm	0–10 cm	10–20 cm	0–10 cm	10–20 cm	0–10 cm	10–20 cm
TP (%)	46.82	46.33	62.59	61.30	56.74	54.32	3.64	3.14	6.42	5.77
pH (H_2O)	8.12	8.02	8.52	8.63	8.36	8.40	0.11	0.10	1.26	1.16
SOC (g/kg)	2.58	1.88	22.32	16.32	5.76	4.22	3.46	2.16	60.12	51.27
$CaCO_3$ (g/kg)	5.33	5.83	13.83	15.60	10.26	11.17	1.48	1.51	14.38	13.49
Fed (g/kg)	6.14	6.53	10.46	10.32	8.61	8.70	0.83	0.78	9.66	8.94
Feo (g/kg)	0.20	0.21	0.98	0.98	0.38	0.37	0.12	0.12	31.67	33.31
Clay (%)	9.64	9.58	24.91	26.52	14.45	16.40	2.79	3.67	19.34	22.36
Silt (%)	55.35	56.98	69.20	72.86	61.97	62.52	3.13	3.30	5.06	5.28
Sand (%)	5.90	6.32	34.37	33.43	23.58	21.08	5.45	5.92	23.13	28.07
MWD (mm)	0.21	0.26	2.82	3.23	1.35	1.38	0.63	0.69	46.72	50.32
$WSA_{>0.25}$ (%)	16.41	14.96	76.61	83.92	44.59	43.15	13.64	14.51	30.59	33.64
K factor	0.0159	0.0156	0.0425	0.0484	0.0223	0.0220	0.0062	0.0072	27.71	32.85
LST (°C)	16.21		25.49		20.32		2.37		11.65	
ΔLST (°C)	27.40		43.40		35.35		3.46		9.79	
NDVI	−0.02		0.17		0.10		0.04		37.42	
TVDI	0.19		0.99		0.54		0.22		40.53	
H (m)	1043.10		1407.50		1235.11		94.18		7.62	
Slope (°)	0		30		14.35		9.20		64.09	
Aspect (°)	9		352		162.13		109.07		67.27	
TWI	1.55		9.00		3.89		1.19		30.74	

Min: minimum; Max: maximum; Std. dev: standard deviation; CV: coefficient of variation.
TP: total porosity; SOC: soil organic carbon; $CaCO_3$: calcium carbonate; Fed: dithionite-extractable iron: poorly ordered and crystalline forms of iron; Feo: oxalate-extractable iron: poorly ordered forms of iron; Clay: clay content; Silt: silt content; Sand: sand content; MWD: mean weight-diameter; $WSA_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor. LST: land surface temperature; ΔLST: the land surface temperature difference between seasons; NDVI: normalized difference vegetation index; TVDI: temperature vegetation dryness index; H: elevation; TWI: topographic wetness index.

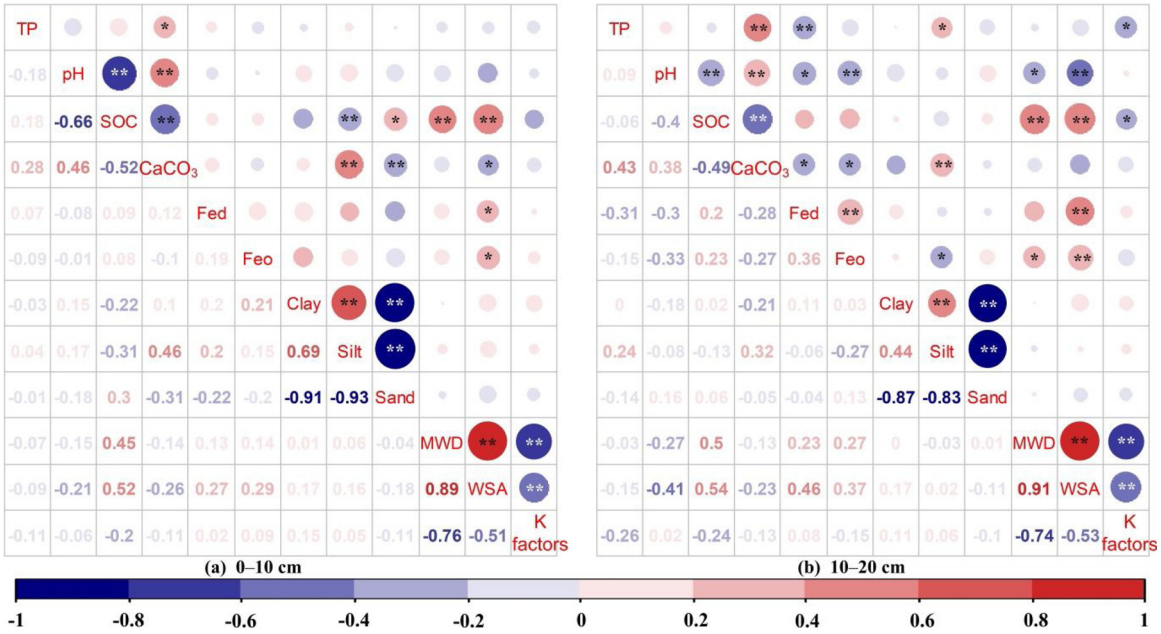


Fig. 1. Heat map of the correlation matrix for soil physicochemical properties in soil samples (n = 70) based on Pearson's correlation coefficients: (a) 0–10 cm; (b) 10–20 cm. *, significant correlations at $P < 0.05$, **, significant correlations at $P < 0.01$. TP: total porosity; SOC: soil organic carbon; $CaCO_3$: calcium carbonate; Fed: dithionite-extractable iron; Feo: oxalate-extractable iron; Clay: clay content; Silt: silt content; Sand: sand content; MWD: mean weight-diameter; $WSA_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor.

$$K = 7.954 \times \left\{0.0017 + 0.0494 \times \exp\left[-0.5 \times \left(\frac{\log GMD + 1.675}{0.6986}\right)^2\right]\right\} \quad (4)$$

where $\bar{x}_i$ is the mean diameter of size class i (mm), and w_i is the percentage of aggregates in size class i.

The disturbed soil samples were collected for physical and chemical analyses. Our primary goal was to study the aggregate stability of surface and subsurface soils. Therefore, the soil properties significantly related to the soil aggregate stability in the Loess Plateau were selected for the study, mainly including soil organic carbon (SOC), calcium carbonate ($CaCO_3$), dithionite-extractable iron (Fed), oxalate-extractable iron (Feo), pH, and soil textural measures (Algayer et al., 2014; An et al., 2013; Regelink et al., 2015; Wei et al., 2014). The disturbed soil samples were air-dried and passed through 2-mm, 0.25-mm, and 0.15-mm plastic sieves. SOC was determined by the Walkley-

Table 2
Pearson's correlation coefficients between MWD, $WSA_{>0.25}$ and K factor and landscape metrics and environmental factor.

Category	Factors	MWD (mm)		$WSA_{>0.25}$ (%)		K factor	
		0–10 cm	10–20 cm	0–10 cm	10–20 cm	0–10 cm	10–20 cm
Class level (Residential)	PLAND	−.027	.111	.115	.168	.157	−.017
	LPI	−.138	.062	−.035	.087	.223	−.030
	ED	.072	.091	.202	.162	.030	−.003
	AI	−.206	−.030	−.104	−.014	.308*	.113
Class level (Farmland)	PLAND	−.384*	−.282	−.319*	−.187	.598**	.533**
	LPI	−.455**	−.355*	−.380*	−.251	.655**	.590**
	ED	−.091	−.089	−.131	−.090	.188	.207
	AI	−.209	−.236	−.203	−.204	.243	.229
Class level (Woodland)	PLAND	.139	−.001	.082	−.099	−.330**	−.252*
	LPI	.045	−.058	.017	−.141	−.195	−.213
	ED	.285*	.120	.280*	.136	−.343**	−.225
	AI	−.074	−.116	−.109	−.187	−.049	−.083
Class level (Grassland)	PLAND	.228	.248*	.161	.232	−.370**	−.315**
	LPI	.244*	.286*	.190	.281*	−.328**	−.287*
	ED	.227	.188	.183	.198	−.398**	−.272*
	AI	.055	.195	.087	.200	−.191	−.245*
Landscape level	LPI	−.133	−.176	−.128	−.214	.162	.126
	ED	.241*	.151	.272*	.198	−.234	−.151
	CONTAG	−.084	−.148	−.126	−.199	.083	.110
	SHDI	−.019	.097	.105	.180	.190	.025
	AI	−.227	−.124	−.251*	−.168	.217	.121
Environmental factors	LST	.008	−.066	−.105	−.144	−.131	−.039
	ΔLST	−.007	.031	.105	.141	.322**	.346**
	NDVI	.311**	.182	.285*	.182	−.319**	−.281*
	TVDI	.037	−.013	−.074	−.094	−.193	−.099
	B_4/B_3	−.315**	−.245*	−.268*	−.207	.285*	.308**
	B_6/B_7	.293*	.215	.279*	.203	−.293*	−.319**
	TWI	−.112	−.136	−.065	−.157	.219	.250*
	Elevation	−.200	−.090	−.363**	−.287*	−.097	−.171
	Slope	.323**	.364**	.249*	.262*	−.482**	−.465**
	Aspect	−.173	−.023	−.155	.134	.054	.117

PLAND: percentage of landscape area; LPI: largest patch index; ED: edge density; AI: aggregation index; CONTAG: contagion; SHDI: shannon's diversity index. LST: land surface temperature; ΔLST: the land surface temperature difference between seasons; NDVI: normalized difference vegetation index; TVDI: temperature vegetation dryness index; B_i is the reflectance of Band i; TWI: topographic wetness index. MWD: mean weight-diameter; $WSA_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor.

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

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

Black method (Nelson and Sommers, 1996). The HCl method was used to determine the calcium carbonate content (Horváth et al., 2005). Fed and Feo were extracted using dithionite-citrate-bicarbonate (DCB) solution (Mehra and Jackson, 1958) and oxalic acid/ammonium oxalate at pH 3.0 (McKeague and Day, 1966), respectively. Soil pH was measured in a soil-distilled water (1:2.5 soil: H_2O) suspension with a pH meter. The soil particle-size distribution was measured using a Mastersizer 2000 (Malvern Instruments, Malvern, England).

2.3. Acquisition of auxiliary data

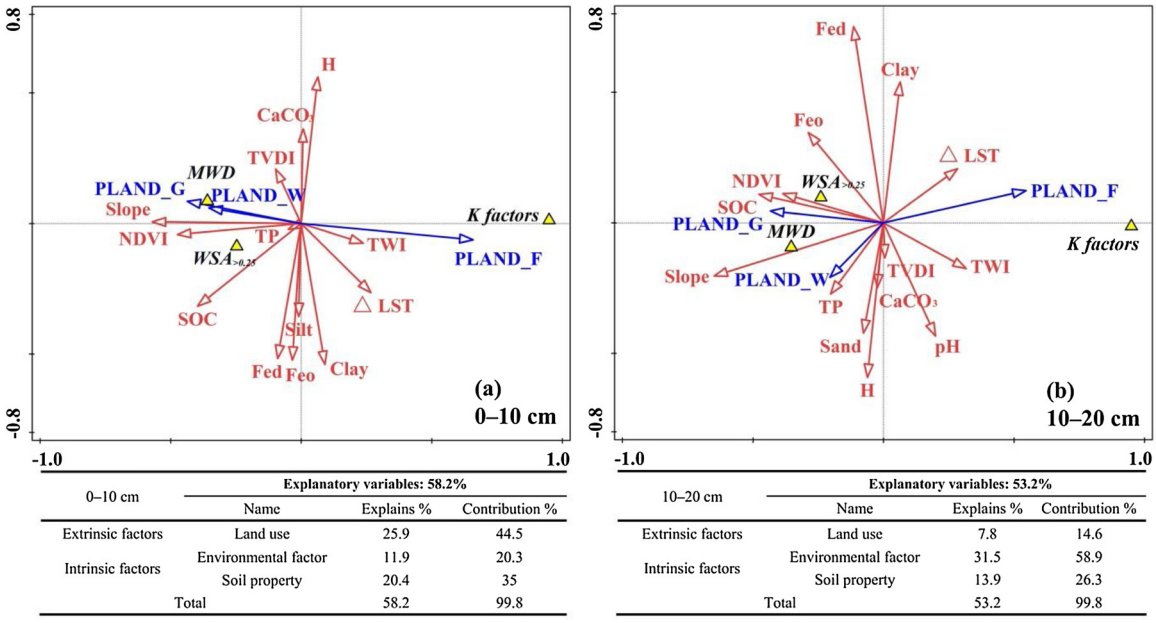
2.3.1. Landscape analysis

It has been demonstrated that soil erosion and soil structure are related to the land use type and landscape structure (Shi et al., 2013). In the past few decades, a large number of landscape metrics have been developed and widely used in characterizing landscape patterns and correlating landscape structures with ecological processes (Liu et al., 2016; McGarigal and Marks, 1995; Turner, 2005). The landscape metrics of cropland in each subplot were applied to quantify the intensity of human activities (Liu et al., 2013). Therefore, we selected thirteen commonly used landscape metrics to associate the soil structure with landscape patterns first (Huang et al., 2013; Ouyang et al., 2010). Then, according to the results of Shi et al. (2013), six landscape metrics closely related to soil erosion were finally selected. These landscape metrics describe the landscape structure and the complexity of this structure, such as size, shape, distance, connectivity, and diversity of land patches (Botequilha Leitão and Ahern, 2002). They can help us to calculate the landscape composition and configuration, which are characteristics of landscape structure. Landscape metrics could also provide important information about the effects of ecological processes and human activities on the landscape (McGarigal et al., 2002). The landscape metrics are presented in Table S1, and the detailed calculation and description were presented by McGarigal et al. (2002).

An aerial photo with 40 cm resolution was used in the landscape analysis (Figure S2). It provided detailed information of the impervious surface area (ISA, including roads and residential areas), grassland, woodland, and farmland. The data were provided by the Loess Plateau Data Center, National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China (http://loess.geodata.cn). The landscape metrics were computed at both the class-level and landscape-level to represent the landscape composition and spatial configuration (McGarigal and Marks, 1995). By combining the semivariogram results of Ye et al. (2018), we additionally generated a 250 × 250 m sample plot around the sampling sites (Hou et al., 2015). The landscape metrics of each plot were computed by the software FRAGSTATS 4.1 (McGarigal et al., 2012). Pearson's correlation analysis was used to explore the relation between landscape metrics and the aggregate stability. The significance of each correlation coefficient was confirmed using a two-tailed Student's t-test.

2.3.2. Remote sensing data and applications

Bands and indices derived from remote sensing data have been widely used in the prediction of soil properties (Boettinger et al., 2008; Mirzaee et al., 2016). Given that the soil properties are relatively



0–10 cm	Explanatory variables: 58.2%		
	Name	Explains %	Contribution %
Extrinsic factors	Land use	25.9	44.5
Intrinsic factors	Environmental factor	11.9	20.3
	Soil property	20.4	35
Total		58.2	99.8

10–20 cm	Explanatory variables: 53.2%		
	Name	Explains %	Contribution %
Extrinsic factors	Land use	7.8	14.6
Intrinsic factors	Environmental factor	31.5	58.9
	Soil property	13.9	26.3
Total		53.2	99.8

Fig. 2. Canonical correspondence analysis ordination diagram illustrating the relationships of auxiliary data with MWD, $WSA_{>0.25}$, and K factor: (a) 0–10 cm; (b) 10–20 cm. Arrows indicate auxiliary data; triangles represent MWD, $WSA_{>0.25}$ and K factor. MWD: mean weight-diameter; $WSA_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor; TP: total porosity; SOC: soil organic carbon; Fed: dithionite-extractable iron: poorly ordered and crystalline forms of iron; Feo: oxalate-extractable iron: poorly ordered forms of iron; Clay: clay content; Silt: silt content; Sand: sand content; $CaCO_3$: calcium carbonate; PLAND_F, PLAND_G, and PLAND_W: the proportion of the landscape occupied by farmland, grassland and woodland, respectively. ΔLST: the land surface temperature difference between seasons; NDVI: normalized difference vegetation index; TVDI: temperature vegetation dryness index; H: elevation; TWI: topographic wetness index.

Table 3
Multiple linear regression models for predicting MWD, $WSA_{>0.25}$ and K factor using soil data, environmental data, landscape metrics, and the coefficient of determination (R^2).

Parameters	Layer (cm)	Model	R^2	P-value
MWD (mm)	0–10	$MWD = 0.347^{***} + 0.81\ SOC^{***} - 0.358\ PLAND\_F^{**} - 0.343\ TP^{**} + 0.264\ Silt^{**} + 0.166\ Slope^{*}$ (SSPFs)	0.471	0.000
		$MWD = 0.538^{***} - 0.215\ B_4/B_3^{*} - 0.22\ H^{**} + 0.255\ Slope^{**}$ (ESPFs)	0.312	0.001
	10–20	$MWD = 0.01^{*} + 0.819\ SOC^{***} + 0.206\ Slope^{**} + 0.239\ PLAND\_G^{**}$ (SSPFs)	0.436	0.000
		$MWD = 0.21^{***} + 0.279\ Slope^{**}$ (ESPFs)	0.183	0.001
$WSA_{>0.25}$ (%)	0–10	$WSA_{>0.25} = 0.73^{***} + 0.898\ SOC^{***} - 0.465\ Sand^{***} - 0.254\ PLAND\_F^{**} - 0.325\ TP^{***} + 0.151\ Slope^{**}$ (SSPFs)	0.571	0.000
		$WSA_{>0.25} = 0.56^{***} - 0.396\ H^{***} + 0.26\ Slope^{**}$ (ESPFs)	0.288	0.000
	10–20	$WSA_{>0.25} = 0.109^{*} + 0.587\ SOC^{***} + 0.326\ Fed^{***} + 0.21\ PLAND\_G^{**} + 0.148\ Slope^{*} - 0.272\ pH^{*}$ (SSPFs)	0.578	0.000
		$WSA_{>0.25} = 0.472^{***} - 0.307\ H^{**} + 0.247\ Slope^{**}$ (ESPFs)	0.258	0.000
K factor	0–10	$K = 0.337^{***} + 0.398\ PLAND\_F^{***} - 0.229\ Slope^{**} - 0.261\ NDVI^{*} + 0.276\ TWI^{*}$ (SSPFs & ESPFs)	0.536	0.000
	10–20	$K = 0.295^{***} - 0.259\ Slope^{***} + 0.264\ PLAND\_F^{**} + 0.326\ TWI^{*} - 0.189\ B_6/B_7^{*}$ (SSPFs & ESPFs)	0.506	0.000

SOC: soil organic carbon; PLAND_F and PLAND_G: the proportion of the landscape occupied by farmland and grassland, respectively; TP: total porosity; B_i is the reflectance of Band i; H: elevation; Fed: dithionite-extractable iron; NDVI: Normalized difference vegetation index; TWI: topographic wetness index; MWD: mean weight-diameter; $WSA_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor. SSPFs: soil spatial prediction functions; ESPFs: environmental spatial prediction functions. *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$.

constant in a short period of time (Shi et al., 2013) and the effects of the environmental factors on soil cannot be represented by one image (Yan, 2015), nearly one year of remote sensing data were selected to better explore the effects of environmental variables on soil. They were extracted from thirteen cloud-free Landsat 8 images (Figure S3). The normalized difference vegetation index (NDVI) was used to represent the vegetation cover (Wang et al., 2017). NDVI was derived by the reflectance ratio from the near-infrared band (NIR) and red band (R) as:

$$NDVI = \frac{(NIR - R)}{(NIR + R)} \quad (5)$$

The areas with NDVI larger than 0.1 were defined as vegetated areas (Wang et al., 2017) (Figure S4).

The land surface temperature reflects the temperature conditions of the interface between the atmosphere and soil surface and plays an important role in soil processes (Dimoyiannis, 2009; Oztas and Fayetorbay, 2003). Hence, the LST of Zhifanggou (Figures S5 and S6) was derived from one-year Landsat 8 data by a practical split-window algorithm (Du et al., 2015). Then, TVDI, which is related to the soil moisture, was calculated based on an empirical parameterization of the correlation between NDVI and LST by Sandholt et al. (2002) (Figure S7). The carbonate index (the ratio of bands 3 and 4) and clay index (the ratio of bands 6 and 7) are introduced into the analysis (Boettinger et al., 2008; Mirzaee et al., 2016; Taghizadeh-Mehrjardi et al., 2016). Their abbreviations and descriptions are listed in Table S1. Image data processing was performed in ENVI 5.1 software.

2.3.3. Terrain analysis

The terrain analysis was performed with a 5 m × 5 m digital elevation model (DEM). The topographic derivatives were calculated in ArcGIS 10.2, including elevation (H, m, above mean sea level), slope (S, °), aspect (A, °; 0° and 360° represent the East, 90° the North, 180° the West, and 270° the South), and topographic wetness index (TWI, a spatial estimation of soil hydrological and physical properties) (Figure

Table 4
Cross-validation indices of multiple stepwise regression (MSR) and regression kriging (RK) for MWD (mm), $WSA_{>0.25}$ (%) and K factor in different soil layers.

Parameters	Layer (cm)	Methods	R^2	NRMSE	AIC
MWD (mm)	0–10	MSR (SSPFs)	0.440	0.325	−24.326
		MSR (ESPFs)	0.365	0.360	−25.716
		RK (ESPFs)	0.584	0.266	−38.457
		CK (ESPFs)	0.618	0.273	−37.346
	10–20	MSR (SSPFs)	0.514	0.322	−28.163
		MSR (ESPFs)	0.201	0.471	−23.356
		RK (ESPFs)	0.379	0.448	−20.871
		CK (ESPFs)	0.389	0.403	−19.701
$WSA_{>0.25}$ (%)	0–10	MSR (SSPFs)	0.459	0.191	−92.374
		MSR (ESPFs)	0.265	0.119	−93.698
		RK (ESPFs)	0.616	0.155	−106.626
		CK (ESPFs)	0.547	0.194	−94.959
	10–20	MSR (SSPFs)	0.714	0.142	−105.869
		MSR (ESPFs)	0.285	0.226	−92.344
		RK (ESPFs)	0.782	0.162	−106.380
		CK (ESPFs)	0.532	0.204	−94.508
K factor	0–10	MSR (SSPFs & ESPFs)	0.456	0.193	−221.522
		RK (ESPFs)	0.492	0.184	−223.434
		CK (ESPFs)	0.747	0.165	−230.179
	10–20	MSR (SSPFs & ESPFs)	0.522	0.229	−213.967
		RK (ESPFs)	0.568	0.210	−217.637
		CK (ESPFs)	0.593	0.261	−210.587

MWD: mean weight-diameter; $WSA_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor. R^2: the coefficient of determination; NRMSE: the normalized root mean squared error; AIC: the Akaike information criterion. MSR: multiple stepwise regression; RK: regression kriging. SSPFs: soil spatial prediction functions; ESPFs: environmental spatial prediction functions.

S8) (Moore et al., 1991). The abbreviations and descriptions of these topographic attributes are listed in Table S1.

2.4. Statistical analysis

The descriptive statistics of the data including minimum (Min), maximum (Max), mean, standard deviation (SD), and coefficient of variation (CV) were obtained by using the statistical software SPSS 18.0. The CV is considered as a criterion to represent the variability in the factors: least (< 15%), moderate (15–35%), and most variable (> 35%) (Wilding, 1985). Initially, the Kolmogorov-Smirnov test was used to examine the normality of all data. Pearson's correlation analysis was performed using RStudio3.2.2 (http://www.r-project.org/) to explore the relationships of MWD, $WSA_{>0.25}$ and K factor with the auxiliary data. The cross-variograms were further calculated to analyze their spatial correlations by using GS + 9.0 (Gamma Design Software) (Rosemary et al., 2017).

The canonical correspondence analysis method, which combines correspondence analysis with multiple regression analysis, can be used to extract the best synthetic gradients from the field data of species and environmental features. Then, a linear combination of environmental variables is formed to isolate the niches of the species to the greatest extent (Klami et al., 2013). The correlations of soil aggregate stability indices (MWD, $WSA_{>0.25}$ and K factor) with soil properties and natural and human factors were determined via CCA by CANOCO 5 (Microcomputer Power, Ithaca, USA). CCA was also used to explore the relationships between soil structure indices (MWD, $WSA_{>0.25}$, K factor, and aggregate size distribution) and binding agents (SOC, Fed, Feo, Clay, and $CaCO_3$).

2.5. Prediction methods

Multiple stepwise regression (MSR) in SPSS 18.0 was performed to predict MWD, $WSA_{>0.25}$, and K factor. Environmental spatial prediction functions (ESPFs) and soil spatial prediction functions (SSPFs) are two prediction methods based on the data sets of input variables. They were used to predict aggregate stability indices from readily available data by including environmental data, such as topographic attributes, NDVI, LST, TVDI, and landscape metrics (in the case of ESPFs), and by combination of soil properties, environmental variables and landscape metrics (in the case of SSPFs) (Zolfaghari et al., 2016). The input variables (x) were used as independent variables in the MSR analyses. All input data were rescaled using the following equation (Besalatpour et al., 2013):

$$x_i = 0.1 + 0.8\left[\frac{x - x_{min}}{x_{max} - x_{min}}\right] \quad x_{min} < x < x_{max} \tag{6}$$

where x_i is the rescaled data, and x_{min} and x_{max} are the minimum and maximum observed values, respectively. The training data set was randomly chosen from 70% of the total data set and the remaining 30% of the samples were used as the validation set.

Regression kriging combines the MSR and residual kriging (Hengl et al., 2007):

$$Z = Z^* + \varepsilon \tag{7}$$

where Z is the predicted value of RK, Z^* is the regression prediction value, and ε is the residual from the regression model at each site.

First, the regression model is built to fit the available data. Then, ε is computed. The residual kriging is used to fit the residual of the MSR. The influence factors with significant correlation coefficients were chosen as co-variables for co-kriging interpolation. All auxiliary data were resampled to 5 m resolution for mapping the spatial distributions of MWD, $WSA_{>0.25}$, and K factor in ArcGIS 10.2 based on the best-fit model from MSR by RK.

2.6. Evaluation of prediction accuracy

The accuracy of the derived SSPFs and ESPFs was evaluated by the coefficient of determination (R^2), the normalized root mean squared error (NRMSE) and the Akaike information criterion (AIC). The NRMSE is a measure of the scatter between the observed and predicted values (Gérard et al., 2008). The AIC, a measure of the relative goodness of fit for statistical models (Arnold, 2010), was also used to evaluate the efficacy of SSPFs and ESPFs. Lower NRMSE and AIC values and greater R^2 values indicate a better performance of the model.

$$\text{NRMSE} = \left[\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}\right]^{0.5} \times \frac{100}{\bar{y}} \tag{8}$$

$$\text{AIC} = 2\,\text{k} + \text{nln}\left[\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}\right] \tag{9}$$

where $\hat{y}_i$ is the predicted value of observation i, y_i is the measured value, $\bar{y}$ is the mean of y_i, k is the number of regression coefficients, and n is the number of observations.

3. Results and discussion

3.1. Descriptive statistics

Table 1 shows the least (CV < 15%) or moderate variabilities (15% < CV < 35%) of the soil physical and chemical properties that can be used to develop the prediction models. SOC shows a high variability over the study area (CV > 35%). Most of the soils are a silt loam texture with an alkaline pH and a relatively low SOC content (Ye et al., 2018). The soil particle size distribution exhibits a relatively low variability, especially for the silt fraction (CV = 5.06% in 0–10 cm and 5.28% in 10–20 cm soil layers). The different forms of iron show significant differences in variability. Moreover, relatively higher variabilities were observed for SOC (CV = 60.12% in 0–10 cm and 51.27%

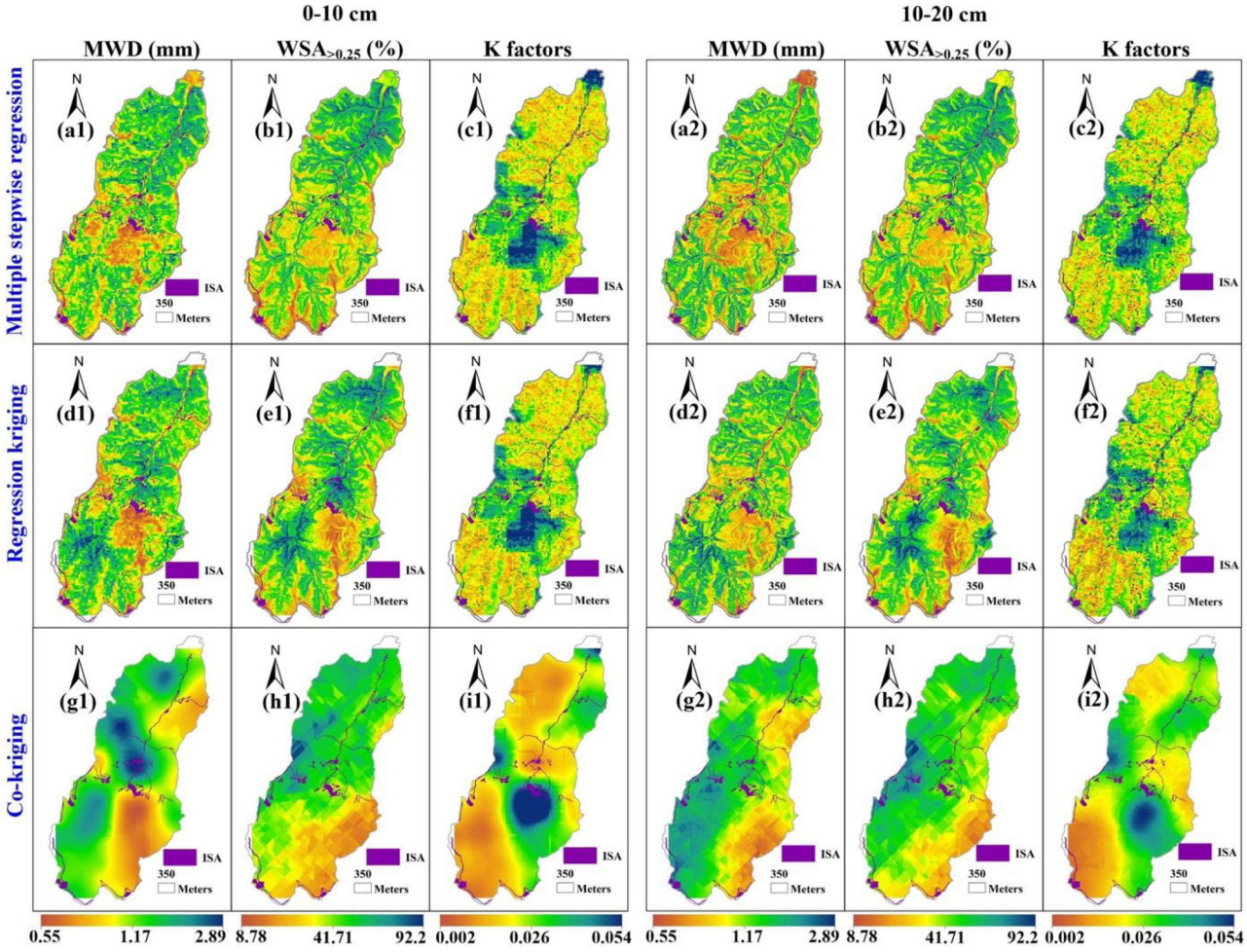


Fig. 3. Spatial distributions of MWD, WSA $_{>0.25}$, and K factor in both 0–10 cm (a1-i1) and 10–20 cm (a2-i2) soil layers. They are predicted by multiple stepwise regression (a1-c1, a2-c2), regression kriging (d1-f1, d2-f2), and co-kriging (g1-i1, g2-i2). MWD: mean weight-diameter; WSA $_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor.

in 10–20 cm soil layers), presumably due to the variability in pedogenic processes, land use types and management practices. In contrast, $CaCO_3$ shows a relatively stable variation. The three aggregate stability indices (MWD, WSA $_{>0.25}$ and K factor) also show moderate to high variabilities over Zhifanggou catchment, which can be attributed to the combined effects of soil properties, natural factors and land use type and landscape structure. The descriptive statistics for environmental factors and land use type and landscape structure as the auxiliary variables are also presented in Table 1. Overall, there are high variabilities in topographic conditions, annual vegetation cover and soil aridity.

3.2. Factors influencing soil aggregate stability

3.2.1. Influence of soil intrinsic properties on soil aggregate stability

The correlation matrix of soil physicochemical properties in soil samples based on Pearson's correlation coefficients is shown in Fig. 1. Under all land use types, there are significant correlations among MWD, WSA $_{>0.25}$, and K factor. For both MWD and WSA $_{>0.25}$, significantly positive correlations are found with SOC in the 0–10 cm soil layer, while they show significantly positive correlations with SOC and Feo and negative correlations with pH in the 10–20 cm soil layer. Moreover, WSA $_{>0.25}$ has significantly negative correlations with $CaCO_3$ and positive correlations with Fed and Feo in the 0–10 cm soil layer, while it only has a positive correlation with Fed in the 10–20 cm soil layer. For K factor, significantly negative correlations are found with TP and SOC in the 10–20 cm soil layer. Spatial cross-correlation indicates that aggregate stability indices vary closely with the influence factors depending on the distances between samples (Figure S9). Their signs of spatial correlation coefficients can shift from positive to negative as the separation distance increases. There are no or weak spatial correlations between aggregate stability indices and almost all influence factors for distances more than 465 m, which means that the calculated correlation coefficients from regression analysis can be true within the distance less than 465 m. At zero distance, their correlations are equal to their Pearson's correlation coefficients. Soil organic carbon is the major cement for soil aggregation, and its content is closely related to soil aggregation (Wei et al., 2014). In addition, Ca^{2+}, which is mainly from the dissolution of $CaCO_3$, acts as an important aggregate-binding agent. When microaggregates are formed by the binding of Ca^{2+}, the soil $CaCO_3$ content exhibits a corresponding decrease. Apart from $CaCO_3$, soil pH is also strongly and negatively correlated with soil aggregate stability because increasing pH leads to clay dispersion resulting from the increased repulsion of negatively charged clay particles.

The canonical correspondence analysis ordination diagram illustrates the relationships of binding agents (SOC, Fed, Feo, Clay, and $CaCO_3$) with MWD, WSA $_{>0.25}$, K factor and aggregate size distribution (Figure S10). In general, SOC and $CaCO_3$ mainly affect MWD, WSA $_{>0.25}$, K factor, and macro aggregates (> 5 mm and 5–2 mm), while Fed, Feo and Clay have significant effects on aggregates with diameters smaller than 2 mm. Clay content is not significantly correlated with MWD, WSA $_{>0.25}$, and K factor, but significantly and

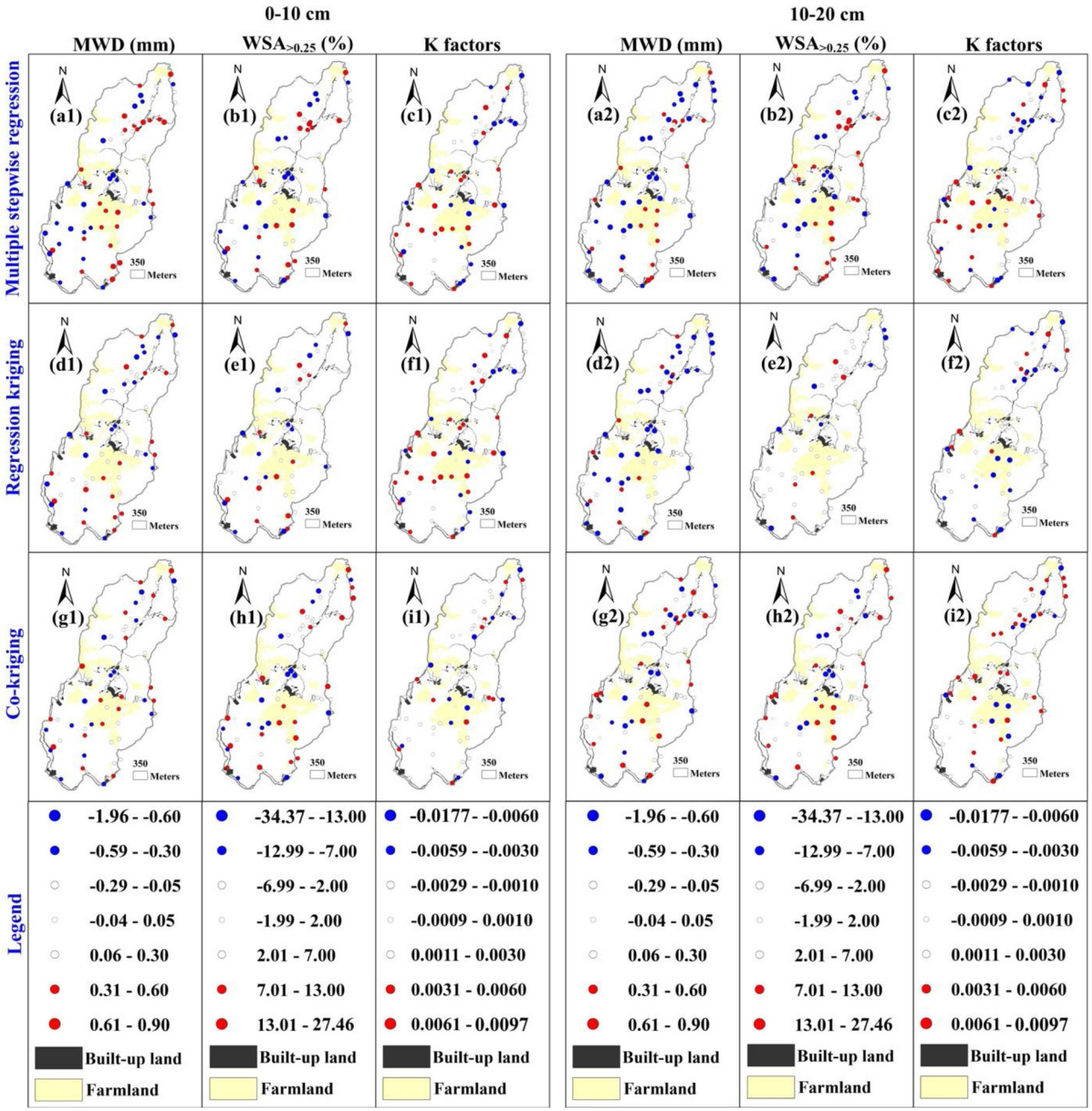


Fig. 4. Prediction residuals of MWD, WSA $_{>0.25}$ and K factor estimated by multiple stepwise regression (a1-c1, a2-c2), regression kriging (d1-f1, d2-f2), and co-kriging (g1-i1, g2-i2) in 0–10 cm (a1-i1) and 10–20 cm (a2-i2) soil layers. MWD: mean weight-diameter; WSA $_{>0.25}$: percentage of water-stable aggregates greater than 0.25 mm; K factor: soil erodibility factor.

positively related to aggregates with diameters smaller than 2 mm (Fig. 1 and Figure S10). Hence, these five binding agents were taken into account when performing MSR. The results show that Clay and $CaCO_3$ are the dominant binding agents in farmland, which is characterized by aggregates with diameters smaller than 2 mm, while for other land use types, all the binding agents contribute to soil aggregation (Figure S10).

3.2.2. Influence of land use type and landscape structure on soil aggregate stability

The correlation coefficients between aggregate stability indices and landscape metrics at the class and landscape levels are presented in Table 2. The mean weight-diameter, WSA $_{>0.25}$, and K factor are highly correlated with the landscape metrics for farmland, woodland and grassland. For farmland, MWD and WSA $_{>0.25}$ are negatively correlated to the percent of land use (PLAND) and largest patch index (LPI) in the 0–10 cm soil layer and are only negatively correlated with LPI in the 10–20 cm soil layer. However, K factor is positively correlated with PLAND and LPI in the different soil layers. For woodland, MWD and WSA $_{>0.25}$ are only positively correlated with edge density (ED) in the 0–10 cm soil layer. The K factor of woodland is negatively correlated with PLAND and ED. For grassland, MWD and WSA $_{>0.25}$ are significantly and negatively correlated with LPI. The K factor is significantly and negatively correlated with PLAND, LPI, ED, and the aggregation index (AI). Their correlations are much more significant in farmland and grassland than in woodland and residential land (the land where the villagers live). Moreover, the correlations between MWD, WSA $_{>0.25}$ and K factor and landscape metrics are stronger in the

0–10 cm soil layer than in the 10–20 cm soil layer.

At the landscape level, ED and AI are significantly correlated with aggregate stability in the 0–10 cm soil layer, but not in the 10–20 cm soil layer. The K factor is insignificantly correlated with landscape metrics in both soil layers. In most cases, the correlation coefficient signs of MWD and $WSA_{>0.25}$ with landscape metrics are opposite to those of K factor, though these correlation coefficients do not reach significant levels.

The relationships between MWD, $WSA_{>0.25}$ and K factor and landscape metrics can provide important implications for revegetation and agricultural land use management to mitigate the negative effects of human activity on soil structure. Although previous studies (Ozgoz et al., 2013; Zhao et al., 2017) have suggested that farming negatively affects soil structure, our research reveals that this negative effect can be mitigated by programming the land use composition and spatial configuration at the class and landscape levels, respectively. At the class level, the landscape metrics for residential land have no significant correlation with MWD, $WSA_{>0.25}$ and K factor, possibly due to the long distance between the residential land and the sampling sites. However, the signs of the correlation coefficients reveal a decrease in soil stability and increase in soil erodibility when residential land is dominant or well-connected. Farmland has similar but stronger effects on the soil structure than residential land, with stronger correlations in the 0–10 cm soil layer than in the 10–20 cm soil layer. Therefore, special attention should be given to these two land use types. Even though the areas of farmland and residential land cannot be further reduced considering the need to ensure local and regional food supply, soil erosion still can be mitigated by increasing the patch number and limiting the patch area and connectivity. Our results indicate that woodland and grassland, particularly the latter, are always beneficial to increasing soil aggregate stability. The size and connectivity of vegetation are important factors to influence the positive effects of these two types of land use. At the landscape level, MWD and $WSA_{>0.25}$ are positively correlated with ED, while negatively correlated with AI, indicating that a blending of different land use types helps to mitigate the negative effects of human activity on the soil structure. These results indicate that interspersing revegetation into farmland or residential land patches can effectively alleviate the negative effects of human activity.

3.2.3. Influence of environmental factors on soil aggregate stability

The Pearson's correlation coefficients between topography, hydrothermal conditions, and vegetation cover and soil aggregate stability are presented in Table 2. In general, MWD, $WSA_{>0.25}$ and K factor are significantly related to ΔLST, NDVI, B_4/B_3, B_6/B_7, TWI, Elevation, and Slope. The mean weight-diameter and $WSA_{>0.25}$ are positively correlated with NDVI, B_6/B_7, and Slope but negatively correlated with B_4/B_3 and weakly correlated with the other covariates (not significant). These results are not in agreement with the findings presented by Ayoubi et al. (2012), who reported a good soil structure in the lower slope. Our results indicate the importance of revegetation in areas with a steep slope for soil protection in natural environments. The topographic factors affect soil properties by influencing the vegetation activity, soil microbes, microclimate, hydrothermal conditions and land use (Teng et al., 2017). Moreover, significant temperature variations among the four seasons promote soil erosion (Table 2). The band reflectance ratios have significant correlations with MWD, $WSA_{>0.25}$ and K factor. The signs of the correlation coefficients between K factor and environmental factors are opposite to those between MWD or $WSA_{>0.25}$ and environmental factors. The results of the correlation analysis for MWD, $WSA_{>0.25}$, K factor and environmental factors are also supported by the results of CCA (Fig. 2). The three parameters are arranged close to the first axis on the CCA ordination biplot. The NDVI, Slope and landscape structure (PLAND_farmland, PLAND_grassland and PLAND_woodland) are the dominant factors associated with this axis. The first paper in this series speculated that the spatial variability of soil aggregate stability was mainly controlled by intrinsic factors (such as parent materials, terrain attributes and soil types) and the effects of extrinsic factors (land use and farming practice) could not be ignored, especially for K factor (Ye et al., 2018). This work further indicates that soil intrinsic properties play important roles in affecting MWD and $WSA_{>0.25}$ in both 0–10 and 10–20 cm soil layers, especially for $WSA_{>0.25}$. The topography, hydrothermal conditions, vegetation cover, and land use pattern play more important roles than soil properties in affecting K factor, especially in the 0–10 cm soil layer (Tables 2 and 3).

3.3. The comparison of SSPFs and ESPFs

The prediction functions for MWD, $WSA_{>0.25}$ and K factor are presented in Table 3. The MSR analysis shows moderate to high correlations between MWD, $WSA_{>0.25}$ and K factor and the covariates ($0.183 \leq R^2 \leq 0.578$); the NRMSE and AIC results suggest a better performance of the model for $WSA_{>0.25}$ and K factor (Table 4). Among the independent variables, the most relevant variables for predicting MWD and $WSA_{>0.25}$ by the stepwise models are SOC, TP, Slope, PLAND, Fed, Silt, Sand and H. However, NDVI is excluded from the stepwise models despite having significant correlations with MWD and $WSA_{>0.25}$ (Tables 2 and 3), which may be attributed to the interrelation between these covariates. For K factor, the most relevant variables are PLAND of farmland, TWI and Slope, and NDVI. The MSR models including soil data as predictors (SSPFs) explain 47.1% and 57.1% of the variability in MWD and $WSA_{>0.25}$ in the 0–10 cm soil layer, respectively. Exclusion of auxiliary soil variables reduces the performance of the SSPFs for the prediction of MWD and $WSA_{>0.25}$ but does not affect the prediction for K factor in different soil layers. Therefore, the prediction models based on the combination of soil properties and natural and human factors would improve their performance in the prediction for MWD and $WSA_{>0.25}$ but not for K factor. The application of easily available auxiliary variables such as derivatives from DEM (H, Slope, Aspect, and TWI) and remote sensing (LST, NDVI, TVDI, and landscape pattern) can improve the performance of the models in predicting the MWD, $WSA_{>0.25}$ and K factor at the landscape level (Table 3). The topography, hydrothermal conditions, and vegetation cover have a direct or indirect influence on the soil properties (Ayoubi et al., 2012; Oztas and Fayetorbay, 2003; Soinne et al., 2016). Overall, the models based on the data of soil properties and environmental variables can produce satisfactory results in predicting MWD, $WSA_{>0.25}$ and K factor, particularly for the 0–10 cm soil layer.

3.4. Spatial predictions of MWD, $WSA_{>0.25}$ and K factor based on ESPFs

The spatial distribution maps of MWD, $WSA_{>0.25}$ and K factor by using RK, CK and MSR approaches are shown in Fig. 3. The maps produced by MSR, RK, and CK reveal the details of the spatial variations of these three parameters. Fig. 3 displays the detailed trends of MWD, $WSA_{>0.25}$ and K factor in the study area. The maps show lower values of MWD and $WSA_{>0.25}$ and higher values of K factor in farmland, which is consistent with the findings of the first paper in this series (Ye et al., 2018). The maximum NRMSE and AIC values generated by these two approaches are both for the predictions of MWD. However, RK and CK generated lower values of NRMSE and AIC and higher values of R^2 (Table 4) than MSR, indicating that they can provide more satisfactory predictions. The comparison of RK and CK demonstrates that CK did not improve predictions for MWD and $WSA_{>0.25}$ compared to RK. Hence, RK and CK are considered to be an accurate and adequate for spatial interpolation of soil properties when multiple covariables are taken into account and for different indices, it is necessary to choose appropriate methods to accurately interpolate their spatial distribution. In the residual distribution maps (Fig. 4), the sites with high residual values have similar conditions, mostly near the built-up land (transportation land and residential land), farmland and the edge of this catchment. In the built-up land and farmland, there are frequent and intensive human

activities, which will cause soil erosion. The estimation uncertainty of MWD, $WSA_{>0.25}$ and K factor is high in these sites, which is a prevalent situation in the prediction of soil properties (Wang et al., 2013). This may explain the fact that it is difficult to accurately estimate the values of soil properties in these places.

Dematte et al. (2007) reported that the quantitative evaluation of soil properties from remote sensing data is a difficult task due to the complexity of soils. The prediction models in this study based on the combination of soil properties and natural and human factors for MWD, $WSA_{>0.25}$ and K factor using MSR, RK and CK can be considered as satisfactory. Our results show advantages over those of Besalatpour et al. (2013), because MSR was used to extract the most important factors and the land use types were taken into account. The performance of the models for aggregate stability predictions is equivalent to that of Jaksik et al. (2015) and Annabi et al. (2017), who used multiple linear regressions to predict soil aggregate stability based on the data of soil properties and terrain attributes. Jaksik et al. (2015) used topographical and soil data and Annabi et al. (2017) only used soil properties for the prediction. Although their results provided satisfactory predictions, they ignored the impact of many important factors, such as landscape structure and vegetation information. The high dependence on soil data means that the modeling must involve soil sampling. However, we found that K factor do not depend on soil properties, making it possible to predict the soil structure based on more easily available data. In addition, the inclusion of landscape structure, temperature difference between seasons, aridity, and vegetation activity improves the model performance.

4. Conclusions

Our study aims to select appropriate and accessible environmental factors for predicting the spatial distributions of soil aggregate stability indices. These factors can serve as auxiliary variables to improve the accuracy of the interpolation of MWD, $WSA_{>0.25}$ and K factor. The satisfactory predictions of soil aggregate stability based on the data of soil properties and environmental factors confirm the superiority of incorporating landscape metrics, topography, temperature difference between seasons, aridity, and vegetation cover along with the soil properties as predictors. The basic statistical results indicate that these three aggregate stability indices are synergetically affected by the soil properties, natural factors and land use type and landscape structure. The relationships between MWD, $WSA_{>0.25}$ and K factor and various readily available environmental factors reveal that variations in the spatial distribution of aggregate stability indices are affected by the joint effects of soil properties, land use type and landscape structure, terrain conditions, vegetation activity and hydrothermal conditions. Although the exclusion of soil variables (ESPFs) reduces the performance of the SSPFs for MWD and $WSA_{>0.25}$, the predictions are still satisfactory. The predictions for K factor in different soil layers do not depend on soil variables. Residual analysis shows that the sites with high residual values are located close to built-up land or agricultural land, indicating that the high uncertainty of the models based on environmental factors is usually due to the impact of anthropogenic factors. These results reveal that the choice of auxiliary variables plays a vital role in the model building, especially the quantification of land use type and landscape structure that significantly influence the spatial variation of soil properties. The spatial distribution maps of MWD, $WSA_{>0.25}$ and K factor adequately reveal the heterogeneity of local environmental variables and landscape structure. Our findings are of practical value to land managers and policy makers for landscape planning, management and decision making.

Acknowledgments

This investigation was supported by National Natural Science Foundation of China (No. 41330852) and the National Key Basic Research Program of China (No. 2015CB150504). We also like to thank "Loess Plateau Data Center, National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China (http://loess.geodata.cn)" for basic data. Mr. Zuoxiong Liu and American Journal Experts were acknowledged to polish the language.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.still.2019.03.009.

References

Algayer, B., Wang, B., Bourennane, H., Zheng, F., Duval, O., Li, G., Le Bissonnais, Y., Darboux, F., 2014. Aggregate stability of a crusted soil: differences between crust and sub-crust material, and consequences for interrill erodibility assessment. An example from the Loess Plateau of China. Eur. J. Soil Sci. 65 (3), 325–335.
Amezketa, E., 1999. Soil aggregate stability: a review. J. Sustain. Agric. 14 (2-3), 83–151.
An, S.S., Darboux, F., Cheng, M., 2013. Revegetation as an efficient means of increasing soil aggregate stability on the Loess Plateau (China). Geoderma 209, 75–85.
Annabi, M., Raclot, D., Bahri, H., Bailly, J.S., Gomez, C., Le Bissonnais, Y., 2017. Spatial variability of soil aggregate stability at the scale of an agricultural region in Tunisia. Catena 153, 157–167.
Arnold, T.W., 2010. Uninformative parameters and model selection using Akaike's information criterion. J. Wildl. Manage. 74 (6), 1175–1178.
Ayoubi, S., Mokhtari Karchegani, P., Mosaddeghi, M.R., Honarjoo, N., 2012. Soil aggregation and organic carbon as affected by topography and land use change in western Iran. Soil Tillage Res. 121, 18–26.
Barik, K., Aksakal, E.L., Islam, K.R., Sari, S., Angin, I., 2014. Spatial variability in soil compaction properties associated with field traffic operations. Catena 120, 122–133.
Besalatpour, A.A., Ayoubi, S., Hajabbasi, M.A., Mosaddeghi, M.R., Schulin, R., 2013. Estimating wet soil aggregate stability from easily available properties in a highly mountainous watershed. Catena 111, 72–79.
Boettinger, J., Ramsey, R., Bodily, J., Cole, N., Kienast-Brown, S., Nield, S., Saunders, A., Stum, A., 2008. Landsat Spectral Data for Digital Soil Mapping, Digital Soil Mapping With Limited Data. Springer, pp. 193–202.
Botequilha Leitão, A., Ahern, J., 2002. Applying landscape ecological concepts and metrics in sustainable landscape planning. Landscape Urban Plann. 59, 65–93.
Chen, Y.J., Day, S.D., Wick, A.F., McGuire, K.J., 2014. Influence of urban land development and subsequent soil rehabilitation on soil aggregates, carbon, and hydraulic conductivity. Sci. Total Environ. 494, 329–336.
Dematte, J.A.M., Galdos, M.V., Guimaraes, R.V., Genu, A.M., Nanni, M.R., Zullo, J., 2007. Quantification of tropical soil attributes from ETM+/LANDSAT-7 data. Int. J. Remote Sens. 28 (17), 3813–3829.
Dimoyiannis, D., 2009. Seasonal soil aggregate stability variation in relation to rainfall and temperature under Mediterranean conditions. Earth Surf. Process. Landf. 34 (6), 860–866.
Du, C., Ren, H.Z., Qin, Q.M., Meng, J.J., Zhao, S.H., 2015. A practical split-window algorithm for estimating land surface temperature from Landsat 8 data. Remote Sens. 7 (1), 647–665.
Gérard, F., Mayer, K.U., Hodson, M.J., Ranger, J., 2008. Modelling the biogeochemical cycle of silicon in soils: application to a temperate forest ecosystem. Geochim. Cosmochim. Acta 72 (3), 741–758.
Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma 152, 195–207.
Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120 (1), 75–93.
Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-kriging: from equations to case studies. Comput. Geosci. 33, 1301–1315.
Horváth, B., Opara-Nadi, O., Beese, F., 2005. A simple method for measuring the carbonate content of soils. Soil Sci. Soc. Am. J. 69 (4), 1066–1068.
Hou, X.Y., Song, B., Zhao, S., Ding, S.Y., Liang, G.F., Dong, C.F., 2015. Effects of agricultural landscape dynamics on the species diversity of ground arthropods in woodlands. Acta Ecol. Sin. 35 (23), 7659–7668 (in Chinese with English abstract).
Huang, J.L., Li, Q.S., Pontius, R.G., Klemas, V., Hong, H.S., 2013. Detecting the dynamic linkage between landscape characteristics and water quality in a subtropical coastal watershed, Southeast China. Environ. Manage. 51 (1), 32–44.
Jaksik, O., Kodesova, R., Kubis, A., Stehlikova, I., Drabek, O., Kapicka, A., 2015. Soil aggregate stability within morphologically diverse areas. Catena 127, 287–299.
Jia, X.X., Shao, M.A., Wei, X.R., Horton, R., Li, X.Z., 2011. Estimating total net primary productivity of managed grasslands by a state-space modeling approach in a small catchment on the Loess Plateau, China. Geoderma 160, 281–291.
Kemper, W., Rosenau, R., 1986. Aggregate stability and size distribution. In: Klute, A. (Ed.), Methods of Soil Analysis, Part 1: Physical and Mineralogical Methods (Second Edition). Soil Science Society of America, Madison, Wisconsin, USA, pp. 425–442.
Klami, A., Virtanen, S., Kaski, S., 2013. Bayesian canonical correlation analysis. J. Mach. Learn. Res. 14 (Apr), 965–1003.
Li, Q.Q., Yue, T.X., Wang, C.Q., Zhang, W.J., Yu, Y., Li, B., Yang, J., Bai, G.C., 2013. Spatially distributed modeling of soil organic matter across China: an application of artificial neural network approach. Catena 104, 210–218.
Li, Y.Y., Liu, L., An, S.S., Zeng, Q.C., Li, X., 2016. Research on the effect of vegetation and slope aspect on the stability and erodibility of soil aggregate in loess hilly region

based on le bissonnais method. J. Nat. Resour. 31, 287–298 (in Chinese with English abstract).
Liu, H.Y., Yin, Y., Piao, S.L., Zhao, F.J., Engels, M., Ciais, P., 2013. Disappearing lakes in Semiarid Northern China: drivers and environmental impact. Environ. Sci. Technol. 47 (21), 12107–12114.
Liu, Y.L., Feng, Y.H., Zhao, Z., Zhang, Q.W., Su, S.L., 2016. Socioeconomic drivers of forest loss and fragmentation: a comparison between different land use planning schemes and policy implications. Land Use Policy 54, 58–68.
Ma, L.G., Ma, F.L., Li, J.D., Gu, Q., Yang, S.T., Wu, D., Feng, J., Ding, J.L., 2017. Characterizing and modeling regional-scale variations in soil salinity in the arid oasis of Tarim Basin, China. Geoderma 305, 1–11.
McGarigal, K., Marks, B.J., 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. USDA-Forest Service, Portland, Oregon, USA.
McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002. FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. http://www.umass.edu/landeco/research/fragstats/fragstats.html.
McGarigal, K., Cushman, S.A., Ene, E., 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Available at the following web site:. Computer software program produced by the authors at the University of Massachusetts, Amherst. http://www.umass.edu/landeco/research/fragstats/fragstats.html.
McKeague, J., Day, J., 1966. Dithionite-and oxalate-extractable Fe and Al as aids in differentiating various classes of soils. Can. J. Soil Sci. 46 (1), 13–22.
Mehra, O., Jackson, M., 1958. Iron oxide removal from soils and clays by a dithionite-citrate system buffered with sodium bicarbonate. National Conference on Clays and Clays Minerals. pp. 317–327.
Mirzaee, S., Ghorbani-Dashtaki, S., Mohammadi, J., Asadi, H., Asadzadeh, F., 2016. Spatial variability of soil organic matter using remote sensing data. Catena 145, 118–127.
Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrol. Processes 5 (1), 3–30.
Mulder, V., De Bruin, S., Schaepman, M., Mayr, T., 2011. The use of remote sensing in soil and terrain mapping—a review. Geoderma 162, 1–19.
Mummey, D.L., Clarke, J.T., Cole, C.A., O'Connor, B.G., Gannon, J.E., Ramsey, P.W., 2010. Spatial analysis reveals differences in soil microbial community interactions between adjacent coniferous forest and clearcut ecosystems. Soil Biol. Biochem. 42 (7), 1138–1147.
Nelson, D.W., Sommers, L.E., 1996. Total carbon, organic carbon and organic matter. In: Spark, D.L., Page, A.L., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis. Part 3: Chemical Methods. ASA/SSSA, Madison, WI, USA, pp. 961–1010.
Ou, Y., Rousseau, A.N., Wang, L.X., Yan, B.X., 2017. Spatio-temporal patterns of soil organic carbon and pH in relation to environmental factors—a case study of the Black Soil Region of Northeastern China. Agric. Ecosyst. Environ. 245, 22–31.
Ouyang, W., Skidmore, A.K., Hao, F.H., Wang, T.J., 2010. Soil erosion dynamics response to landscape pattern. Sci. Total Environ. 408 (6), 1358–1366.
Ozgoz, E., Gunal, H., Acir, N., Gokmen, F., Birol, M., Budak, M., 2013. Soil quality and spatial variability assessment of land use effects in a Typic Haplustoll. Land Degrad. Dev. 24 (3), 277–286.
Oztas, T., Fayetorbay, F., 2003. Effect of freezing and thawing processes on soil aggregate stability. Catena 52 (1), 1–8.
Regelink, I.C., Stoof, C.R., Rousseva, S., Weng, L.P., Lair, G.J., Kram, P., Nikolaidis, N.P., Kercheva, M., Banwart, S., Comans, R.N.J., 2015. Linkages between aggregate formation, porosity and soil chemical properties. Geoderma 247, 24–37.
Rosemary, F., Indraratne, S., Weerasooriya, R., Mishra, U., 2017. Exploring the spatial variability of soil properties in an Alfisol soil catena. Catena 150, 53–61.
Sandholt, I., Rasmussen, K., Andersen, J., 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sens. Environ. 79 (2–3), 213–224.
Shi, Z.H., Ai, L., Li, X., Huang, X.D., Wu, G.L., Liao, W., 2013. Partial least-squares regression for linking land-cover patterns to soil erosion and sediment yield in watersheds. J. Hydrol. 498, 165–176.
Soinne, H., Hyvaluoma, J., Ketoja, E., Turtola, E., 2016. Relative importance of organic carbon, land use and moisture conditions for the aggregate stability of post-glacial clay soils. Soil Tillage Res. 158, 1–9.
Taghizadeh-Mehrjardi, R., Nabiollahi, K., Kerry, R., 2016. Digital mapping of soil organic carbon at multiple depths using different data mining techniques in Baneh region. Iran. Geoderma 266, 98–110.
Teng, M.J., Zeng, L.X., Xiao, W.F., Huang, Z.L., Zhou, Z.X., Yan, Z.G., Wang, P.C., 2017. Spatial variability of soil organic carbon in Three Gorges Reservoir area. China. Sci. Total Environ. 599-600, 1308–1316.
Turner, M.G., 2005. Landscape ecology: what is the state of the science? Annu. Rev. Ecol. Evol. Syst. 36, 319–344.
Veihe, A., 2002. The spatial variability of erodibility and its relation to soil types: a study from northern Ghana. Geoderma 106 (1–2), 101–120.
Wang, K., Zhang, C.R., Li, W.D., 2013. Predictive mapping of soil total nitrogen at a regional scale: a comparison between geographically weighted regression and cokriging. Appl. Geogr. 42, 73–85.
Wang, X.Y., Wang, T., Liu, D., Guo, H., Huang, H.B., Zhao, Y.T., 2017. Moisture-induced greening of the South Asia over the past three decades. Glob. Change Biol. Bioenergy 23 (11), 4995–5005.
Wei, G.X., Zhou, Z.F., Guo, Y., Dong, Y., Dang, H.H., Wang, Y.B., Ma, J.Z., 2014. Long-Term Effects of Tillage on Soil Aggregates and the Distribution of Soil Organic Carbon, Total Nitrogen, and Other Nutrients in Aggregates on the Semi-Arid Loess Plateau, China. Arid Land Res. Manage. 28 (3), 291–310.
Wilding, L.P., 1985. Spatial variability: its documentation, accommodation, and implication to soil surveys. In: Nielsen, D.R., Bouma, J. (Eds.), Soil Spatial Variability. Pudoc, Wageningen, the Netherlands, pp. 166–194.
Yan, A., 2015. Spatial Distribution and Storages Estimation of Soil Organic Carbon and Soil Inorganic Carbon in Xinjiang, China. China Agricultural University (in Chinese with English abstract).
Ye, L.P., Tan, W.F., Fang, L.C., Ji, L.L., Deng, H., 2018. Spatial analysis of soil aggregate stability in a small catchment of the Loess Plateau, China: I. Spatial variability. Soil Tillage Res. 179, 71–81.
Zhao, J.S., Chen, S., Hu, R.G., Li, Y.Y., 2017. Aggregate stability and size distribution of red soils under different land uses integrally regulated by soil organic matter, and iron and aluminum oxides. Soil Tillage Res. 167, 73–79.
Zolfaghari, Z., Mosaddeghi, M.R., Ayoubi, S., 2016. Relationships of soil shrinkage parameters and indices with intrinsic soil properties and environmental variables in calcareous soils. Geoderma 277, 23–34.