CLEAN

Soil Air Water

Renewables

Sustainability

Environmental Monitoring

Dongli She^{1,2}
 Jiaying Zheng¹
 Ming'an Shao³
 Luis Carlos Timm⁴
 Yongqiu Xia²

Research Article

Multivariate Empirical Mode Decomposition Derived Multi-Scale Spatial Relationships between Saturated Hydraulic Conductivity and Basic Soil Properties

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

²State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, P. R. China

³Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, Beijing, P. R. China

⁴Department of Rural Engineering, Faculty of Agronomy, Federal University of Pelotas, Pelotas, RS, Brazil

Saturated hydraulic conductivity (K_s) is affected by various factors operating at different scales. This study identified the multi-scale spatial relationships between K_s and selected basic soil properties (soil organic matter [SOM], clay, silt, and sand contents, and bulk density) along two landscape transects (with various soil textures and land use covers) on the Loess Plateau. Multivariate empirical mode decomposition (MEMD) yielded four different intrinsic mode functions (IMFs) for the multivariate data series of each transect according to the scale of occurrence. The dominant scales in terms of explained variance of K_s were IMF1 (scale: 403 m) for transect 1, and IMF1 and IMF2 (scale: 407 and 775 m) for transect 2. The multi-scale correlation between K_s and soil properties was more complex for transect 1 due to a more fragmented landscape. For each IMF or residue, K_s was predicted using the identified factors that significantly affected it at that IMF scale or residue. The summation of the four predicted IMFs and the residue predicted K_s at the measurement scale, and was more accurate than predictions based on simple multiple linear regressions between K_s and the other soil properties. Soil particle size components were the main contributors in explaining K_s variability for both landscape transects, mostly due to their contributions from IMF1; however, SOM was also a major contributor for transect 2, mainly due to contributions from IMF2. Using MEMD has great potential in characterizing scale-dependent spatial relationships between soil properties in complicated landscape ecosystems.

Keywords: Intrinsic mode functions; Landscape transects; Multivariate empirical mode decomposition; Scale; Soil hydraulic property

Received: March 8, 2014; *revised:* June 12, 2014; *accepted:* June 18, 2014

DOI: 10.1002/clean.201400143

1 Introduction

Saturated hydraulic conductivity (K_s) is one of the most important soil quality indicators [1]. Quantifying K_s is essential for describing and modeling hydrological and biological processes, as well as for improving soil erosion control and other soil-water management practices, such as reducing soil-water pollution [2–4]. However, K_s in soils exhibits a great degree of heterogeneity, and obtaining sufficient and reliable K_s data by direct in situ determination methods (e.g., by using double ring or tension disk infiltrometers, etc.) is expensive and time consuming. Hence, one of the challenges is to develop a better

understanding of the nature of the spatial variability and the scaling properties of K_s and to develop efficient strategies for the accurate estimation of K_s distributions at various scales [2, 3, 5].

The variability of K_s in agricultural fields has been studied extensively at different scales [1, 6–8]. These studies showed that inherent soil heterogeneity and extrinsic factors simultaneously affected spatial distribution of K_s at a given scale. Therefore, the establishment of rigorous relationships between K_s and the factors affecting it has become an alternative method to indirectly understand the nature of K_s spatial variability. However, the relationships between the spatial distributions of these affecting factors and those of K_s itself were detected to be highly changed among the various scales studied in the past [3, 6, 9]. Any individual factor might have different effects on K_s at different scales. Thus, when considering multiple factors, superimposing processes of different natures that act simultaneously among various scales often gave rise to chaotic and non-linear types of K_s distribution [10, 11]. Consequently, accuracy and reliability were usually low when using statistically derived models to estimate K_s distributions.

The Loess Plateau of China is a region with deep loess deposits and is susceptible to severe soil erosion. Vegetation restoration is

Correspondence: Dr. Y. Xia, State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, P. R. China
E-mail: yqxia@issas.ac.cn

Abbreviations: BD, bulk density; CV, coefficient of variation; EMD, empirical mode decomposition; IMF, intrinsic mode function; K_s , saturated hydraulic conductivity; MEMD, multivariate empirical mode decomposition; NRMSE, normalized root mean square error; R^2 , coefficient of determination; RMSE, root mean squared error; SOM, soil organic matter

currently the principal means by which soil erosion is being controlled and ecosystem recovery is being achieved in this area [12]. However, changes in land use and vegetation cover have led to a patchwork or mosaic land use type pattern that has caused the distribution of hydrological response variables (e.g., soil moisture, K_s) to become more chaotic and non-linear [4, 13]. So, research on the scale-dependent relationships between these hydrological variables and the factors affecting them in different landscapes could be very interesting and important, especially to the soil–water management practices during the vegetation restoration process.

Recent studies on the multi-scale influences of environmental factors on the distributions of hydraulic variables have used Pearson linear correlation analysis [13], wavelet analysis [14], or multi-fractal analysis [15]. These methods assumed that the distributions of the hydraulic variables and of the related processes were linear but this is contrary to the nature of their distribution. In order to avoid using this assumption, Hu and Si employed multivariate empirical mode decomposition (MEMD) [16], which is an extended empirical mode decomposition (EMD) algorithm [17], to identify the scale-specific factors that control soil moisture condition on the Loess Plateau. The MEMD was pioneered by Huang et al. who developed it as a way of analyzing data empirically as a form of wavelet analysis [18]. In 2004, Flandrin and Goncalves attempted to lend some statistical weight to the method [19]. Despite the demonstrable usefulness of the method, it has only existed for 16 years and has yet to be widely applied to resolve soil property scaling. Hu and Si, using this analysis for two transects with different land use (korshinsk peashrub and bunge needlegrass), the overall spatial patterns of soil moisture in the 0–6 cm soil layer and of environmental attributes (organic carbon contents, sand, silt, clay, and elevation) were separated into different intrinsic mode functions (IMFs) with similar scales [16]. The assumption was that the distribution of soil water at any given scale would be affected by the environmental attributes that were identified by MEMD for the same scale. When Hu and Si predicted the IMF of soil moisture for each scale using the corresponding IMF of the environmental factors, they obtained more accurate overall

estimations of soil water than would have been predicted by simple linear regressions [16].

There is limited information on the relationships between K_s and some basic soil attributes scale-dependently (e.g., texture/particle-size distribution, soil organic matter [SOM], and bulk density [BD]) that are usually available from soil survey databases and especially not at larger scales where the land use distributes as a mosaic pattern. Therefore, the scope of this presented research was to detect the multi-scale spatial relationships between K_s and basic soil properties along two landscape transects (where soil and vegetation types changed) in a catchment on the Loess Plateau of China. Specific objectives included: (i) comparing the dominant scale of variation of the non-linear and non-stationary K_s between the two landscape transects; and (ii) estimating the measurement scale K_s distribution of the two landscape transects based on the scale-specific relationships between K_s and basic soil properties.

2 Materials and methods

2.1 Site description

Data were collected from the Liudaogou catchment (110°21' to 110°23' E and 38°46' to 38°51' E) of the Loess Plateau, which is located in Shenmu County, Shaannxi Province, China (Fig. 1). The study area is characterized by a semiarid climate, with a mean annual rainfall of 437 mm, whereof about 70% falls in intense storms from June to September. The mean annual temperature is 8.4°C. The desiccation degree is 1.8 with a potential evapotranspiration of 785 mm/year. The dominant soils are loessal mein soil (Calcaric regosol), red loessal soil (Eutric regosol), warp soil (Calcaric fluvisol), and aeolian sand soil (Calcaric arenosol) [20]. The main soil textural classes include sandy loam, loamy sand, and sand, which account 70, 17, and 13% of proportions throughout the entire study area, respectively. The study area is representative of the “crisscross zone”, which is subjected to severe erosion by both wind and water. Extensive destruction of natural vegetation had occurred in the

Figure 1. Geographic location of the study watershed in Shenmu County on the Loess Plateau of China and of the sampled transects.

watershed due to long-term human activity and severe wind–water erosion. Widespread vegetation restoration projects have been implemented in the region during the past decades resulting in the establishment of a patchwork of land use in the catchment. More detailed information about the vegetation in the watershed can be found in She et al. [13].

2.2 Experimental design

Soil saturated hydraulic conductivity was measured along two south–north transects passing through different landscape types on either side of the Liudaogou catchment along the main gully. Transects 1 and 2 were 3411 m (131 m intervals per sampling point, totally 27 sampling points) and 3587 m (124 m intervals per sampling points, totally 30 sampling) long, respectively; the southern ends of the transects were designated as the sampling origin points (Fig. 1). The most obvious differences between the two transects were in the distributions of the soil particle size composition and land use types (Table 1). The dominant soil textures were silt loam and sandy loam for transect 1 and sandy loam and sand for transect 2, which accounted for approximately 93 and 73% of all soil textural classes along each transect, respectively. Extensive destruction of natural vegetation and widespread re-vegetation projects have resulted in the establishment of a patchwork or mosaic pattern of land use [12]. Transect 1 passed through three vegetation types that covered >54% of the total catchment area, that is, grassland and forage, which together comprised 40% of the distribution along the transect, and shrubland (12%). Transect 2 was dominantly covered by woodland (23%) and shrubland (31%). Cropland, man-made structures, orchards, and gullies covered the remaining area. The land use along the two transects constituted a mosaic structure. Vegetation cover changed more frequently along transect 1 (about every 50–100 m) compared to transect 2 due to a greater degree of anthropogenic influences. The topography along the two transects was similar. Based on data from the sampling sites (see below), transect 1 had a mean elevation

of 1181 m (coefficient of variation, $CV = 2.8\%$) and the elevations ranged between 1126 and 1251 m; the mean slope gradient was 10.5° ($CV = 68\%$) and the slopes ranged between 0.5 and 23.3° . Transect 2 had a mean elevation of 1199 m ($CV = 2.5\%$) and the elevations ranged between 1124 and 1241 m; the mean slope gradient was 10.9° ($CV = 36\%$) and the slopes ranged between 0.5 and 18.7° .

At each selected sampling site, three undisturbed soil samples were collected using steel cylinders, 5 cm in length and 5 cm in diameter, from the soil surface layer (0–10 cm) at about 2.5–7.5 cm, in September 2007. Soil saturated conductivity (K_s) was measured by the constant head method [21], and soil bulk density was measured after oven-drying the cores at $105\text{--}110^\circ\text{C}$ for 24 h following the K_s measurements [12]. Soil organic matter content was measured by the Walkley–Black method [22]. Soil particle size (clay [<0.002 mm], silt [$0.002\text{--}0.05$ mm], and sand [>0.05 mm]) was determined by using a MasterSizer2000 laser particle size analyzer (Malvern Instruments, UK).

2.3 Data analysis

The data series of K_s along with those of the other five basic soil attributes composed a multivariate data series. The multivariate data series of each of the two transects was decomposed into different IMFs using MEMD. MEMD analysis is developed from the standard EMD, which has been used in signal processing and neural signal analysis [23, 24]. The MEMD analysis was carried out using the MATLAB program (MathSoft), which could make projections of multiple input data with various directions in an n -dimensional space to produce multiple n -dimensional envelopes, aligning “common scales” present within multivariate data. Based on the MEMD algorithm, each “common scale” was identified in the common oscillatory modes for all variates within the n -variate intrinsic mode function [25]. MEMD then grouped similar scales among the different data sources to stand for the real scales of the related processes. A more detailed description of MEMD analysis theory can be found in Rehman and Mandic [25] and Hu and Si [16].

Table 1. Comparison of land use composition and soil texture along the two transects

Transect	Land use			Percentage distribution of soil textures (%)			
	Land use type	Main plant species	Percentage distribution (%)	Silt loam	Sandy loam	Loamy sand	Sand
Transect 1	Grassland	Bunge needlegrass (<i>Stipa bungeana</i> Trin.)	24.5	55.6	37.0	3.7	3.7
	Forage	Alfalfa (<i>Medicago sativa</i>)	15.7				
	Shrubland	Korshinsk peashrub (<i>Caragana korshinskii</i> Kom.)	12.2				
	Orchard	Apricot (<i>Prunus armeniaca</i>)	6.9				
	Fallow	Annual grass	9.8				
	Cropland	Millet (<i>Setaria italica</i> (L.) Beauv.), Bean (<i>Phaseolus vulgaris</i>)	13.5				
	Man-made structure	N/A	4.2				
	Gully channels	N/A	11.8				
Transect 2	Other	N/A	1.4	20.0	40.0	6.7	33.3
	Woodland	Locust (<i>Robinia pseudoacacia</i>), Simon poplar (<i>Populus simonii</i>)	23.2				
	Grassland	Bunge Needlegrass (<i>Stipa bungeana</i> Trin.)	20.4				
	Shrubland	Korshinsk peashrub (<i>Caragana korshinskii</i> Kom.)	31.2				
	Fallow	Annual grass	14.4				
	Gully channels	N/A	9.7				
	Other	N/A	1.1				

All other statistical analyses including data frequency distributions and normality tests were carried out by SPSS (version 17.0) and Microsoft Excel (version 2007) packages. Pearson correlation coefficients between K_s and the factors considered as being likely to affect it were calculated based on the entire datasets from each transect, that is, at the measurement scale, as well as for different IMFs or residues. Soil hydraulic conductivity at each scale and residue was predicted from the relevant IMFs and residues of the basic soil properties by stepwise multiple linear regression models. The predicted K_s at the measurement scale was then derived by adding all predicted values for each IMF and residue of K_s , using the following equation:

$$K_s^p = \sum_i^m \text{IMF}_i^p + R_e^p \quad (1)$$

where K_s^p is the predicted value of K_s at the measurement scale; m is the number of IMF; IMF_i^p is the predicted K_s value for each IMF; and R_e^p is the predicted K_s value for residue.

Finally, the quality of K_s prediction was evaluated by the coefficient of determination (R^2), root mean squared error (RMSE) and normalized root mean square error (NRMSE) from the estimated and measured values of K_s , using the following equations [26]:

$$R^2 = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - O)^2} \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (3)$$

$$\text{NRMSE} = \text{RMSE}/O \quad (4)$$

where n is the number of K_s observations; O_i and P_i are the observed and predicted K_s values at each sampling site i , respectively; and O is the mean of all observed K_s values.

3 Results and discussion

3.1 Spatial patterns of K_s at the measurement scale

The spatial distributions of K_s and of the other soil properties are shown in Fig. 2, and the descriptive statistical results on the overall variability are presented in Table 2. K_s in transect 1 (mean K_s value = 1.52 cm/h) and its coefficients of variation (CV = 57.5%) were significantly smaller than those in transect 2 (mean K_s value = 2.02 cm/h, CV = 74.7%). These coefficients of variation of K_s were comparable to those (CV = 98.3%) reported by Gao et al. for a relatively small watershed (0.5 km²) on the Loess Plateau [27]. However, they were notably lower than those (CV = 206%) reported by Wang et al. for the entire Loess Plateau (620 000 km²) [28]. The greater variability of K_s along transect 2, especially between 400 and 2500 m, as compared with transect 1, was attributed partly to the greater variability in the soil properties along that transect (Fig. 2).

The degree of linear association between K_s and the soil properties at the observation scale were evaluated using Pearson's correlation analysis. For transect 1, K_s was significantly related to clay, silt, and sand contents, but not to SOM and BD, implying that at this scale the

Figure 2. Spatial distribution of saturated hydraulic conductivity and other soil properties along the sampled transects (blue for transect 1, magenta for transect 2).

Table 2. Means and variances of saturated hydraulic conductivity (K_s) and other soil properties, with the correlation coefficients (R) between them and K_s as well as the stepwise multiple linear regression models at the measurement scale

	Mean	CV (%)	R with K_s
Transect 1			
K_s (cm/h)	1.52	57.5	N/A
SOM (g/kg)	3.22	26.0	-0.157
Clay (%)	5.77	43.1	-0.437*
Silt (%)	48.52	27.2	-0.710*
Sand (%)	45.71	33.8	0.678*
BD (g/cm ³)	1.36	7.7	0.344
$K_s = 4.466 - 0.095 \text{ Silt}$ ($F^{(3)} = 23.144$; Adjusted $R^2 = 0.630$)			
Transect 2			
K_s (cm/h)	2.02	74.7	N/A
SOM (g/kg)	2.43	39.0	-0.661**
Clay (%)	3.13	82.7	-0.681**
Silt (%)	31.07	62.2	-0.767**
Sand (%)	65.80	33.2	0.759**
BD (g/cm ³)	1.45	7.2	0.526**
$K_s = 72.706 - 0.831 \text{ Silt} - 0.682 \text{ Sand}$ ($F = 26.761$; Adjusted $R^2 = 0.640$)			

*Significant at $p < 0.05$; ** significant at $p < 0.01$.

^aThe F statistic for the accepted model.

variability in K_s distribution was mainly due to the soil texture and was influenced by the balance of the effects of clay, silt, and sand contents. More important local scale processes, such as biological activities and tillage induced perturbations, apparently obscured

the other expected relationships so that they were found to be non-significant. However, all selected soil properties were significantly correlated with K_s in transect 2, and all correlation coefficients were higher than those for transect 1. Stepwise multiple linear regressions were used to obtain models with one or more factors that could significantly predict K_s (Table 2). Factors that were selected by this procedure (silt content alone for transect 1, and silt and sand contents for transect 2) explained 63 and 64% of the variance in K_s for transects 1 and 2, respectively.

3.2 Multi-scale influences of soil properties on K_s distribution

To examine the correlation at different scales, the multivariate data series of K_s and soil properties for the two studied transects were separated into four different modes of oscillation (IMF1-IMF4) by MEMD (Fig. 3). Numbers and widths of oscillations for the IMFs having the same numerical numbers differed between K_s and the other soil properties. Thus, the identified scales of K_s and the soil attributes derived from the momentary frequencies of the IMFs using the Hilbert transform varied for any given IMF of either transect [16, 25]. The differences in the identified scales increased when the IMFs became larger, which meant that there were no common scales present among K_s and the basic soil properties at large scales (Table 3).

The scale of a specific IMF was equal to the mean scale of K_s and the other soil properties for that specific IMF. The mean scales of IMF1 and IMF2 were similar for the two transects but those of IMF3 and IMF4 differed greatly (Table 3). The IMF1 might represent localized

Figure 3. IMFs and residues for the multivariate data series of saturated hydraulic conductivity and other soil properties along two landscape transects (blue for transect 1, magenta for transect 2) separated by multivariate empirical mode decomposition analysis. The numbers in parentheses are the percentage of variance explained by each (IMF) for two transects.

Table 3. Scale of each IMF of K_s and other soil properties

		K_s	SOM	Clay	Silt	Sand	BD	Mean ^{a)}	CV (%) ^{b)}
Transect 1	IMF1	450	489	401	358	358	359	403	13.9
	IMF2	750	929	733	730	730	725	766	10.5
	IMF3	1374	1785	1713	1347	1698	1294	1535	14.3
	IMF4	2545	3359	2257	3457	3455	2258	2889	20.7
Transect 2	IMF1	433	407	386	387	387	443	407	6.2
	IMF2	730	749	761	759	759	890	775	7.4
	IMF3	1124	1424	1284	1315	1311	1294	1292	7.5
	IMF4	3004	3559	2529	2139	2137	2505	2646	20.8

^{a)} Mean scale of K_s and other soil properties.

variations over relatively short distances of about 400 m (403 and 407 m for transects 1 and 2, respectively). The mean values of the identified scales for IMF2 were also similar for transects 1 (766 m) and 2 (775 m). Furthermore, the scales of the K_s series in either transect were also similar for IMFs with lower numerical values, implying that some local processes controlling K_s also occurred at similar scales. For IMF3 and IMF4, the scales for transect 1 (1535 and 2889 m, respectively) were notably larger than those for transect 2 (1292 and 2646 m). This implied that the variations in land use cover and soil types could change the scales at which K_s and the other soil properties operated.

The distributions of variance among the different IMFs were different between the two transects (Fig. 3). For transect 1, the greatest percentage of the total variance of K_s (40.9%) was contributed by IMF1 whereas ~85% of the variance was explained by IMF1 and IMF2 together for transect 2. From Fig. 2, we could compare that these scales were similar with the distances of the soil particle composition (~400 m) and SOM (~800 m) cycles of the two transects (Fig. 2). The variance distributions were similar for K_s and the other soil properties for each transect (Fig. 3). Some variation (19.7 and 12.5% of the total variance of K_s for transects 1 and 2, respectively) was found in the residue, which might be due to the inability to identify larger scales that would exceed the length of the transects by the MEMD analysis. In addition, the sum of the percentage variances of the four IMFs and the residue for each variable did not equal 100%. Hu and Si suggested that this might be attributed to the shortage of robust orthogonality between the different IMFs of a given variable [16].

Table 4. Correlation coefficients between saturated hydraulic conductivity and soil properties for each IMF and residue based on multivariate empirical mode decomposition for all the variables

		SOM	Clay	Silt	Sand	BD
Transect 1	IMF1	-0.153	-0.133	-0.439*	0.394*	0.226
	IMF2	-0.154	-0.082	-0.641**	0.585**	0.460**
	IMF3	-0.177	-0.183	-0.216	0.210	0.945**
	IMF4	0.879**	-0.875**	-0.937**	0.929**	-0.614**
	Residue	-0.965**	-0.999**	-0.999	0.999	0.983**
Transect 2	IMF1	-0.615**	-0.599**	-0.664**	0.659**	0.215
	IMF2	-0.889**	-0.666**	-0.766**	0.757**	0.696**
	IMF3	-0.699**	-0.643**	-0.761**	0.746**	0.759**
	IMF4	0.317**	-0.651**	-0.515	0.533**	0.994**
	Residue	0.994	-0.989	-0.986	0.987	1.000

*Significant at $p < 0.05$; **significant at $p < 0.01$.

Multi-scale correlation coefficients between K_s and soil properties varied at the transect measurement scales (Table 4). For transect 1, the relationships between K_s and soil properties were more complex than those for transect 2, which was partly explained by the greater diversity in landscapes along transect 1 due to a greater degree of human activity. For transect 1, K_s was correlated significantly with silt and sand for IMF1, while BD became an additional significant factor for IMF2 and was the only significant factor for IMF3; K_s was significantly correlated with all five soil properties for IMF4. However, for transect 2, K_s was significantly correlated with all five soil properties for all four IMFs with the exception of the correlation between K_s and BD for IMF1. The correlation coefficients between K_s and the soil properties were greater for transect 2 than those for transect 1, which suggested the feature of the simpler land use pattern along transect 2 leading to simpler relationships between K_s and all five soil properties. In addition, the relationships between K_s and the soil attributes were significant ($|r| > 0.9$) for the residues for both transects, which indicated a strong influence of the selected soil properties on K_s .

3.3 K_s prediction based on its multi-scale relationships with soil properties

Table 5 shows the equations for predicting K_s for each IMF that were derived by stepwise multiple linear regression based on its multi-scale relationships with soil properties. The adjusted R^2 values for all prediction models ranged from 0.54 to 1.00, and increased with the IMF number. These trends in the adjusted R^2 values indicated that at larger scales the models were usually more accurate when predicting K_s and that the influences of the selected soil properties might be more dominant in affecting soil hydrological processes. The K_s for the measurement scale was estimated by adding all predicted IMFs and the residue. The determining coefficients (R^2) between the measured K_s and the estimated K_s at the measurement scale were 0.798 and 0.834 for transects 1 and 2, respectively, which were higher than those (0.63 and 0.64) obtained when using the model derived from the entire dataset at the measurement scale. The RMSEs and NRMSEs for the results obtained by MEMD analysis were notably lower than those obtained by simple multiple linear regressions (Table 6). Based on this comparison of the evaluation parameters, it can be concluded that the K_s prediction using MEMD was better than the K_s estimations based on the original data at the measurement scale.

Figure 4a shows the R^2 between each estimated IMF (residue) and the overall K_s at the measurement scale. A comparison of the coefficients, which reflect the relative importance of each scale (IMF)

Table 5. Predictive equations and regression statistics (F -value and adjusted R^2) for soil saturated hydraulic conductivity for each IMF using stepwise multiple linear regression based on multivariate empirical mode decomposition

Model			F^a	Adjusted R^2
Transect 1	IMF1	$-0.002 + 0.454 (0.098^{b})$ Clay $- 0.107 (1.019)$ Silt	9.0	0.54
	IMF2	$0.042 + 0.199 (0.062)$ Clay $- 0.064 (0.211)$ Silt	17.2	0.56
	IMF3	$0.038 + 8.497 (0.586)$ BD	210.3	0.89
	IMF4	$-0.016 + 1.166(0.030)$ Clay $- 0.283 (0.005)$ Silt $- 5.378 (0.684)$ BD	5.9E+3	1.00
	Residue	$3.906 + 0.335 (0.000)$ SOM $- 0.073 (0.000)$ Silt	4.1E+6	1.00
Transect 2	IMF1	$0.088 - 1.198 (0.247)$ SOM $- 0.782 (0.278)$ Silt $- 0.642 (0.244)$ Sand	21.9	0.68
	IMF2	$-0.006 - 1.325 (1.171)$ SOM $+ 5.368 (0.051)$ BD	70.9	0.83
	IMF3	$0.00 - 0.419 (0.021)$ SOM $- 1.349 (0.026)$ Silt $- 1.175 (0.021)$ Sand $+ 11.134 (0.195)$ BD	1.6E+4	1.00
	IMF4	$0.00 - 0.219 (0.006)$ SOM $+ 1.097 (0.013)$ Clay $+ 0.115 (0.002)$ Sand $+ 11.999 (0.066)$ BD	3.1E+5	1.00
	Residue	$-10.710 - 0.424 (0.010)$ SOM $+ 0.188 (0.004)$ Clay $+ 9.164 (0.029)$ BD	8.5E+6	1.00

^{a)} F statistics for the accepted models.

^{b)} Number in parenthesis is the standardized regression coefficient.

to the predictive K_s at the measurement scale, indicates that IMF1 was the main contributor to the overall prediction of K_s for transect 1, while it was clearly IMF2 for transect 2. The notion of spatial continuity of the landscape pattern could be used to explain these differences. Note that continuity in the spatial distribution of K_s for transect 1 was expected because of the relative lower variations in the soil properties (Table 2). However, vegetation cover also transformed more frequently (about every 50–100 m) along transect 1 due to the greater degree of human activity, which resulted in the occurrence of a high number of small cropland patches along the transect. Local scale processes, such as tillage induced perturbations, resulted in the variation or fluctuation in macropore distribution, soil particle composition and root biomass [3], which caused a similar discontinuity in the K_s distribution at the smaller scale (IMF1). In contrast, the lesser degree of human activity along transect 2 resulted in less landscape diversity and fragmentation than those along transect 1 [29]. The discontinuity of the landscape along transect 2 resulted mainly from the divisions created by erosion gullies. In this case, the vegetation cover on slopes separated by an erosion gully was similar with scales almost equal to the scale of IMF2 along transect 2. In contrast, the soil texture changed markedly for different vegetation-slope patches. Under these conditions, the sources of K_s variation would be better detected at the medium scale (IMF2) for transect 2. In addition, the predicted residue, which might only represent larger scales, illustrated a relatively low amount of the total variation in K_s for the two transects, which indicated that the leading processes had similar scales of measurement scale, especially for transect 2.

Figure 4b shows the R^2 between the total predictive K_s from different scales for each variable and the overall K_s at the measurement scale, which indicates that the relative importance of each explanatory variable for the K_s prediction at the

measurement scale was different. Silt content and BD were the dominant explanatory factors for K_s prediction for transect 1. The variations of these two variables mainly resulted from the combined influences of inherent soil heterogeneity, vegetation, topography, and human activities, which also have significant influences on K_s distribution [17]. As shown in Table 5, the silt content played a more important role in predicting K_s distributions at the scales of IMF1, as indicated by the absolute values of the standardized regression coefficients in parentheses following the parameter's coefficient. This result indicated that the effect of silt on predicted K_s values at the measurement scale was most obvious at the smaller scales, for example, at 400 m. Soil bulk density was important in K_s prediction at the scales of IMFs 3 and 4. This result indicated that the effect of BD on the predicted K_s values at the measurement scale was most obvious at the larger scales, for example, at 1535 and 2889 m. Then, at the measurement scale, the correlation coefficient between K_s and BD was negligible. This meant that the contribution of BD would be overlooked when using the traditional regression method. SOM and sand had little contribution to the overall K_s prediction for transect 1. However, SOM was clearly one of the dominant explanatory factors for the overall K_s prediction for transect 2. This difference between the two transects was supported by the different correlation coefficients between K_s and SOM at the measurement scale (Table 2). IMF2 for SOM was the most important IMF in predicting K_s , indicated by the high absolute values of its standardized regression coefficients (Table 5). Therefore, the greatest effect of SOM in the measurement scale K_s prediction was mainly produced at the medium scales, for example, at 775 m. Furthermore, the soil particle composition (silt and sand) also presented their contributions to the overall K_s prediction, mainly from IMF1 (scale of 407 m), as for transect 1. However, although BD had a significant relationship with K_s at the measurement scale, it presented little

Table 6. Values of evaluation parameters used to assess the overall K_s prediction accuracy

Prediction method		RMSE	NRMSE	R^2
Prediction by MEMD	Transect 1	0.224	0.147	0.798
	Transect 2	0.169	0.084	0.834
Prediction by stepwise multiple linear regression	Transect 1	0.471	0.309	0.627
	Transect 2	0.429	0.213	0.643

R^2 , coefficient of determination.

Figure 4. Coefficient of determination between (a) saturated hydraulic conductivity at the measurement scale and predicted IMFs (residue) or (b) total saturated hydraulic conductivity predicted by each variable.

contribution to the overall K_s estimation. A possible reason for this could be that BD variation mainly occurred at larger scales than the measurement scale (Fig. 3).

This research evidently showed that the relationships between K_s and some basic soil properties in terms of their spatial distribution and variability are a function of the scale. These relationships were more complex in a fragmented landscape, resulting in greater changes on the spatial scale of the biological and hydrological processes associated with each land use type. Single scale studies prevent understandings on hydrological processes that operate at multiple scales [3, 28]. By using MEMD analysis, the common scale for different IMFs of the multivariate data series of K_s and soil properties was revealed. The relative importance of each scale, or each affecting factor (soil property), to the spatial variability of K_s was different for different landscape transects, which has important implications for designing sampling strategies for the accurate estimation of K_s distributions at various scales. Given that the two transects have considerably different soil textures and land use covers, the results from this study could be useful in understanding K_s and the factors that affect it in complicated landscape ecosystems. Furthermore, using MEMD to identify different scale effects is useful in understanding hydraulic and other properties within a complex landscape.

4 Conclusions

In order to investigate the effect of scale on the relationships between K_s and five other soil properties, MEMD was used to separate the data in IMFs that reflected different scales of occurrence. Four different IMFs for a multivariate data series of K_s and the other soil properties along two landscape transects having various soils and land use covers were obtained. The dominant scales in terms of the percentages of total variances in K_s were identified (40.9% was explained by IMF1 [scale: 403 m] for transect 1, and about 85% by IMF1 [scale: 407 m] and IMF2 [scale: 775 m] together for transect 2). Multi-scale correlation coefficients between K_s and soil properties varied with measurement transects, which were more complex in transect 1 compared with transect 2 due to more landscape fragmentation resulting from more human activity along transect 1. Saturated hydraulic conductivity for a specific IMF (of a scale) or residue was predicted from the dominating variables identified at that scale or residue. Summing all the predicted IMFs and residue resulted in predictions of K_s at the measurement scale. IMF1 presented the main contribution to the overall prediction of K_s for transect 1, while IMF2 explained most of the variation in overall K_s

for transect 2. Soil texture was the dominating contributor to explain the K_s variation for both landscape transects, although different aspects of soil particle size composition were identified (silt for transect 1, silt and sand for transect 2), mainly by its contributions from IMF1 (scale: ~400 m). However, SOM was also a major contributor for K_s prediction for transect 2, mainly from IMF2 (scale: 775 m). The overall K_s estimation using MEMD outperformed that using simple multiple linear regression between K_s and the selected soil properties, showing that single scale (observation scale) analysis was not enough to fully describe the complex relationships between K_s and soil properties that are valid across all spatial scales.

Acknowledgements

We acknowledge the financial support provided by the National Nature Science Foundation of China through grants No. 41471180, the Open Funding Project (No. Y412201423) of the State Key Laboratory of Soil and Sustainable Agriculture (Institute of Soil Science, Chinese Academy of Sciences), the Excellent Creative Talents Support Program of Hohai University, and the Advanced Science and Technology Innovation Team in Colleges and Universities in Jiangsu Province.

The authors have declared no conflict of interest.

References

- [1] M. M. Moustafa, A geostatistical approach to optimize the determination of saturated hydraulic conductivity for large-scale subsurface drainage design in Egypt, *Agric. Water Manage.* **2000**, *42*, 291–312.
- [2] J. S. C. Mbagwu, Saturated hydraulic conductivity in relation to physical properties of soils in the Nsukka Plains, southeastern Nigeria, *Geoderma* **1995**, *68*, 51–66.
- [3] T. B. Zeleke, B. C. Si, Scaling relationships between saturated hydraulic conductivity and soil physical properties, *Soil Sci. Soc. Am. J.* **2005**, *69*, 1691–1702.
- [4] W. Hu, M. A. Shao, B. C. Si, Seasonal changes in surface bulk density and saturated hydraulic conductivity of natural landscapes, *Eur. J. Soil Sci.* **2012**, *63*, 820–830.
- [5] N. J. Jarvis, L. Zavattaro, K. Rajkai, W. D. Reynolds, P. A. Olsen, M. McGeachan, M. Mecke, et al., Indirect estimation of near-saturated hydraulic conductivity from readily available soil information, *Geoderma* **2002**, *108*, 1–17.
- [6] V. Comegna, P. Damiani, A. Sommella, Scaling the saturated hydraulic conductivity of a vertic ustorthens soil under conventional and minimum tillage, *Soil Till. Res.* **2000**, *54*, 1–9.

- [7] J. A. Sobieraj, H. Elsenbeer, G. Cameron, Scale dependency in spatial patterns of saturated hydraulic conductivity, *Catena* **2004**, 55, 49–77.
- [8] C. T. Petersen, A. Trautner, S. Hansen, Spatio-temporal variation of anisotropy of saturated hydraulic conductivity in a tilled sandy loam soil, *Soil Till. Res.* **2008**, 100, 108–113.
- [9] E. Perfect, R. W. Gentry, M. C. Sukop, J. E. Lawson, Multifractal Sierpinski carpets: Theory and application to upscaling effective saturated hydraulic conductivity, *Geoderma* **2006**, 134, 240–252.
- [10] J. D. Phillips, Stability implications of the state factor model of soils as a nonlinear dynamical system, *Geoderma* **1993**, 58, 1–15.
- [11] A. B. McBratney, Some considerations on methods for spatially aggregating and disaggregating soil information, *Nutr. Cycl. Agroecosyst.* **1998**, 50, 51–62.
- [12] D. L. She, M. A. Shao, L. C. Timm, K. Reichardt, Temporal changes of an alfalfa succession and related soil physical properties on the Loess Plateau, China, *Pesquisa Agropecuária Bras.* **2009**, 44(2), 189–196.
- [13] D. L. She, M. A. Shao, W. Hu, S. E. Yu, Variability of soil water-physical properties in a small catchment of the Loess Plateau, China, *Afr. J. Agric. Res.* **2010**, 5(22), 3041–3049.
- [14] Q. J. Liu, Z. H. Shi, N. F. Fang, H. D. Zhu, L. Ai, Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet-ANN approach, *Geomorphology* **2013**, 186, 181–190.
- [15] D. Wang, B. J. Fu, W. W. Zhao, H. F. Hu, Y. F. Wang, Multifractal characteristics of soil particle size distribution under different land-use types on the Loess Plateau, China, *Catena* **2008**, 72(1), 29–36.
- [16] W. Hu, B. C. Si, Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition, *Geoderma* **2013**, 193–194, 180–188.
- [17] D. L. She, D. D. Liu, S. Z. Peng, M. A. Shao, Multiscale influences of soil properties on soil water content distribution in a watershed on the Chinese Loess Plateau, *Soil Sci.* **2013**, 178(10), 530–539.
- [18] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, et al., The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. London, Ser. A* **1998**, 454, 903–995.
- [19] P. Flandrin, P. Gonclaves, Empirical mode decompositions as data-driven wavelet-like expansions, *Int. J. Wavelets Multi.*, 2, 1–20.
- [20] FAO, UNESCO, *Soil Map of the World, World Soil Resources*, Report 60, FAO, Rome **1988**.
- [21] Y. Y. Li, M. A. Shao, Change of soil physical properties under long-term natural vegetation restoration in the Loess Plateau of China, *J. Arid Environ.* **2006**, 64, 77–96.
- [22] D. W. Nelson, L. E. Sommers, Total carbon, organic carbon and organic matter, in *Methods of Soil Analysis* (Eds.: A. L. Page, R. H. Miller, D. R. Keeney), Part 2, Agronomy Monograph Vol. 9, 2nd ed., ASA-SSA, Madison, WI **1982**, pp. 534–580.
- [23] J. Fleureau, A. Kachenoura, L. Albera, J. C. Nunes, L. Senhadji, Multivariate empirical mode decomposition and application to multichannel filtering, *Signal Process.* **2011**, 91, 2783–2792.
- [24] M. Hu, H. L. Liang, Intrinsic mode entropy based on multivariate empirical mode decomposition and its application to neural data analysis, *Cognit. Neurodyn.* **2011**, 5, 277–289.
- [25] N. Rehman, D. P. Mandic, Multivariate empirical mode decomposition, *Proc. R. Soc. London, Ser. A* **2010**, 466, 1291–1302.
- [26] X. W. Han, M. A. Shao, R. Horton, Estimating van Genuchten model parameters of undisturbed soils using an integral method, *Pedosphere* **2010**, 20(1), 55–62.
- [27] L. Gao, M. A. Shao, Y. Q. Wang, Spatial scaling of saturated hydraulic conductivity of soils in a small watershed on the Loess Plateau, China, *J. Soil Sediment* **2012**, 12, 863–875.
- [28] Y. Q. Wang, M. A. Shao, Z. P. Liu, R. Horton, Regional-scale variation and distribution patterns of soil saturated hydraulic conductivities in surface and subsurface layers in the loessial soils of China, *J. Hydrol.* **2013**, 487, 13–23.
- [29] Z. L. Gao, X. M. Mu, Spatio-temporal change of land use/coverage in loess wind-water erosion crisscross region—take Liudaogou watershed as example, *J. Soil Water Conserv.* **2004**, 18(5), 146–150. (In Chinese with English abstract).