Geoderma xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Geoderma

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

Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS

Feng Liu [a,b,f], Xiaoyuan Geng [c], A-Xing Zhu [a,d,*], Walter Fraser [e], Arnie Waddell [e]

[a] State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, PR China
[b] Graduate University of the Chinese Academy of Sciences, Beijing 100049, PR China
[c] Agri-Environmental Services Branch (AESB), Agriculture and Agri-Food Canada (AAFC), 960 Carling Av. Ottawa, ON, Canada K1A 0C6
[d] Department of Geography, University of Wisconsin, Madison, Wisconsin, 53706, USA
[e] Cereal Research Centre, 195 Dafoe Road, Winnipeg, Manitoba, Canada R3T 2M9
[f] State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, PR China

ARTICLE INFO

Article history:
Received 30 April 2010
Received in revised form 7 May 2011
Accepted 7 May 2011
Available online xxxx

Keywords:
Low relief areas
Soil texture mapping
Land surface dynamic feedbacks
MODIS
SoLIM

ABSTRACT

In low relief areas such as plains, easily obtained soil forming factors generally do not co-vary with soil conditions over space to the level that they can be used effectively in digital soil mapping. Mapping variation of soil properties over such areas remains a challenge. This paper presents an approach to mapping soil texture using environmental covariates derived from temporal responses of the land surface to a rainfall event (dynamic feedbacks) collected through remote sensing techniques. The approach consists of four steps: (1) construction of a set of environmental covariates from dynamic feedbacks of the land surface, captured daily from MODIS (Moderate Resolution Imaging Spectroradiometer) images over a short period (6–7 days) after a major rain event; (2) derivation of environmental classes based on the set of environmental covariates using a fuzzy c-means clustering; (3) Identification of typical soil texture value for each of the environmental classes from a dataset of field soil samples; (4) mapping of spatial variation of soil texture through a linearly weighted averaging function. The approach was applied to produce soil texture maps in a low relief area situated in south-central Manitoba, Canada. Its performance was assessed through comparison with soil texture maps generated from 1:20,000 traditional soil survey. The assessment was based on 34 field sample sites, independent of the samples used for prediction. The error values (9.42 for MAE and 12.56 for RMSE) of A-horizon percentage of sand from the proposed approach are less than these from the detailed soil survey (10.59 for MAE and 15.12 for RMSE). Similar results were obtained for A-horizon percentage of clay. In addition, the difference between the results of multiple linear regression analysis without and with the MODIS derived variables further demonstrated the effectiveness of the variables at differentiating patterns of soil texture. These indicated that the proposed approach is effective for mapping the variation of soil texture over the low relief area and it could be used to map other soil property variation over similar areas.

1. Introduction

Based on the soil-landscape model concept (Jenny, 1941), difficult to measure soil types and soil properties can be associated with some easily obtained soil forming factors. This makes it possible to use the spatial variations of the easily obtained environmental factors to predict soil spatial variation. Obviously, the prediction requires that the easily obtained environmental factors are good indicators of spatial variation of soils. However, over low relief areas, such as plains and gently undulating terrains, relatively easily obtained environmental factors such as landform and vegetation generally cannot effectively indicate soil spatial variation (Ding et al., 1989; McKenzie and Ryan, 1999; Zhu et al., 2010a).

To overcome the difficulty of mapping soil texture in areas of low relief, some attempts have been made to predict the variation of soil texture using multispectral remote sensing (Coleman et al., 1993; Sullivan et al., 2005). The multispectral techniques are generally characterized by wide bandwidths and a limited number of spectral bands. Odeh and McBratney (2000) used AVHRR (Advanced Very High Resolution Radiometer) images acquired in the absence of vegetation to develop multiple linear regression models for predicting topsoil clay content in the lower Namoi Valley of New South Wales in Australia. Similarly, Demattê et al. (2009) applied multiple spectral bands (TM2, TM5, and TM7) of Landsat Thematic Mapper imagery to estimate the variation of surface clay content in a bare soil area located in the region of Barra Bonita, Brazil.

* Corresponding author at: State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, CAS, Beijing 100101, PR China. Tel.: +86 10 64888961; fax: +86 10 64889630.
E-mail addresses: liuf@lreis.ac.cn (F. Liu), axing@lreis.ac.cn, azhu@wisc.edu (A.-X. Zhu).

0016-7061/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2011.05.007

In contrast to the above, major efforts have also been made to map soil texture using hyperspectral remote sensing (Clark and Swayze, 1996; Selige et al., 2006). With narrow bandwidths (less than 20 nm) and many (tens or hundreds) spectral bands, the hyperspectral techniques possess high spectral resolution for delineating material characteristics. Lagacherie et al. (2008) used the data from laboratory, field and airborne hyperspectral measurements on bare soils to estimate topsoil clay. Significant relationships were observed between clay and reflectance values. They found that the performance of clay estimation decreased from the laboratory to airborne scales. Based on the same data, Gomez et al. (2008) argued that the partial least squares regression method was appropriate for accurate topsoil clay mapping with airborne hyperspectral data.

The multi/hyperspectral remote sensing methods only detect the top few millimeters of the soil surface. Furthermore, due to the difficulty of excluding the confounding influence of vegetation, their applications are usually limited to areas with exposed soil surfaces. How to retrieve soil texture information from deeper soil thicknesses and extend the application from bare soils to partially vegetated areas remains to be a challenge.

In an attempt to address this challenge, Zhu et al. (2010a) examined the idea that change patterns of the land surface in areas of low relief with partial vegetation cover, captured daily by the MODIS sensor over a short period after a major rain event, are related to soil spatial variation. MODIS was chosen because it has high temporal resolution when compared with other hyperspectral sensors such as EO-1 Hyperion. They used rainfall as the input to land surface. It was assumed that the occurrence of the input across a certain spatial extent was uniform. After the rain, the land surface begins a drying process from wet to dry, and it is this process that was considered as the feedback of the land surface in response to the rainfall. The characteristics of the drying process at a given location during the few days immediately after a rain event are referred to as the *land surface dynamic feedback pattern* for that location. Locations with different soil conditions should exhibit different feedback to the input if other surface conditions are the same. They found that areas with different soil types exhibited significantly different dynamic feedback patterns and areas with the same soil type had similar dynamic feedback patterns. Also, the more similar the soil types the more similar their feedback patterns. These findings provide a foundation for the development and application of new environmental covariates for digital soil mapping. The objective of this paper was to examine the utility of so derived environmental covariates for mapping spatial variation of soil texture over areas with low relief.

2. Materials and methods

2.1. Study area and data sets

The study area is situated in south-central Manitoba, Canada (Fig. 1). It occurs within the Manitoba Plain, and covers approximately 430 km^2 (49.45°N–49.61°N and 97.80° W–98.15°W). According to climate data from Graysville station (Environment Canada, 1993), mean annual temperature is 2.7 °C, and mean annual precipitation is 538.7 mm. The physiography is a direct result of advancing continental ice sheets during the last stages of the Pleistocene Period. The entire area was covered by several major advances of ice before the ice retreated to the north and glacial lakes were formed (Michalyna et al., 1988). The area is characterized by level to gentle undulating lacustrine sediments overlying glacial deposits. Thus, soil parent materials are glaciolacustrine in origin derived from sandstone, shale and limestone sedimentary bedrocks. Slope gradient mainly ranges from 0 to 1.5% with a mean of 0.6%. The gentle terrain makes it practically impossible to distinguish different slope locations (upper, middle and lower) on the digital elevation model (DEM) or in the field. According to the Soil Landscapes of Canada (SLC) maps, major soil subgroups are those of the Black Chernozem, Regosol, Humic Vertisol and Humic Gleysol Great Groups of the Canadian System of Soil Classification (Soil Classification Working Group, 1998). Soil texture exhibits great spatial variation and varies from sands to clays. Over 90% of the land is devoted to agricultural production and the major crops are wheat, barley and canola (Land Resource Unit, 1999). This area was chosen mainly because, in large portions of the area, landform and vegetation conditions are quite uniform but soil texture shows a high degree of spatial variability.

Two main data sets were used. One set was time series MODIS imagery. The other was field soil sample data. The MODIS sensor on board the polar orbiting satellite "Terra" was used to capture land surface feedbacks. Seven observation windows mainly designed for land surface observation were selected: blue (459–479 nm), green (545–565 nm), red (620–670 nm), near infrared (NIR1: 841–876 nm, NIR2: 1230–1250 nm), and shortwave infrared (SWIR1: 1628–1652 nm; SWIR2: 2105–2155 nm). MODIS daily surface reflectance data (MOD09GA and MOD09GQ, v005) were used in this research. The data were obtained though the NASA (National Aeronautics and Space Administration) Warehouse Inventory Search Tool (WIST, http://wist.echo.nasa.gov/api/). Among the bands, red and NIR1 bands are at 250 m spatial resolution, and the other five bands are with spatial resolution of 500 m. In order to utilize the spatially detailed spectral data from the red and NIR1 bands, data from the other bands were resampled from 500 m by 500 m to 250 m by 250 m pixel size. In addition, quantitative soil texture data from laboratory analysis at 51 field sites were used. Each site (pit) is about 1.5 m long, 1 m wide and 1.5 m deep. They were labeled using the identification number (ID) from 1 to 51 (Fig. 1). The accuracy of their spatial locations is about +/−10 m. A-horizon percentage of sand (2–0.05 mm size fraction, % by weight) and percentage of clay (<0.002 mm size fraction, % by weight) were determined in the laboratory using the conventional sieve-pipette method (Gee and Bauder, 1986; Sheldrick and Wang, 1993). The data were provided by the Manitoba Land Resource Unit, Agriculture and Agri-Food Canada.

2.2. Method overview

Fig. 2 illustrates the concept underlying use of environmental covariates derived from land surface dynamic feedbacks to map soil texture. There are four main steps. First, dynamic feedbacks of the land surface captured daily by MODIS imagery over a short period after a major rain event were collected to construct environmental covariates. Second, the covariates were used to generate environmental classes and their fuzzy membership distributions using a fuzzy clustering method. Third, typical soil texture values of the generated environmental classes were obtained through a spatial overlay between the membership distributions and the field soil sample sites. Finally, soil texture was predicted based on the membership distributions and typical soil texture values of the environmental classes.

2.3. Construction of environmental covariates from land surface dynamic feedbacks

2.3.1. Obtaining land surface dynamic feedbacks

Three requirements are needed to effectively extract feedbacks. First, there needs to be long period (over 1 month) of little or no rain prior to a rainfall event over the study area so that the land surface is dry. Second, the magnitude of rainfall should be large enough to force the land surface to produce a clear response. Third, no significant rainfall should occur in the area over the seven days or so following the rainfall event (referred to as the observation period). According to daily meteorological observations of the area, there was a major rainfall event (25.2 mm) between May 8 and 11, 2002. Before that the area had experienced a dry period (over 30 days) with little precipitation. Furthermore, the period which preceded the rainfall event was characterized by dramatically increasing air temperature. As a result, the land surface was in a dry state before the rainfall event. In addition, soils over much of the area were

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007

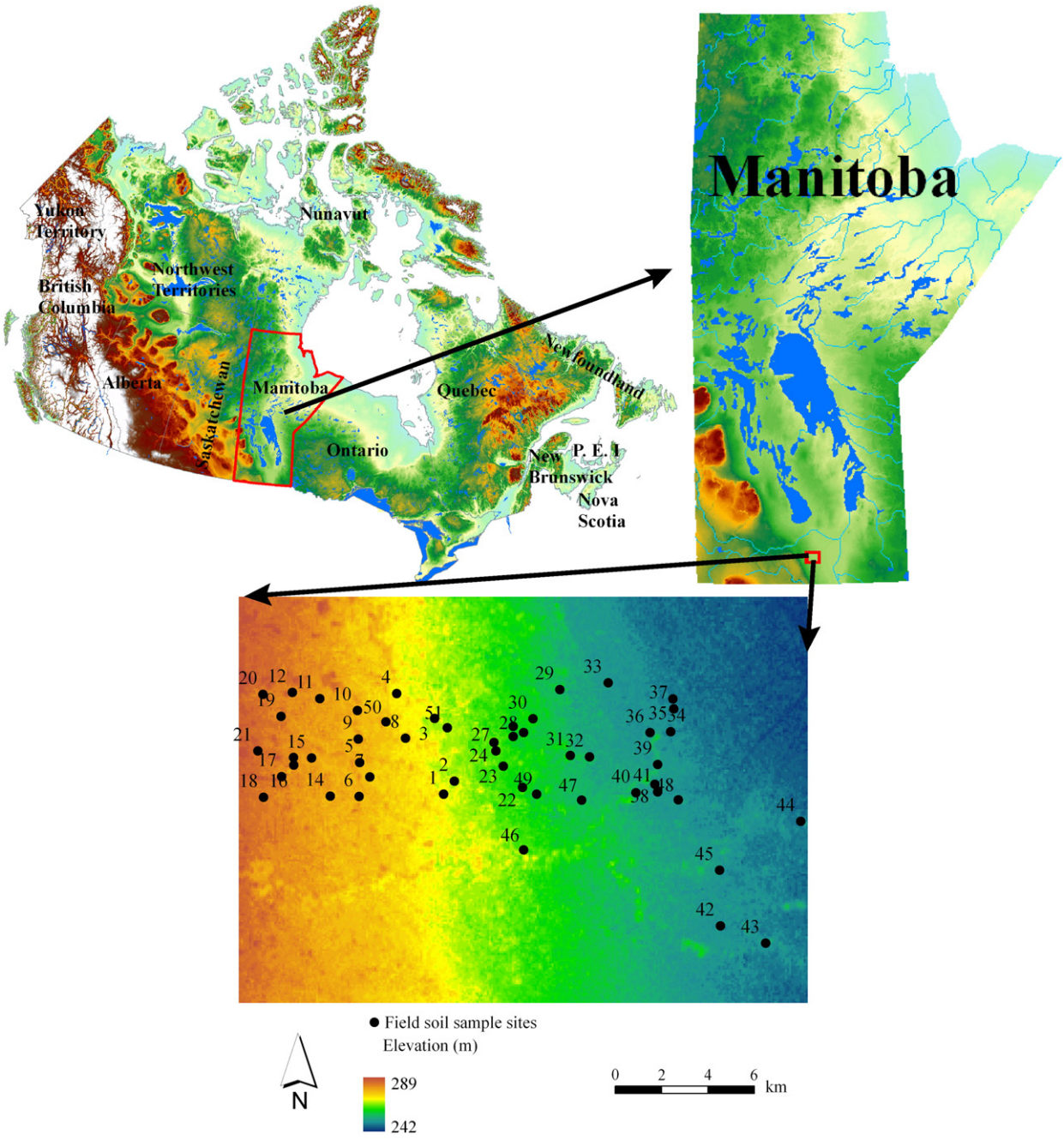

Fig. 1. Location of the study area and field soil sample sites with a DEM (digital elevation model) background.

exposed as crops had not yet emerged. There was only sparse vegetation that occurred in the northwestern portion of the area. The days from May 12–21 (Days 132–141 of the year (DOY)) were chosen as the observation period (Fig. 3). The MODIS sensor was used to capture land surface dynamic feedbacks over this period. The feedback signatures were represented as surface reflectance data which formed a time series of MODIS images 133, 134, 138, 139, 140 and 141.

2.3.2. Organization and representation of land surface dynamic feedbacks

For each pixel in the study area, the captured feedbacks contain responses from seven spectral bands over six dates. They were organized in a "day-band-spectral response" format and expressed as a surface, referred to as the "spectral–temporal response surface" (Fig. 4). The two horizontal axes represent wavelength and time respectively, and the vertical axis represents the spectral response. The spectral response can be the digital number (DN), surface reflectance, or surface radiance. Surface reflectance was used in this study. The time corresponds to a series of dates over the observation period and the wavelength corresponds to a series of bands. In this way, a spectral–temporal response surface was generated for every pixel in the study area.

2.3.3. Construction of environmental covariates

We adopted 2-D discrete wavelet analysis to create a set of environmental covariates from spectral–temporal response surfaces (Fig. 5). Daubechies's wavelet with two vanishing moments (Nason, 2008) was used in the analysis. Due to the limited number of pixels on every spectral–temporal response surface, even the edges of the

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007

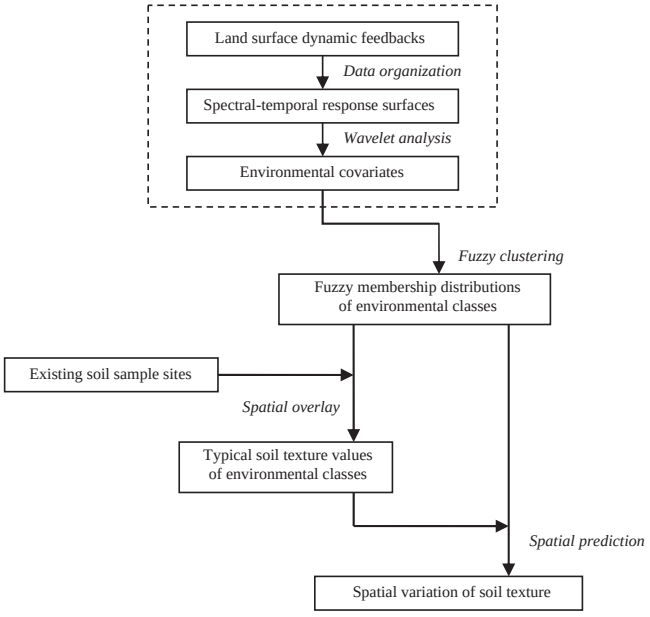

Fig. 2. Basic idea for using environmental covariates derived from land surface dynamic feedbacks to map soil texture over low relief areas.

response surface contain important response information. However, near the edges of spectral–temporal response surface, it is difficult to calculate wavelet transformation coefficients because wavelet filters usually overlap near the edges. This is the edge effect in 2-D wavelet analysis. To deal with this problem and preserve local spatial structure of response surface at the same time, every response surface was symmetrically mirrored to create a larger surface (Lark and Webster, 2004). Then, single-level discrete 2-D wavelet decomposition on every extended response surface was performed to obtain approximation coefficients matrix "*cA*", horizontal details coefficients matrix "*cH*", vertical details coefficients matrix "*cV*", and diagonal details coefficients matrix "*cD*" (Misiti et al., 2009). The values of mean and standard deviation of the coefficients matrices were used to represent the structure features of a response surface (Lee and Lou, 2003; Montoya Zegarra et al., 2008; Wu et al., 2000). As a result, for every pixel in the study area, we extracted a feature vector {*cAmean*, *cAstd*, *cHmean*, *cHstd*, *cVmean*, *cVstd*, *cDmean*, *cDstd*} from its spectral–temporal response surface. For the whole of the study area, the elements of feature vectors of all pixels formed a set of environmental covariates characterizing spatial variation of the land surface dynamic feedbacks.

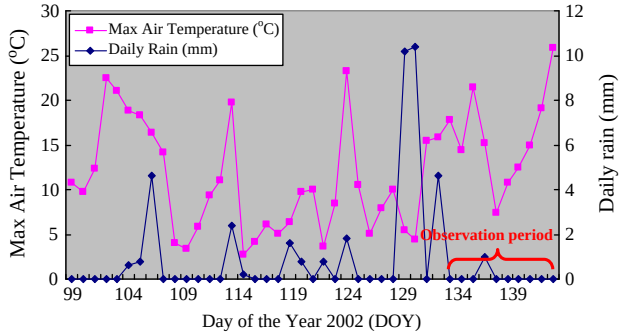

Fig. 3. Observation period for capturing land surface dynamic feedbacks.

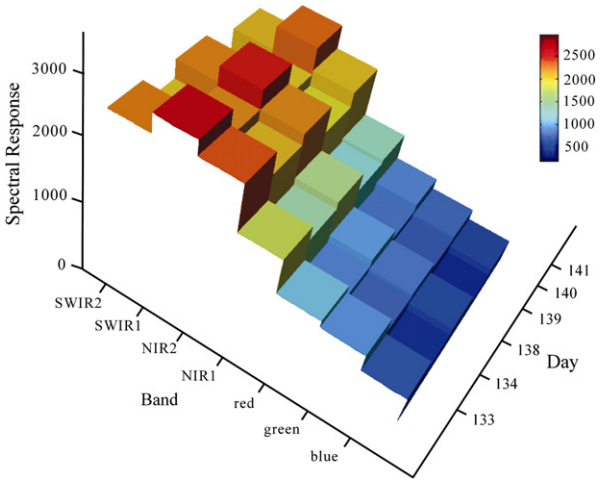

Fig. 4. Spectral–temporal response surface of a pixel in the study area.

2.4. Derivation of environmental classes using fuzzy c-means clustering

In addition to the control of soil conditions, land surface dynamic feedbacks are influenced by landform and vegetation. To reveal soil variation through the analysis of spatial differences in the dynamic feedbacks among geographic locations, it is necessary to control all other major influencing factors other than soil conditions. Thus, in this research, the entire study area was stratified into four landform-vegetation units based on elevation and NDVI (Fig. 6). The elevation data were from a SRTM (Shuttle Radar Topography Mission) DEM at 90 m resolution. NDVI data were calculated from surface reflectance values of the red and NIR1 bands of the MODIS data acquired on May 12, 2002. Areas with NDVI less than 0.3 were considered bare land or contained very sparse vegetation coverage. Among the units, unit 1 represented the areas with elevation less than 260 m and NDVI less than 0.3; unit 2 corresponded to areas with elevation greater than 260 m and NDVI less than 0.3; unit 3 was the areas with elevation less than 260 m and NDVI greater than 0.3; unit 4 was the areas with elevation greater than 260 m and NDVI greater than 0.3. The stratification resulted in relatively

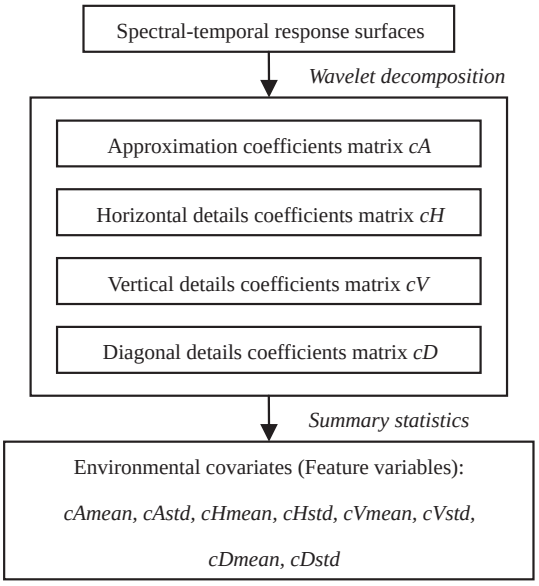

Fig. 5. Derivation of a set of environmental covariates using wavelet analysis.

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007

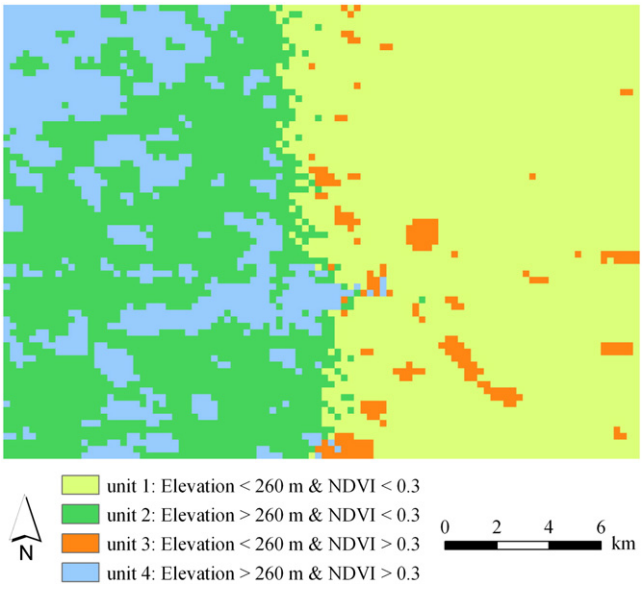

Fig. 6. Stratifying the study area into different landform-vegetation units.

homogeneous landform-vegetation units. Within any given unit, the differences in dynamic feedbacks among locations can be primarily attributed to differences in soil conditions.

For each unit, a fuzzy c-means clustering (FCM) was performed on the environmental covariates to identify different environmental classes (Zhu et al., 2008). FCM is an unsupervised classification method which optimally partitions a dataset into a given set of classes and then computes the fuzzy membership of each data element in each class (Bezdek et al., 1984). The resulting environmental classes represented different environmental combinations based on the constructed environmental covariates. Before the FCM operations were performed, data standardization was performed for each unit to stretch every variable to a range between 0 and 100. The cluster number and fuzzy weighting exponent were two key user assigned parameters in the FCM algorithm. Two cluster validity measurements including partition coefficient and normalized entropy were employed to determine the optimal values for these parameters (Bezdek et al., 1984; English, 2001). In this study, the optimal value of the fuzzy weighting exponent was considered to be 1.5. The optimal cluster numbers were determined to be 6, 4, 1 and 3 for unit 1, 2, 3 and 4, respectively. Based on these parameter values, the fuzzy membership distributions of environmental classes within each landform-vegetation unit were derived.

2.5. Identification of typical soil texture values of environmental classes

An effective way of obtaining representative soil texture values for every derived environmental class is to carry out field sampling at locations that represent the environmental class well. The optimal or best typical locations can be identified using the fuzzy membership distribution of the environmental class (Zhu et al., 2008, 2010b). In general, pixels with a high fuzzy membership value are regarded as good examples of typical locations. However, in this study area, field soil sample data already existed for both typical and less-typical sites. The expense of new field sampling could be avoided if we could make good use of the existing samples to obtain typical soil texture values for the derived environmental classes. Based on this idea, we used the fuzzy membership distributions of the derived environmental classes to identify the most typical sites from the available sample set.

For a given environmental class, its fuzzy membership distribution was spatially overlain with the soil sample sites. Sites with high membership values were considered most likely to be typical. In this research, sites with a membership value greater than 0.8 were regarded as typical. For those situations where the membership values of all the pixels of the environmental class were less than 0.8, the sites with membership values greater than 0.7 were chosen. In total, 17 typical sites were recognized. Table 1 lists these typical sites and their membership values to the associated environmental classes.

In cases where only one site was identified as typical for a particular environmental class, the soil texture values (A-horizon percentage of sand and percentage of clay) at that site were taken as the typical values of the environmental class. In those cases where more than one typical site were identified, either the average of soil texture at all sites identified as typical or the soil texture values at the single site with the highest membership were considered as the typical values of the environmental class. The former option was adopted if several sites identified as typical had similar membership values. The latter option was used if one site had much higher membership than the others. In this way, we identified typical soil sites for all the environmental classes and then obtained typical soil texture values of the classes in the study area (Table 1).

2.6. Mapping spatial variation of soil texture through a linearly weighted averaging function

Environmental conditions are characterized by the derived covariates. It was assumed that the more the local environmental conditions resemble the typical condition of an environmental class, the more the local soil conditions were similar to the typical soil conditions of the environmental class. Based on this assumption, for each landform-vegetation unit, we used a linearly weighted averaging function (Zhu

Table 1
Typical soil sample sites and their A-horizon soil texture (percentage of sand and percentage of clay) for each environmental class within landform-vegetation units.

Landform-vegetation unit	The derived environmental Class	Typical soil sites: ID (membership value)	Percentage of sand (% by weight)	percentage of clay (% by weight)
1	1	33 (0.85), 34 (0.87)	8	56
	2	30 (0.82)	69	19
	3	38 (0.87)	7	38
	4	47 (0.79)	13	51
	5	22 (0.86)	85	14
	6	31 (0.84)	52	24
2	1	5 (0.82), 24 (0.81)	85	9
	2	2 (0.80)	72	18
	3	13 (0.81)	81	12
	4	3 (0.85)	91	6
3	1	26 (0.77), 44 (0.75)	52	25
4	1	1 (0.83)	94	4
	2	11 (0.85)	84	9
	3	46 (0.82)	70	19

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007

et al., 1997) to map spatial variation of A-horizon soil texture (percentage of sand and percentage of clay). The texture value at each pixel within any given unit was computed as the weighted average of typical texture values of the environmental classes belonging to this unit and fuzzy membership values of the pixel to these environmental classes. The function can be described as follows:

$$V_{ij} = \frac{\sum_{K=1}^{n} \left(S_{ij}^{K} \cdot V^{K}\right)}{\sum_{K=1}^{n} S_{ij}^{K}} \quad (1)$$

Where V_{ij} is the predicted value of A-horizon soil texture (percentage of sand or percentage of clay) at pixel (i, j) within any given landform-vegetation unit; V^k is typical value of A-horizon soil texture (percentage of sand or percentage of clay) of environmental class k; S_{ij}^k is fuzzy membership value of environmental class k at pixel (i, j); n is the number of environmental classes within the unit.

2.7. Assessment of prediction results

The prediction results of A-horizon soil texture were assessed through comparison with A-horizon soil texture maps derived from a 1:20,000 traditional soil survey. The assessment was based on 34 field sample sites, independent of the samples used for prediction. Two assessment measures were used. One was the mean absolute error (MAE) between the observed and the predicted values, and the other was the root mean square error (RMSE). Both MAE and RMSE measure the average magnitude of the errors between the observed and the predicted values. But MAE is less sensitive to large errors than RMSE. MAE was estimated by:

$$MAE = \sum_{k=1}^{n} \frac{|P_i - O_i|}{n} \quad (2)$$

where P_i is the predicted value of A-horizon percentage of sand (percentage of clay) at validation site i, O_i is the observed value of A-horizon percentage of sand (percentage of clay) at validation site i, n is the number of the validation sites. The RMSE was estimated by:

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

The value of MAE (or RMSE) ranges from 0 to ∞. The lower the values of MAE and RMSE the better the performance of the prediction.

2.8. Assessment of MODIS derived variables

To further examine the effectiveness of the MODIS derived environmental variables for mapping soil texture, we conducted the following comparative analyses. First, for each landform-vegetation unit, we used the data in Table 1 to calculate an average of the typical soil texture over the environmental classes within the unit. The average was considered as a 'typical' composition for the unit. With the same approach proposed in this paper, the 'typical' compositions shown in Table 2 were used to predict soil texture variation for the units. Based on the same 34 field sample sites, we compared the errors of this prediction with those of the prediction using fourteen separate environmental classes listed in Table 1.

Table 2
The averages of typical A-horizon soil texture (percentage of sand and percentage of clay) over the environmental classes within each landform-vegetation unit.

Landform-vegetation unit	Average of typical soil texture over the classes	
	Percentage of sand (% by weight)	Percentage of clay (% by weight)
1	39	34
2	82	11
3	52	25
4	83	11

Second, linear regression analyses of A-horizon soil texture on the units and the MODIS derived variables were performed based on all 51 field samples. It was assumed that the response of soil texture to the units and the derived variables was linear and static over space. Before the regression analyses, special treatments were conducted on the data of the soil texture and the units. The percentages of sand, clay and silt in A-horizon soil are compositional data with a constant sum constraint (100%). Due to the compositional constraints (Aitchison, 1986; Odeh et al., 2003), the additive log-ratio (alr) transformation described by Lark and Bishop (2007) was used to transform the compositional data before analysis. We computed the alr transformation of the percentages of sand and clay, with the percentage of silt as the denominator of the ratio. In addition, the attribute of the units is discrete and qualitative. To facilitate mathematical considerations, it was represented by dummy variables which take the value of either 0 or 1 (Greene, 2003). Three dummy variables (*D1*, *D2* and *D3*) were constructed for the four units. Specifically, *D1* represented the unit 1, which was equal to 1 if a field sample was collected from this unit or 0 if it was from other units. With the same method, *D2* and *D3* were also obtained for the units 2 and 3 respectively. If *D1*, *D2* and *D3* were all equal to 0, the field sample was collected from the unit 4; otherwise, it was from other units. In the following regression analysis, the three dummy variables and other eight MODIS derived variables (*cAmean*, *cAstd*, *cHmean*, *cHstd*, *cVstd*, and *cDstd*) were considered as explanatory variables for the alr-transformed sand and clay variables (*alr(Sand)* and *alr(Clay)*) respectively. It should be noted that the alr-transformed values can be back-transformed and then re-expressed as percentages based on the techniques provided by Lark and Bishop (2007).

The regression analyses include three steps. To clearly describe the procedure, we took the alr-transformed sand for example. In the first step, a multiple linear regression (Model 1) was used to model the relationships between the alr-transformed sand and the three dummy variables. It can be formulated as follows:

$$alr(Sand) = a_0 + a_1 D1 + a_2 D2 + a_3 D3 + \varepsilon_1 \quad (4)$$

where $alr(Sand)$ is the alr-transformed variable for the percentage of sand, a_0 is a constant term, a_1, a_2 and a_3 are partial regression coefficients, ε_1 is a random error term. In the second step, the eight MODIS variables were added to the list of explanatory variables. Multiple linear regression analysis (Model 2) was used again to model the relationships between the alr-transformed sand and the dummy variables and the MODIS variables. It can be described as follows:

$$\begin{aligned} alr(Sand) = {} & b_0 + b_1 D1 + b_2 D2 + b_3 D3 + b_4 cAmean + b_5 cAstd \\ & + b_6 cHmean + b_7 cHstd + b_8 cVmean + b_9 cVstd \\ & + b_{10} cDmean + b_{11} cDstd + \varepsilon_2 \end{aligned} \quad (5)$$

where b_0 is a constant term, b_1, b_2, b_3, …, b_{11} are partial regression coefficients, ε_2 is a random error term. Finally, we computed the difference in the residual sum of squares for the two models, and obtain a mean square through dividing this difference by the associated degree of freedom. This degree of freedom is equal to the difference in the residual degrees of freedom for the two models. Then, *F*-statistic was constructed through dividing this mean square by the residual mean square for the Model 2. The *F*-test was used to further show the effectiveness of the MODIS derived variables at differentiating patterns of soil texture.

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007

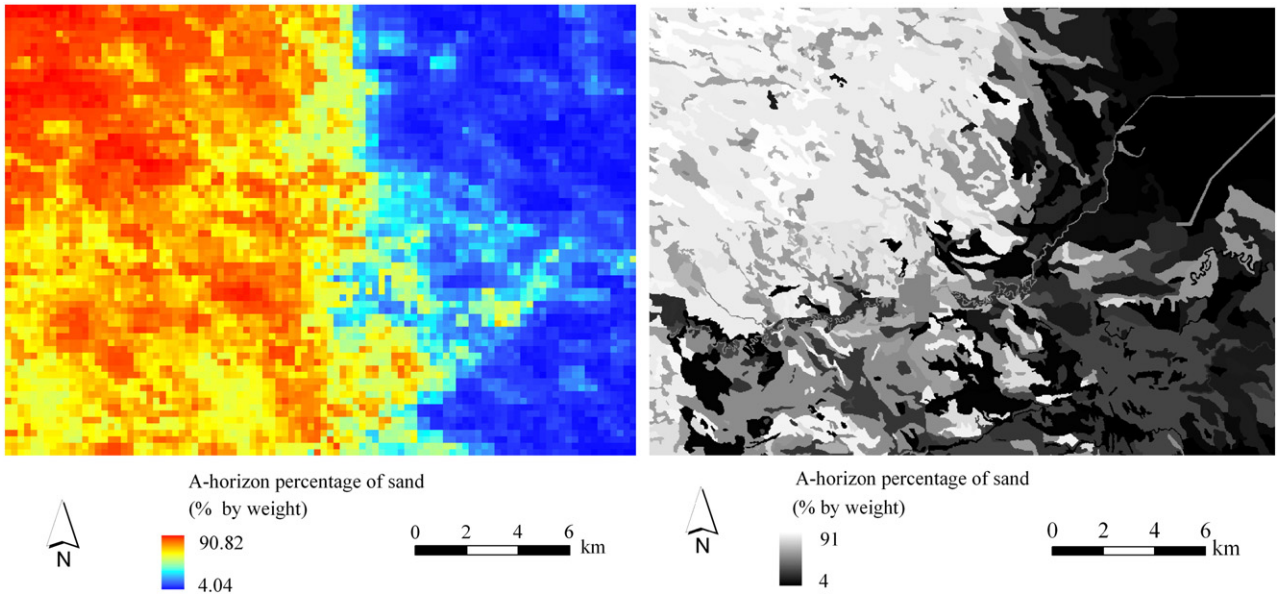

Fig. 7. (left) Predicted map of A-horizon percentage of sand; (right) map of percentage of sand derived from 1:20,000 traditional soil survey.

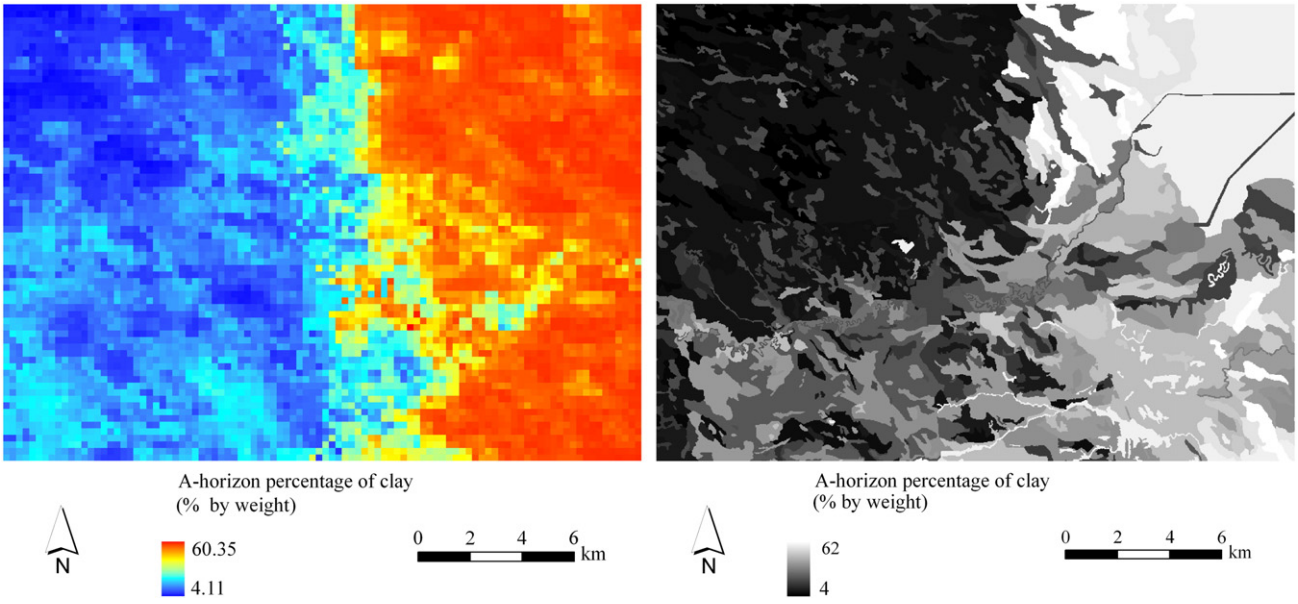

Fig. 8. (left) Predicted map of A-horizon percentage of clay; (right) map of percentage of clay derived from 1:20,000 traditional soil survey.

3. Results and discussion

Figs. 7 (left) and 8 (left) are the maps of the predicted A-horizon soil texture. It can be seen that there are obvious differences in A-horizon soil texture between the western and the eastern parts of the study area. The western part has higher percentage of sand and lower percentage of clay while the eastern part shows the opposite pattern. Specifically, in the northwestern portion, sandy soils are predominant with percentage of sand greater than 80 and percentage of clay less than 15. In the northeastern and southeastern portions, clay soils are prevalent with percentage of clay greater than 50. In other portions, mixed distributions of the soils with different percentage of sand and percentage of clay occur.

For comparison with the prediction results, maps of A-horizon soil texture derived from 1:20,000 traditional soil survey are also shown in Figs. 7 (right) and 8 (right). Based on the 34 field sample sites, the MAE and RMSE were calculated for both maps. Table 3 lists the MAE and RMSE values.

For the A-horizon percentage of sand, the error values (9.42 for MAE and 12.56 for RMSE) of the results from the dynamic feedback

Table 3
Accuracy assessments of A-horizon soil texture maps from the proposed approach and from 1:20,000 traditional soil survey, respectively.

Quantitative measures	The proposed approach		Detailed traditional soil survey	
	Percentage of sand	Percentage of clay	Percentage of sand	Percentage of clay
MAE	9.42	7.33	10.59	7.5
RMSE	12.56	9.08	15.12	12.08

RMSE: root mean square error; MAE: mean absolute error.
Number of validation sites: 34.

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007

approach are less than those (10.59 for MAE and 15.12 for RMSE) from the detailed traditional soil survey. Similarly, for the A-horizon percentage of clay, the error values (7.33 for MAE and 9.08 for RMSE) of the results from the proposed approach are less than those (7.5 for MAE and 12.08 for RMSE) from the detailed traditional soil survey. This indicates that the prediction results are at least as good as, if not more accurate than, the soil texture maps derived from the detailed traditional soil survey.

It is observed that the major patterns displayed by the prediction results are similar to the soil texture maps derived from the detailed traditional soil survey. The obvious difference is that spatial variation of soil texture from the former is more continuous than the latter. This is due to limitations of traditional soil survey map formats based on discrete polygons, which have been widely discussed by the digital soil mapping community (McBratney et al., 1992; Zhu et al., 1997).

On the other hand, our comparative analyses for further assessment of the MODIS derived variables show that the error values of the prediction using the 'typical' compositions shown in Table 2 are 13.38 (MAE) and 17.64 (RMSE) for percentage of sand, and 8.09 (MAE) and 11.83 (RMSE) for percentage of clay. These values are remarkably greater than those (9.42 (MAE) and 12.56 (RMSE) for percentage of sand, 7.33 (MAE) and 9.08 (RMSE) for percentage of clay) of the prediction using the fourteen separate classes. The remarkable difference is mainly because the fourteen classes derived from the MODIS variables effectively captured the soil texture variation within each unit. The comparative analyses also show that the adjusted R^2 value of Model 1 is 0.33 while that of Model 2 is 0.747. Compared to Model 1, Model 2 can account for much more variation of the sand contents at the 51 sites. This means that the MODIS variables notably increase the proportion (from 33% to 74.7%) which can be explained by regression modeling. Besides, Table 4 lists the results of the analysis of variance for the two regression models. The values of the residual sum of squares and the associated degree of freedom for Model 1 are 91.302 and 47, respectively, while for Model 2 they are 28.663 and 39 respectively. The residual mean square for Model 2 is 0.735. Based on these values, a mean square $((91.302-28.663)/(47-39)=7.83)$ and thus the constructed *F*-statistic $(7.83/0.735=10.653)$ were obtained using the method described in the *2.8* section. According to tables of critical values of *F* distribution (Bernstein and Bernstein, 1999), the value of *F* distribution with the first degree of freedom 8 and the second degree of freedom 39 at the significance level of 0.005 is about 3.35. Obviously, the $F_{0.005}(8, 39)$ is much less than our obtained *F*-value. This indicates that there is a significant improvement from Model 1 to Model 2. The improvement should be attributed to the addition of the MODIS variables to the regression modeling. Furthermore, Table 5 lists the estimates of the regression coefficients for the two models. In Model 1, only the dummy variable *D1* is significant at the level of about 0.005 (two-tailed). In Model 2, all the dummy variables are not significant, and all the MODIS derived variables except *cVmean*, *cVstd* and *cDstd* are significant at the level of about 0.01 or 0.05 (two-tailed). Similar results were obtained for the alr-transformed clay. Therefore, the MODIS derived variables are capable of characterizing most of variation of soil texture over the study area.

Table 4
Analysis of variance for the two regression models (Model 1 and Model 2).

Regression models		Sum of squares	Degree of freedom	Mean square	F-values	p-values
Model 1	Regression	53.73	3	17.91	9.22	0.000
	Residual	91.302	47	1.943		
	Total	145.032	50			
Model 2	Regression	116.369	11	10.579	14.394	0.000
	Residual	28.663	39	0.735		
	Total	145.032	50			

Dependent variable of both Model 1 and 2: *alr(Sand)*.
Explanatory variables of Model 1: *D1*, *D2* and *D3*, which represent the landform-vegetation strata.
Explanatory variables of Model 2: *D1*, *D2*, *D3*, *cAmean*, *cAstd*, *cHmean*, *cHstd*, *cVstd*, and *cDstd*; the new added variables were derived from time series MODIS data.

Table 5
Regression coefficient estimates for the two regression models (Model 1 and Model 2).

Regression models	Variables	Coefficients	Standard errors	t-values	p-values
Model 1	*(constant)*	1.700	0.387	4.398	0.000
	D1	−1.567	0.502	−3.123	0.003
	D2	0.874	0.520	1.680	0.100
	D3	−0.191	0.893	−0.214	0.832
Model 2	*(constant)*	−8.179	1.636	−5.001	0.000
	D1	0.172	0.465	0.369	0.714
	D2	0.335	0.413	0.812	0.421
	D3	0.235	0.571	0.412	0.683
	cAmean	0.001	0.001	1.937	0.059
	cAstd	0.002	0.001	2.528	0.016
	cHmean	−0.034	0.015	−2.234	0.031
	cHstd	0.013	0.005	2.648	0.012
	cVmean	−0.052	0.421	−0.124	0.902
	cVstd	0.000	0.002	−0.351	0.728
	cDmean	−0.054	0.023	−2.284	0.028
	cDstd	−0.016	0.012	−1.336	0.189

Dependent variable of both Model 1 and 2: *alr(Sand)*.
Explanatory variables *D1*, *D2* and *D3* are dummy variables, which represent the landform-vegetation strata.
Explanatory variables *cAmean*, *cAstd*, *cHmean*, *cHstd*, *cVstd*, and *cDstd* were derived from time series MODIS data.

It should be acknowledged that stratifying the study area into relatively homogeneous landform-vegetation units was important to the success of the soil texture prediction over the study area using the proposed approach. In addition to the control of soil conditions, land surface dynamic feedbacks are influenced by landform and vegetation. However, within any given landform-vegetation strata, the differences in dynamic feedbacks among locations can be primarily attributed to the differences in soil conditions within that strata. This stratification enables the FCM techniques to produce classes of dynamic feedback patterns that are primarily controlled by soil conditions. Further, this makes it possible to identify the most typical soil sample sites based on the fuzzy membership distributions of the produced feedback patterns. Stratification is not always obligatory, for application of the dynamic feedback approach but it is likely to be necessary for application to any areas that display significant differences in landform and vegetation conditions. The difference between the results of the linear regression analysis without and with the MODIS derived variables provides another proof that the MODIS variables and the described land surface feedback techniques are effective at differentiating patterns of soil texture. Therefore, it can be recognized that the variables, extracted from analysis of land surface dynamic feedbacks, made major contributions to the success of the soil texture prediction.

4. Conclusions

In low relief areas, easily obtained landform and vegetation environmental covariates often do not co-vary with soil conditions over space to the level that they can be effectively used in digital soil mapping. This paper presents an approach to predict the variation of soil texture over such areas using environmental covariates derived from land surface dynamic feedbacks extracted from MODIS data. A case study in a low relief area located in the south-central Manitoba was carried out to demonstrate this approach. It can be concluded from the study that the developed environmental covariates have the ability to reveal soil texture variation. The proposed approach can serve as an effective solution for mapping soil texture and possibly other properties, over similar areas. It should be emphasized that the 250 m resolution of the MODIS data and the 90 m resolution of the SRTM DEM data impose a lower limit on the size of areas for which

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007

differences in texture can be estimated. The same concept could be applied using data sets of finer spatial resolution however these are frequently not widely available. The widespread availability of MODIS and SRTM data sets ensures that the technique described here is likely to be applicable for large portions of the world.

Extensive portions of Canada and other countries consist of low relief plains with topographically gentle undulation where the observable association between landform/vegetation conditions and soil conditions is similar to our study area. The proposed approach has the potential to play an important role in digital soil mapping over such areas.

In this study, the development of environmental covariates was based on land surface dynamic feedbacks collected by remote sensing techniques after a major rainfall event. The selection of the rainfall event and observation period is very important for the efficiency of the feedbacks in indicating soil spatial variation. Also, the FCM method in this study was effective for identifying environmental classes. But, it should be realized that the selection of user assigned algorithm parameters is crucial for obtaining optimal clustering results.

It should be pointed out that this work only examines one way to use the developed environmental covariates to map soil variation. In actuality, like commonly used climate and terrain variables, the developed covariates can be utilized by other digital soil mapping techniques (for example, Soil Land Inference Model (SoLIM)) in similar ways over low relief areas.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (Project Numbers: 40971236 and 41023010), the National Basic Research Program of China (2007CB407207); Program of International S&T Cooperation, the Ministry of Science & Technology of China (Project No.: 2010DFB24140), the 2008 MOE-AAFC PhD Research Program of the China Scholarship Council and Agriculture & Agri-Food Canada; Government Related Interest Program (GRIP) of Agriculture and Agri-Food Canada; the Canadian Space Agency; the Vilas Associate Program and the Hamel Faculty Fellow Program at the University of Wisconsin–Madison. The senior author would like to thank all members of Canadian Soil Information System (CanSIS) at Agriculture & Agri-Food Canada for their friendly and generous support and help. We also thank Dr. Bert VandenBygaart and Dave Howlett for their language editing efforts on the paper. We benefitted greatly from the constructive comments from the anonymous reviewers and the guest editor, Dr. Murray Lark, particularly in relation to the regression analysis in this paper.

References

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman and Hall, London.
Bernstein, S., Bernstein, R., 1999. Elements of Statistics II: Inferential Statistics. McGraw-Hill Professional, New York.
Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: the fuzzy c-means clustering algorithm. Computers and Geosciences 10, 191–203.
Clark, R.N., Swayze, G.A., 1996. Evolution in imaging spectroscopy analysis and sensor signal-to-noise: an examination of how far we have come. Summaries of the sixth annual JPL airborne Earth science workshop. 4–8 March 1996: AVIRIS Workshop, vol. 1, p. 5.
Coleman, T.L., Agbu, P.A., Montgomery, O.L., 1993. Spectral differentiation of soils and soil properties: is it possible from space platforms? Soil Science 155, 283–293.
Demattê, J.A.M., Fiorio, P.R., Ben-Dor, E., 2009. Estimation of soil properties by orbital and laboratory reflectance means and its relation with soil classification. The Open Remote Sensing Journal 2, 12–23.
Ding, Y., Xu, S., Zhu, K., 1989. Application of remote sensing techniques on 1: 500,000 soil mapping in Nanjing, Jiangsu Province, China. Soils 6, 304–306 (In Chinese).
English, E.M., 2001. Assisting knowledge-based inferential soil mapping: The application of Fuzzy c-Means Clustering to Expose Environmental Niches. Master Thesis. University of Wisconsin-Madison.
Environment Canada, 1993. Canadian climatic normals 1961–1990. Prairie Provinces. Atmospheric Environment, Downsview, Ontario.
Gee, G.W., Bauder, J.W., 1986. Particle-size analysis, In: Klute, A. (Ed.), Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods, 2nd edn.: Agronomy, 9. ASA, Madison, Wisconsin, pp. 383–411.
Gomez, C., Lagacherie, P., Coulouma, G., 2008. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. Geoderma 148, 141–148.
Greene, W.H., 2003. Econometric Analysis, Fifth edition. Pearson Education, Inc., Upper Saddle River, New Jersey, p. 07458.
Jenny, H., 1941. Factors of Soil Formation. McGraw-Hill, New York.
Lagacherie, P., Baret, F., Feret, J.B., Netto, J.M., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, filed and airborne hyperspectral measurements. Remote Sensing of Environment 112, 825–835.
Land Resource Unit, 1999. Soils and terrain. An introduction to the land resource. Rural Municipality of Dufferin. Information Bulletin 97–20 (Revised), Brandon Research Centre, Research Branch, Agriculture and Agri-Food Canada.
Lark, R.M., Bishop, T.F.A., 2007. Cokriging particle size fractions of the soil. European Journal of Soil Science 58, 763–774.
Lark, R.M., Webster, R., 2004. Analyzing soil variation in two dimensions with the discrete wavelet transform. European Journal of Soil Science 55, 777–797.
Lee, J.D., Lou, L.P., 2003. Using texture and shape features to retrieve sets of similar medical images. Biomedical Engineering Applications, Basis & Communications 15, 193–199.
McBratney, A.B., De Gruijter, J.J., Brus, D.J., 1992. Spatial prediction and mapping of continuous soil classes. Geoderma 54, 39–64.
McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma 89, 67–94.
Michalyna, W., Podolsky, G., Jacques, E.S.T., 1988. Canada–Manitoba Soil Survey: Soils of the Rural Municipalities of Grey, Dufferin, Roland, Thompson and Part of Stanley (Report D60). Department of Soil Science, University of Manitoba.
Misiti, M., Misiti, Y., Oppenheim, G., Poggi, J., 2009. User's Guide of Wavelet Toolbox™ 4. http://www.mathworks.com/access/helpdesk/help/2009.
Montoya Zegarra, J.A., Leite, N.J., Silva Torres, R., 2008. Wavelet-based fingerprint image retrieval. Journal of Computational and Applied Mathematics. doi:10.1016/j.cam.2008.03.017.
Nason, G.P., 2008. Wavelet Methods in Statistics with R. Springer, New York, pp. 37–41.
Odeh, I.O.A., McBratney, A.B., 2000. Using AVHRR images for spatial prediction of clay content in the lower Namoi Valley of eastern Australia. Geoderma 97, 237–254.
Odeh, I.O.A., Todd, A.J., Triantafilis, J., 2003. Spatial prediction of soil particle-size fractions as compositional data. Soil Science 168, 501–515.
Selige, T., Böhner, J., Schmidhalter, U., 2006. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. Geoderma 136, 235–244.
Sheldrick, B.H., Wang, C., 1993. Particle size distribution. In: Carter, M.R. (Ed.), Soil Sampling and Methods of Analysis: Canadian Society of Soil Science. Lewis Publishers, Ann Arbor.
Soil Classification Working Group, 1998. The Canadian System of Soil Classification, Third edition. Agriculture and Agri-Food Canada Publication 1646.
Sullivan, D.G., Shaw, J.N., Rickman, D., 2005. IKONOS imagery to estimate surface soil property variability in two Alabama physiographies. Soil Science Society of America Journal 69, 1789–1798.
Wu, P., Manjunath, B.S., Newsam, S., Shin, H.D., 2000. A texture descriptor for browsing and similarity retrieval. Signal Processing: Image Communication 16, 33–43.
Zhu, A.X., Band, L., Vertessy, R., Dutton, B., 1997. Derivation of soil properties using a Soil Land Inference Model (SoLIM). Soil Science Society of America Journal 61, 523–533.
Zhu, A.X., Yang, L., Li, B.L., Qin, C.Z., English, E., Burt, J.E., Zhou, C.H., 2008. Purposefully sampling for digital soil mapping. In: Hartemink, A.E., McBratney, A.B., Mendonca Santos, M.L. (Eds.), Digital Soil Mapping with Limited Data. Springer-Verlag, New York, pp. 233–245.
Zhu, A.X., Liu, F., Li, B.L., Pei, T., Qin, C.Z., Liu, G.H., Wang, Y.J., Chen, Y.N., Ma, X.W., Qi, F., Zhou, C.H., 2010a. Differentiation of soil conditions over low relief areas using feedback dynamic patterns. Soil Science Society of America Journal 74, 861–869.
Zhu, A.X., Yang, L., Li, B.L., Qin, C.Z., Pei, T., Liu, B.Y., 2010b. Construction of membership functions for predictive soil mapping under fuzzy logic. Geoderma 155, 164–174.

Please cite this article as: Liu, F., et al., Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS, Geoderma (2011), doi:10.1016/j.geoderma.2011.05.007