Geoderma 316 (2018) 89–99

Contents lists available at ScienceDirect

Geoderma

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

Methods for sandy land detection based on multispectral remote sensing data

Junjun Wu[a], Zhihai Gao[b], Qinhuo Liu[a], Zengyuan Li[b,*], Bo Zhong[a]

[a] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100101, China
[b] Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China

ARTICLE INFO

Handling Editor: Yvan Capowiez

Keywords:
Sandy land detection
Mixed pixel decomposition
Soil particle composition
Remote sensing
Transitional sandy land

ABSTRACT

Sandification is becoming a serious threat to the sustainability of human habitation. The potential of remote sensing in sandy land detection has been previously demonstrated, but transitional sandy land is difficult to detect because of vegetation cover. The aim of this study, therefore, was to propose methods for sandy land detection based on mixed pixel decomposition and soil particle composition to determine the effects of vegetation coverage and transitional sandy land, using Zhenglan Banner of China as the study area and GF-1 multispectral images as the main data. Results showed that the pixel purity index (PPI) is a viable indicator for pure endmember extraction for sandy land detection via remote sensing. A linear spectral unmixing (LSU) model was established to distinguish sandy land coverage from vegetation, alkaline land, etc. Results showed that without considering the vegetation proportion, when the endmember proportion of sandy land accounted for > 50% of the total (except for the vegetation), a pixel would be detected as sandy land, its extraction accuracy was verified to be 86.42% by field data. The results derived from soil particle composition showed that silt was the best indicator for sandy land detection, and clay was secondary to it. Through Partial Least Squares Analysis (PLSA), the percentage of silt content was determined as the dependent variable; R_{676} and R_{774} were selected as independent variables to establish the inversion model according to the model effect weights and VIP value. The minimum Prediction Residual Error Sum of Squares (PRESS) was 0.824 tested by leave one out cross validation. The threshold of silt content was determined finally as 3.5%, namely, when the silt content was < 3.5%, the pixel would be classified as sandy land, its extraction accuracy was 80.86% that was verified using the same field data.

1. Introduction

Sandy land occurs in arid, semi-arid and semi-humid regions, due to climatic change and human activities (UNCCD, 2013; Reed et al., 2013), which is a global problem associated with biological diversity loss, deforestation and soil degradation (Biro et al., 2013; Gisladottir and Stocking, 2005). The growing coverage of sandy land is becoming an important issue and poses a serious threat to the sustainability of human habitation, especially in China (Li et al., 2013). Sandy land detection is sorely needed at both regional and national scales for land resource management and environmental research, but maps of its extent and severity at large scales do not exist (Prince et al., 2009). Detection is the basis for the quantitative monitoring and assessment of sandy land (Zhao et al., 2008; Sun et al., 2014; Stringer and Reed, 2007), and in this respect, remote sensing technology has played an active role (Susan et al., 2009). Detection methods have mainly included supervised/unsupervised classification, object-oriented classification, visual interpretation knowledge, neural network, decision tree, spectral mixture analysis and quantitative retrieval for soil parameters (Wu et al., 2015). Most are qualitative and subjective, but spectral mixture analysis and quantitative retrieval for soil parameters are much more quantitative and objective.

Because of the limited spatial resolution of sensors and the complex features of background, pixels generally consist of several land cover types; such pixels are referred to as mixed pixels. Mixed pixel decomposition makes it possible to acquire land information at the sub-pixel level and to detect soil type quantitatively. At present, linear and non-linear spectral unmixing models are commonly used. Linear spectral unmixing models have the advantages of simple modeling and clear physical meaning; these models can decompose mixed pixels easily and quickly (Ines and Honda, 2005). Non-linear spectral unmixing models can reflect actual features better, but the calculation process makes obtaining results difficult; moreover, the residual error is often large. Gregory's study showed that relative spectral mixing analysis (RSMA)

* Corresponding author.
E-mail address: lizy@caf.ac.cn (Z. Li).

https://doi.org/10.1016/j.geoderma.2017.12.015
Received 10 July 2017; Received in revised form 10 December 2017; Accepted 12 December 2017
Available online 19 December 2017

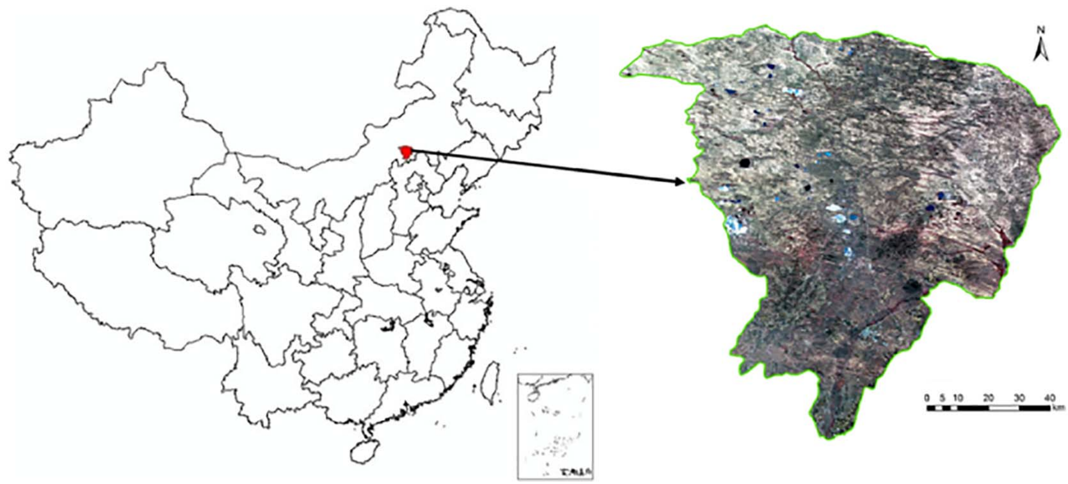

Fig. 1. Location of the study area.

could well predict the dynamic change of soil (Gregory et al., 2013). Regarding the effect of constraint conditions on the objective function, it was believed that the constraint decomposition method could overcome the local minimum problem (Liu et al., 2011) and that the constrained model was better than the unconstrained one for bare land extraction (Sun et al., 2012; Yang et al., 2008). The method for determining the number, type and selection methods of endmembers is crucial for mixed pixel decomposition based on remote sensing imagery. If the number of endmembers is not sufficiently large, not all typical land types can be captured, which would lead to misclassification; if the number is too large, the correlation between endmembers would be enhanced, which could confound the results (Zhao et al., 2011). Generally, endmembers are obtained from measured spectra and remote sensing images (Pan et al., 2010). The measured endmember has the advantage of possessing a continuous spectrum, which can be resampled by most satellite sensors; the greatest drawback, however, is that it is difficult to guarantee synchronization with remote sensing images, whereas an image endmember can better match image spectra. Scholars have developed different methods for extracting endmembers for different purposes, such as minimum volume transform, automated morphology, vertex component analysis, independent principal component analysis, iterative error analysis, blind sources separation methods, and pixel purity index (Loghmari et al., 2002; Naceur et al., 2004). Research has shown that the accuracy of these methods can be improved by using the pure pixel index to determine endmembers (Cai et al., 2009). Generally, mixed pixel decomposition has been mostly applied to hyperspectral data (Sebastian et al., 2013; Li et al., 2010; Antonio Plaza et al., 2004), whereas it is less frequently applied to multispectral data because of the limited bands of these data. In the matter of sandy land detection and monitoring based on hyperspectral data, Okin evaluated the wind erosion status of Mojave Desert in American quantitatively using the AVIRIS and MSS hyperspectral data (Okin et al., 2001); Gregory carried on a contrast research on AVIRIS and EO-1 Hyperion data, and extracted quantitatively three components information in large area, such as green vegetation, withered vegetation and bare soil, and reflected the degree of desertification quantitatively (Gregory et al., 2003). In previous studies based on hyperspectral images, the study area was often kept small because of high data costs. A few studies have shown that mixed pixel decomposition can also be applied to multispectral remote sensing data, such as Landsat ETM +, NOAA-AVHRR and Formosat-2 (Moussa et al., 2012; Wang et al., 2011), although the accuracy is slightly lower. For example, Yan extracted the snow coverage by LSU based on the NOAA-AVHRR multispectral image (Yan and Zhang, 2004), and Husain mapped the land cover by LSU analysis based on Landsat7 ETM + and ASTER multispectral data (Alanazi and Ghrefat, 2013), both of them obtained approving results. A relatively mature method for extracting vegetation information by mixed pixel decomposition has been developed, but its application in detecting soil information is not yet well developed, especially for sandy land; a new approach to researching soil condition is required. Previous studies on sandy land detection have mainly obtained the proportion of sandy land in a given area but have not been able to determine its type precisely; this problem was solved in this study.

Soil particle composition, soil organic matter, soil roughness and other soil indexes were important characteristic parameters to reflect soil essence (Carsten et al., 2011). Soil particle composition affected the soil hydraulic properties, fertility and erosion degree (Meland et al., 2012; Rawlins et al., 2011). The research showed that the silt content in soil particle composition was the best indicator to reflect the soil status of sandy land (Wu et al., 2015), so that the precise inversion of soil particle composition was a necessary prerequisite for sandy land detection. There were a lot of quantitative inversion models for soil characteristic parameters, such as correlation analysis, stepwise regression (Toure and Tychon, 2004), multiple regression (Al-abbas et al., 1972), partial least squares (Wold, 1984) and so on. The research showed that when one or more variables were used to establish the inversion model, the partial least squares model was more effective than the other normal multiple regression model, especially when the independent variables had a high correlation with each other, the advantage of partial least squares was more obvious (Geladi and Qkowlaski, 1986; Höskuldsson, 1988).

Detecting sandy land accurately by traditional classification method from remote sensing image is difficult because of transitional sandy land and high vegetation coverage. In this study, a method for sandy land detection based on mixed pixel decomposition was proposed to solve these problems, and another method based on soil particle composition was also put forward to be compared, taking Zhenglan Banner in the Inner Mongolia region of China as the study area and GF-1 multispectral remote sensing data as the main data.

2. Study area

The study area is located in Zhenglan Banner in Inner Mongolia (41°56′–43°11′N, 115°00′–116°42′E), with a total area of 10,182 km^2 (Fig. 1). The northern region of the study area is located in the hinterland of Otindag sandy land, whose surface is covered by shifting sandy land, semi-fixed sandy land and fixed sandy land. The study area is one of the largest wind areas and the main sources of sandstorms in China. The southern study area is a typical pastoral region; its topography features low mountains and hills. The average elevation is approximately 1300 m, and the climate belongs to the temperate