Advanced Materials Research Vols. 347-353 (2012) pp 3559-3563
Online available since 2011/Oct/07 at www.scientific.net

doi:10.4028/www.scientific.net/AMR.347-353.3559

Assessing the Soil Fertility using Landsat TM Imagery and Geospatial Statistical Analysis

Jinling Zhao[1, 2, a], Dacheng Wang[1, b], Dongyan Zhang[1, c], Juhua Luo[1, d] and Wenjiang Huang[1, e*]

[1]Beijing Research Center for Information Technology in Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing, 100097, P.R. China

[2]Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, 100101, P.R. China

[a]zhaojl@nercita.org.cn, [b]wdc198206@163.com, [c]hello-lion@hotmail.com, [d]luojuhua@126.com, [e*]Corresponding author: huangwj@nercita.org.cn

Keywords: Geospatial Statistical Analysis, Landsat TM imagery, Remote Sensing, Soil Fertility, Support Vector Machine (SVM).

Abstract. This paper aims to investigate the soil fertility of Shunyi District's cropland combing remote sensing and ground census data based on the Geostatistical Analyst of ArcGIS. Firstly, Landsat TM image was used to identify the spatial distribution and estimate the cropland plot area using support vector machine (SVM) classification method, and the overall classification is 91.5 % by 435 field survey points. Then, the survey indicators were added to ArcGIS such as organic matter, available P, available K, total N, soil pH, etc. After exploring the sample data for each indicator, trend surfaces were generated using the optimum prediction models after cross validation. Finally, according to the identified cropland plots, the soil quality index (SQI) was derived to map the soil fertility of the study area. The result shows that the southwestern part and northeastern corner of this district were found to be high in soil pH, which lies in between 8.2 and 8.6. Additionally, wide variability of organic matter, total N, available P and K were noted which can be due to the extent of cultivation in these areas while the change in fertility level could be due to anthropogenic influence. When considering the soil heavy metals, Zn, Fe, Cu and Mn show almost the same distribution.

Introduction

As one of the most important agricultural resources, cropland is the basic survival condition for human being. Accurate information on cropland area and soil fertility is of critical importance for assessing food security [1]. Soil fertility refers to the amount of nutrients in the soil, which is sufficient to support plant life. Loss of soil fertility can cause soil degradation, which may not only undermine soil productivity, but may also affect environmental health [2]. In addition, assessing the soil fertility can be also very essential to accurately spray chemical fertilizers for farmers in order to reduce the enviromental pollution. On the other hand, soil fertility is the measure of the ability of the soil to supply essential nutrients in the right amounts, and at the correct proportion at the right time [3]. However, the fertility of soil is also determined by the quality of soil physical properties. Deterioration of soil structure due to reasons such as soil erosion, poor land management practices, as well as a failure of soils to supply nutrients in the correct amount and and the right time, are indicators of land degredation [4]. What's more, various kinds of human activities have greately changed the cropalnd landscape in recent years. As a result, it is so urgent to find out an effective way to assess the soil fertility especially as the cropland area decreasing and the population growing.

Remote sensing is an essential tool to identify and assess the cropland quality because it facilitates observations across larger extents of Earth's surface than is possible by ground-based observations. With the increasing development of aerospace technology, various sensors mounted on air- and space-borne platforms can yield aerial photographs, satellite imagery, RADAR and LiDAR datasets. Data available from remote sensing vary from the very high-resolution datasets produced irregularly over extents no larger than a single state or province, to regional datasets produced at regular intervals

from satellites (e.g. Landsat, SPOT), to the lower-resolution (> 250 m) datasets now produced across the entire Earth on a daily basis (e.g., AVHRR, MODIS). Those remotely sensed iamgey have been widely applied in the assessment of soil fertility [5-7]. In this study, Shunyi District of Beijing is selected as the study area and Landsat TM image and ground census data were jointly used to investigate the soil fertility of cropland.

Description of the study area

Shunyi District lies to the northeast of central Beijing, about 30 kilometers from the centre. It is located between latitudes 40°00'-40°18' N and longitude 116°28'-116°58' E. The total area is 1,021 km^2 and it has a population of 593,000, of which 419,000 are permanent agricultural residents (Beijing Statistical Office, 2001). Shunyi has a warm temperate wet continental monsoon climate. Its average annual temperature is 11.5 °C, that in January 4.9 °C, and in July 25.7 °C. The lowest temperature in January is -19.1 °C and the highest in July is 40.5 °C. The frost-free period lasts around 195 days. Annual sunshine duration is 2,750 hours and average annual relative humidity is about 50%. Average annual precipitation is about 625 mm, 75% of which falls in summer. It has a fertile soil ranging from sandy to loamy soils and the topography is mostly plains with a small remote hilly area, so Shunyi District is a farming community [8]. In this study area, arable cropping, residential and industrial use and forest are the three major land use types. In the past decades, land use in Shunyi has undergone a quick change due to the fast economic growth and urbanization, characterized by conversion of arable land into built-up and green area.

Data Sources and Preprocessing

A Landsat TM image to cove the whole study area of 2010 was acquired, which has six optical bands (visible plus near- and mid-infrared) with 30 m resolution. It retains as much of the original radiometric and geometric properties as possible and was just systematically processed by radiometric and geometric corrections. Therefore, accurate radiometric and geometric corrections must be firstly performed. In this study, radiometric calibration and atmospheric correction of Landsat TM were processed using the FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes) atmospheric correction module which is integrated in the ENVI image processing & analysis software. In addition, topographic relief also exerts a great influence on identifying croplands, so topographic correction must be also performed. The commercial ERDAS Imagine processing software was used to co-register and orthorectify the Landsat TM image using the orthocorrection model of Leica Photogrammetry Suite (LPS) and ASTER 30 m resolution DEM (Digital Elevation Model). As a result, the RMSE (Root Mean Squared Error) was less than 0.5 pixel which meets the accuracy requirement of geometric correction and the nearest neighbor method was used for image resampling. Additionally, field survey was also performed to derive the ground census data of soil fertility.

Methodology

SVM Classification Method. Support Vector Machine (SVM) is a new popular supervised classification algorithm, because of its ability to "learn" classification rules from a set of training data, and moreover it is fit for processing the high dimensional data [9, 10]. It is a classification technique derived from statistical learning theory and sorts out the classes with a decision surface that maximizes the margin between the classes. The SVM training algorithm promises to obtain the optimal separating hyperplane for a training data set in terms of generalization error. The support vectors are the critical elements of the training set. In this study, the commercially available software ENVI 4.5 was used to perform the SVM classification. Given a set of samples x, x= $\{(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k) | x_i \in R^N, yi \in (-1, +1)\}$ where i=1, 2,…, k and R^N is the N-dimensional space. The general optimal classification function is described as in Eq. 1.

$$f(x) = \text{sign}\left[\sum_{i=1}^{k} a_i' y_i K(x_i, x_j) + b\right] \quad (1)$$

where, x_i and x_j are the training vectors; k is the total number of samples, and $K(x_i, x_j)$ is called the kernel function. The main function of such a kernel is to map the training vectors into a higher dimensional space, and its choice is very crucial for obtaining perfect classification performance. In general, there are mainly four types of kernels: linear, polynomial, radial basis function (RBF) and sigmoid. The RBF is one of the commonly used kernel functions and works well in most cases. Some applications have proved that the SVM detects the study targets more accurately with a lesser false alarm rate [11, 12]. In our experiment, RBF was chosen as the kernel function for SVM classification.

Geostatistical Analysis. Geostatistical analysis is an approach to applying statistical analysis and other informational techniques to geographically based data. Such analysis employs spatial software and analytical methods with terrestrial or geographic datasets, including geographic information systems and geomatics [13]. Geostatistical method uses variogram (semivariance) and Kriging interpolation as the basic tools, which can be used to study the various variables with randomness and structure. Kriging interpolation uses the known data point to estimate the unknown point (X_0). The nature of this method is to derive a weighted average by local estimation (Eq. 2). In this study, ArcGIS Geostatistical Analyst (AGGA) was used to explore the soil fertility of cropland and it builds a bridge between statistics and GIS. AGGA is an extension to ArcGIS desktop that provides tools for spatial data exploration and surface generation and it provides a suite of statistical models and tools for spatial data exploration and surface generation.

$$Z(X_0) = \sum_{i=1}^{n} \lambda_i Z(X_i) \quad (2)$$

where, $Z(X_0)$ is the inner-interpolation value on the X_0 point, and $Z(X_i)$ is the measured value around the X_0 point.

Soil Fertility Assessment. In this study, soil samples were collected and analyzed as an input for the soil fertility capability classification (FCC), and soil fertility index (SFI) was used to assess the soil fertility. In each sampling region, essential features necessary for site characterization were specifically gathered in the area. For each sample, several physicochemical parameters were collected such as organic matter, available P, available K, total N, soil pH, exchangeable Zn, etc. These samples were analyzed for nutrient availability indicators based on the minimum data set (MDS) for FCC and SFI determinations. Here, the SFI was obtained following the conceptual model developed by Andrews *et al*. [14]. In this quantitative soil quality evaluation method, the soil management assessment framework (SMAF) is designed to follow three basic steps: indicator selection, indicator interpretation, and integration into a soil quality index (SQI) value (Eq. 3). The final step is accomplished by summing the scores for each indicator, dividing by the total number of indicators, and then multiplying by 10. When the cropland plots were indentified from Landsat TM image using the ENVI-SVM classification module, they were input into the SMAF assessment framework by integrating soil census data.

$$SQI = \left(\frac{\sum_{i=1}^{n} S_i}{n}\right) \times 10 \quad (3)$$

where, S represents the scored indicator value and n is the number of indicators in the MDS.

Results and Discussions

Cropland Plots Identification. The land of the study area account for most of the agricultural land, so those regions covered by field crops, vegetables and flowers, were assigned to cropland and all other land cover types were classified as non-cropland. Additionally, it must be mentioned that perennial tree crops (including fruit trees) which present very little percentage were classified into

non-crop classes. Owing to the similar spectral reflectance features with forest, it is very difficult to exclude the tree crops from forest without the support of other ancillary data. As a result, an exclusion method was applied to the Landsat TM classification data to produce the crop/non-crop map. At the initial stage, non-vegetation area including water area, built-up area and grassland were derived. Then, forestland map was produced. After extracting the forest and non-vegetation land, the rest land was treated as cropland plots. In order to evaluate the identification accuracy of croplands, 435 sample points was used. The result shows that 398 points are classified as cropland and the overall accuracy is 91.5% (398/435*100%). Fig. 1 shows the original Landsat TM image and the identified croplands. It is clear that the croplands distribute almost everywhere in the study area besides the Southwestern regions and Northeastern corner because of built-up areas and forestland.

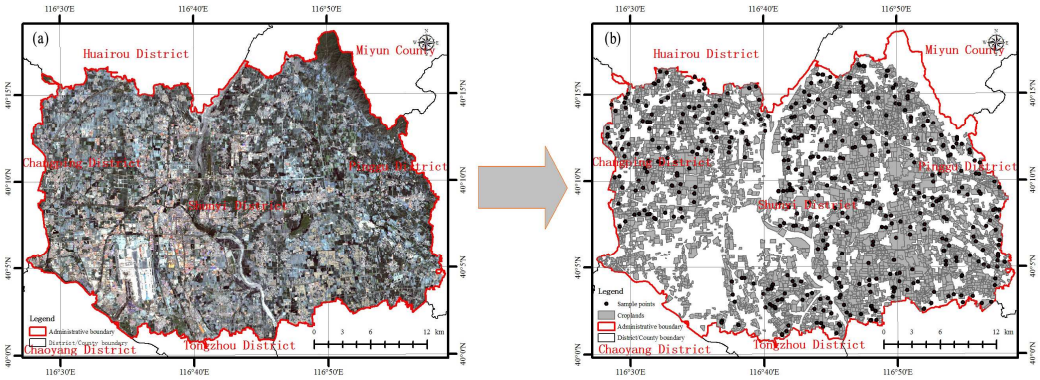

Fig. 1 (a) Original Landsat TM false color composite image by Blue band (TM1), Green band (TM2) and Red band (TM3) (b) is the identified croplands of Shunyi District, Beijing.

Classification of Soil Fertility. In order to derive the SFI, among the survey indicators, organic matter, available P, available K, total N, soil pH, Zn, Fe, Cu, and Mn were selected as the MSFI (Minimum Soil Fertility Indicators). With these indicators, samples collected in the study were analyzed to assess the soil fertility. On the analysis platform of AGGA, the sample points for each indicator were interpolated into trend surfaces for reflecting the spatial distribution and variation of soil nutrient (Fig. 2).

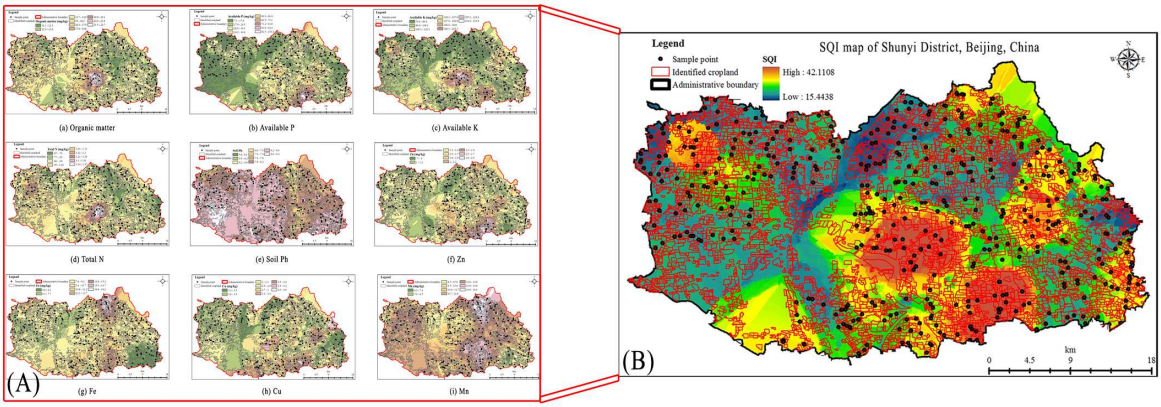

Fig. 2 (A) is the spatial distribution maps for each indicator (B) is the final SQI map to describe the soil fertility of the study area.

As shown in Fig. 2, it is obvious that the southwestern part and northeastern corner of the district were found to be high in soil pH, which lies in between 8.2 and 8.6. Wide variability of available P and K were noted which can be due to the extent of cultivation in these areas, while the change in fertility level could be due to anthropogenic influence. Concerning the soil heavy metals, Zn, Fe, Cu and Mn show almost the same distribution. Likewise, organic matter and total N show also the almost same distribution rules. Owing to the land use types and human influence, the spatial variability for

each soil indicator shows the same or different change trends. As a whole, the soil fertility is better in the west eastern and northeastern parts than that in the western and northern parts. In the northwestern part, it is worse due to the human activities and land was changed to built-up areas.

Conclusions

Compared with traditional labor-intensive survey methods, remote sensing provides a fast, accurate, large-scale, affordable tool for detecting and mapping the soil fertility. In this study, some indicators were collected including organic matter, available P, available K, total N, soil pH, Zn, Fe, Cu, and Mn. Based on the analysis platform of AGGA, the sample points for each indictor were interpolated into trend surfaces. (1) As can be seen in Fig. 1 and Fig. 2, they show that the croplands distribute almost everywhere in the study area except the Southwestern part and Northeastern corner because of built-up areas and forest. (2) Owing to the different land use types and human influence, the spatial variability for each soil indicator shows the same or different change trends. (3) Taken as a whole, the soil fertility is better in the northwestern, southern and northeastern parts than that in the western and northern parts. In the northwestern part, it is worse due to the human activities and land was changed to built-up areas. Therefore, it can provide an effective solution to assess and map soil fertility by integrating remote sensing techniques with geostatistical functions of geographic information system.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (41071276), the Program of Ministry of Agriculture (200903010), the Special Funds for Major State Basic Research Project (2007CB714406) and the Postdoctoral Science Foundation of Beijing Academy of Agriculture and Forestry Sciences (2011).

References

[1] G.L. Hou, H.Y. Zhang, Y.Q. Wang and Z.X. Zhang: J. Nat. Resour. Vol. 25 (2010), p. 1607 (in Chinese)

[2] H.J. Tang, J. J. Qiu, E. V. Ranst and C.S. Li: Geoderma Vol. 134 (2006), p. 200

[3] D.L. Rowell: Soil Sciences: Methods and Applications (Longman, London, England 1994).

[4] Information on http://www.lucideastafrica.org

[5] L.N. Fang and J.P. Song: Progr. Geogr. Vol. 27 (2008), p. 71 (in Chinese)

[6] E.F. Moran, E.S. Brondizio, J.M. Tucker, M.C. de Silva-Forsberg, S. McCracken and I. Falesi: Forest Ecol. Manag. Vol. 139 (2000), p. 93

[7] P.J. Vickery, D.A. Hedges and M.J. Duggin: Remot. Sens. Environ. Vol. 9 (1980), p. 131

[8] J.L. Zhao, X.J. Pan, R.F. Sui, S.R. Munoz, R.D. Sperduto and L.B. Ellwein: Am. J. Ophthalmol. Vol. 129 (2000), p. 427

[9] C. Nello and S.T. John: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods (Cambridge University Press, Cambridge, England, 2000)

[10] V.N. Vapnik: The Nature of Statistical Learning Theory (New York: Springer-Verlag, 1995).

[11] H. Nemmour and Y. Chibani: ISPRS J. Photogramm. Vol. 61 (2006), p. 125

[12] M. Trebar and N. Steele: Comput. Electron. Agr. Vol. 63 (2008), p. 119

[13] Information on http://dictionary.reference.com/browse/geospatial

[14] S.S. Andrew, D.L. Karlen and C.A. Cambardella: Soil Sci. Soc. Am. J. Vol. 68 (2004), p.1945

Renewable and Sustainable Energy
10.4028/www.scientific.net/AMR.347-353

Assessing the Soil Fertility Using Landsat TM Imagery and Geospatial Statistical Analysis

10.4028/www.scientific.net/AMR.347-353.3559