Optik 127 (2016) 1071–1077

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Major crops classification using time series MODIS EVI with adjacent years of ground reference data in the US state of Kansas

Shakir Muhammad[a,b,1,2], Yulin Zhan[a,*,1], Li Wang[a,b,1], Pengyu Hao[a,b,1], Zheng Niu[b,1]

[a] The State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China
[b] University of Chinese Academy of Sciences, Beijing 100039, China

ARTICLE INFO

Article history:
Received 5 January 2015
Accepted 25 October 2015

Keywords:
Optical
MODIS
EVI
Machine learning ABNet
USDA

ABSTRACT

Most methods used large quantity of field data of the same reference year for crops classification which is labor-intensive and time-consuming. In this study we explored the optical application of time series MODIS EVI with adjacent years of ground reference data for classifying major crops on a regional level in US state of Kansas. Time series MODIS EVI data have been obtained between 2008 and 2013. Ground reference data (2008–2013) of the major crops (winter wheat, corn, soybeans, sorghum and alfalfa) in Kansas were acquired from the United States Department of Agriculture (USDA). A machine learning algorithm namely Antibody Network (ABNet) classifier was used to classify the major crops. The ABNet was trained using five years of ground reference data and verified by ground reference data of the other year. For instance, to classify major crops in 2008, ground reference data of (2009–2013) were used as training samples and the data of that year (i.e. 2008) were used as validation. The results evince the classification accuracy in a range from 74.4 to 81.9% and kappa coefficient of 0.6–0.8 respectively. This method can improve remote sensing imagery in the process of classification and can help to alleviate the heavy load of field data in areas where ground data are unavailable.

1. Introduction

With the increase in population the landuse land cover faced severe consequences of change globally. These changes drew serious environmental change issues across the globe which can be entertained by having regional or global level landuse land cover data [1–3]. The direct or indirect impacts of these changes on environment, economy and social system implored that these changes need to be monitored with regular time interval for a particular area in order to provide information that can map the human environment interaction [4]. An updated and regular time based land use land cover dataset are needed for the regions significantly dominated by agricultural activities like US central Great Plains. The cropland is of specific importance because it undergoes a process of continuous change which has its influence on food availability, ground water table, environmental processes, ecology, and economy [5]. Croplands like US central great planes need to be monitor with proper time interval as such places act as the main food sources for the world and its long time sustainability can strengthen the continuity of food supply for the future generations.

Time series satellite data can depict the stages of crop life [6–8]. Time series multi-temporal data has been tested to classify crops [9,10]. Most of the methods depend on the ground reference data collected in the mapping year for training the classifier [5,11–13]. Collecting these data are costly, difficult, and time consuming [14]. To overcome this problem researchers used the data collected in other years as substitute to train the classifier in the case of unavailability of the data in the specific year [15,16] but they did not show any turn to develop methods for classifying multiyear crops. In recent years methods were developed to classify multiyear crops by utilizing single or multiyear field data other than classification year [9,17,18]. Developing such methods are challenging especially for the crop types exhibiting the same crop calendar. These methods prefer the use of vegetation profile from time series vegetation indices as the base for crop discrimination as vegetation indices shows a more uniformity over spectral profiles [9,19].

* Corresponding author at: P.O. Box 9718, No. 20 Datun Road, Olympic Science & Technology Park of CAS, Beijing, China. Tel.: +86 10 64889215; fax: +86 10 64889215.
E-mail addresses: mshakir@radi.ac.cn (S. Muhammad), zhanyl@radi.ac.cn (Y. Zhan), wangli@radi.ac.cn (L. Wang), haopy8296@163.com (P. Hao), niuzheng@radi.ac.cn (Z. Niu).
[1] Address: P.O. Box 9718, No. 20 Datun Road, Olympic Science & Technology Park of CAS, Beijing, China. Tel.: +86 10 64889215; fax: +86 10 64889215.
[2] Tel.: +86 13641064964.

http://dx.doi.org/10.1016/j.ijleo.2015.10.107
0030-4026/

Huge variations are found in the collection of field data for training classification schemes. Some used easily accessible survey points along the roads [20] while some went to more detailed survey of each field [21]. Others collect the relevant information by interviewing the farmers [22], while some used high resolution imagery to interpret different classes of landuse [23]. General land cover classes can be classified by using MODIS 250 m data [24]. This shows that this datasets would have high ability to differentiate crop types in the US central great plains of Kansas where average field size is more than 32.5 ha that constitute of average 5 pixels of MODIS 250 m pixels avoiding the mixing of pixels containing different crops.

The intention of this research was (1) to classify major crops using time series MODIS EVI with adjacent year's reference data in order to lower the cost of field data collecting through regular time interval. Although classification of crops has been done for the region of Kansas using remote sensing datasets but these methods rely on the field data collected in the mapping year which is costly to collect through such vast area. In this study we (2) examine the potential of a Machine learning algorithm namely ABNet classifier with time series MODIS EVI data for classification of major crops. This classifier is subjected to use the ground reference data of years other than classifying year to discriminate the crops i.e. the samples for training and validation are from different years. This research will test the adjustability of ABNet classifier to time series MODIS EVI and ground reference data for classification of major crops in Kansas.

2. Study area and material

2.1. Study area

The sample field data were acquired from the US state of Kansas situated between 37° and 40° N latitude and 94° and 102° W longitude. It is considered as one of the most intensely agriculture area across the United States covering an area of (213,096 km^2). It has border with Nebraska in the north, Missouri in the east, Oklahoma in the south and Colorado in the west. The eastern part contains some mountains and forest while rest is entirely dominated by croplands. The major crop types contain winter wheat (*Triticum aestivum*), corn (*Zea mays*), soybeans (*Glycine max*), sorghum (*Sorghum bicolor*) and Alfalfa (*Medicago sativa*). The climate of the state ranges from humid continental to humid subtropical on the basis of Koppen climate classification.

2.2. Time series MODIS EVI data

The Moderate Resolution Imaging Spectroradiometer (MODIS) provides an opportunity to monitor the landuse and land cover changes over global scale by providing data in different range of temporal and spatial resolutions. Studies concluded that MODIS vegetation indices with 250 m spatial resolutions are suitable to classify different crops in the US state of Kansas [5]. As the average field size in Kansas 32.4 ha or larger which constitute around five or more 250-m MODIS pixels. Moreover, EVI is similar to NDVI in characteristics of vegetation [5] and it is more preferred over NDVI as it has corrections in reference to canopy background like (soil and bare earth) [25]. It can be explain by the equation given below.

$$\mathrm{EVI} = G\frac{\rho\mathrm{NIR} - \rho\mathrm{Red}}{\rho\mathrm{NIR} + C1\rho\mathrm{red} - C2\rho\mathrm{blue} + L}$$

where ρ is the atmospherically corrected reflectance in NIR, red and blue bands, and $C1$ and $C2$ are the coefficient with values of 6.0 and 7.5 respectively while L is the soil adjustment factor with value 1.0 and G is the gain factor with value of 2.5 [26].

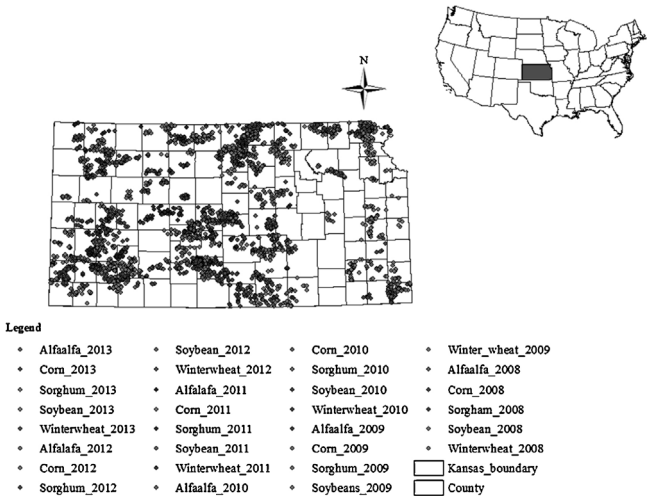

Fig. 1. US state of Kansas on the country map (up right) in dark while (lower left) shows the state of Kansas with all the field samples for all major crops for six years.

MODIS aqua (MYD13Q1) EVI data with 250 m spatial resolution and 16 days temporal resolution for the time span of six years (2008–2013) was acquired from the National Aeronautics and Space Administration (NASA). The acquired MODIS EVI data spanning from January to December constitute the whole growing season went under pre-processing. Kansas area contained within three MODIS 250 m scenes which were mosaicked and re-projected using Modis Re-projection Tool acquired from NASA website. Each remote sensing data contains different types of noises for which researchers developed methods to remove [27,28] TIMESAT software is used to filter the MODIS data in order to acquire a smooth and even EVI data [29,30]. Savitzky Golay filter was applied to the time series MODIS data with best fit. A single year contains 23 images with 16 days interval which were stacked in order of day of the year to retrieve vegetation profiles (Fig. 1).

2.3. Ground reference datasets

Under the auspices of United State Department of Agriculture (USDA), National cropland data layer (CDL) (http://www.nass.usda.gov/research/Cropland/SARS1a.htm) is a cropland mapping program designed to map landuse land cover changes especially agriculture changes and crop acreage across the US. Kansas with the approximate area of 213,000 km^2 is considered as one of the world most intensive agricultural areas. Such data along with MODIS provide an opportunity to develop and test the potential of different classification schemes for classifying these croplands. Polygon based cropland datasets for six years and five major crops (winter wheat, Corn, soybeans, sorghum and alfalfa) were acquired for the state of Kansas. For these five major crops the accuracy of classification in (CDL) system was above 95% for all the six years. This accuracy was resulted by using remote sensing data with classification procedures and validating the results through ground observations [31]. The polygon based data is overlaid over time series MODIS EVI stacked images for selecting the most centrally located pure pixel or 'best case' as did by [8,18]. The process of selection of the most centered pixel within a specific crop polygon decreases the chance of mixing pixels of different crop classes. An extensive 3300 samples for these five major crops were taken throughout the study area. These samples are the best representations of their respective crops which will be used for training and validation (Fig. 2).