ISPRS Journal of Photogrammetry and Remote Sensing 119 (2016) 151–164

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Automated mapping of soybean and corn using phenology

Liheng Zhong [a,*], Lina Hu [b], Le Yu [c], Peng Gong [c,d], Gregory S. Biging [a]

[a] Department of Environmental Sciences, Policy & Management, University of California, Berkeley, CA 94720, United States
[b] Department of Sociology, Tsinghua University, Beijing 100084, China
[c] Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, Beijing 100084, China
[d] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Beijing 100101, China

ARTICLE INFO

Article history:
Received 28 January 2016
Received in revised form 21 May 2016
Accepted 23 May 2016
Available online 13 June 2016

Keywords:
Automated classification
MODIS
Corn
Soybean
Brazil

ABSTRACT

For the two of the most important agricultural commodities, soybean and corn, remote sensing plays a substantial role in delivering timely information on the crop area for economic, environmental and policy studies. Traditional long-term mapping of soybean and corn is challenging as a result of the high cost of repeated training data collection, the inconsistency in image process and interpretation, and the difficulty of handling the inter-annual variability of weather and crop progress. In this study, we developed an automated approach to map soybean and corn in the state of Paraná, Brazil for crop years 2010–2015. The core of the approach is a decision tree classifier with rules manually built based on expert interaction for repeated use. The automated approach is advantageous for its capacity of multi-year mapping without the need to re-train or re-calibrate the classifier. Time series MODerate-resolution Imaging Spectroradiometer (MODIS) reflectance product (MCD43A4) were employed to derive vegetation phenology to identify soybean and corn based on crop calendar. To deal with the phenological similarity between soybean and corn, the surface reflectance of the shortwave infrared band scaled to a phenological stage was used to fully separate the two crops. Results suggested that the mapped areas of soybean and corn agreed with official statistics at the municipal level. The resultant map in the crop year 2012 was evaluated using an independent reference data set, and the overall accuracy and Kappa coefficient were 87.2% and 0.804 respectively. As a result of mixed pixel effect at the 500 m resolution, classification results were biased depending on topography. In the flat, broad and highly-cropped areas, uncultivated lands were likely to be identified as soybean or corn, causing over-estimation of cropland area. By contrast, scattered crop fields in mountainous regions with dense natural vegetation tend to be overlooked. For future mapping efforts, it has great potential to apply the automated mapping algorithm to other image series at various scales especially high-resolution images.

1. Introduction

Soybean and corn are two of the most important commodities in the global crop market. Quantitative and spatially-explicit information of soybean and corn is of great value to investment, market forecast and resource management in the agricultural sector. For the vast cropland in North and South America, soybean and corn are fundamental components in the study of land cover and land use change which affects social welfare (Chavas and Holt, 1996; Qaim and Traxler, 2005), carbon and water cycles (Drinkwater et al., 1998; Anderson et al., 2004; Hill et al., 2006), and biological conservation (Soares et al., 2006; Mosier et al., 2006). Detailed information of the planting area can be incorporated into crop yield models to improve the yield estimate (Xin et al., 2013; Johnson, 2014; Sakamoto et al., 2014; Lobell et al., 2015). Also, the planting area is a useful input to a range of models and studies involving various natural, sociological and economic factors (Howard and Wylie, 2014). Examples include agricultural expansion (Turner et al., 2007; Lambin and Meyfroidt, 2011), the interaction between agriculture and environment (Rounsevell et al., 2003; DeFries et al., 2004; Sakamoto et al., 2009), the trend of farming intensification and bioenergy use (Wang et al., 2014), global food security (Thenkabail et al., 2009), and the impact of climate change on agricultural water demand. All these social-economic and

* Corresponding author at: 222 Mulford Hall, UC Berkeley, CA 94720, United States.
E-mail address: lihengzhong@berkeley.edu (L. Zhong).

http://dx.doi.org/10.1016/j.isprsjprs.2016.05.014
0924-2716/

scientific applications highlight the need to produce soybean and corn maps regularly, reliably and timely at low cost.

For major soybean and corn production areas, a variety of land cover products is available at multiple spatial scales. Global products provide the global distribution of cropland (Friedl et al., 2002; Arino et al., 2008; Monfreda et al., 2008; Ramankutty et al., 2008; Pittman et al., 2010; Portmann et al., 2010; Gong et al., 2013). However, cropland was usually considered as one or a few general categories, and information about individual crop types was not derived directly from remote sensing classification (Yu et al., 2013; B. Wu et al., 2014). Crop-specific mapping of soybean and corn was often conducted at regional scale, mostly focusing on US Corn Belt (Wardlow et al., 2007; Wardlow and Egbert, 2008; Ozdogan and Gutman, 2008; Boryan et al., 2011; Howard and Wylie, 2014) and the tropical and temperate plains in South America (Vieira et al., 2000; Sugawara et al., 2008; Epiphanio et al., 2010; Arvor et al., 2011, 2012; Mercante et al., 2012; Gusso et al., 2012; Wachholz de Souza et al., 2013; Brown et al., 2013; dos Santos et al., 2014). All of these products were based on the spectral features of land cover classes in certain phases using supervised or unsupervised classification methods. The algorithms of these land cover datasets relied on image-specific statistics, training data collection and visual interpretation. As a result, these algorithms may be subject to low cross-year repeatability and consistency and the high cost of human labor. Although in many existing studies the stages of plant development were considered, for example, to select the imaging date with the greatest separability, time-related measurements were not explicitly employed to improve the cross-year repeatability. In South America, the implementation of a classifier to separate corn and soybean over multiple years is especially challenging because of the difficulty of repeatedly collecting reference data that are temporally consistent.

Conventional supervised classification divides all pixels in an image into classes based on training data which are usually a subset of the image. Results are image- or time-specific and additional reference data collection and training are required when the method is applied to other periods. The need of mapping cropland rapidly, consistently and repeatedly calls for an automated mapping algorithm. When completed and published, an automated algorithm provides trained rules that are ready to be used directly by other researchers, and its application could be repeated year after year without re-calibration or re-training (Macdonald and Hall, 1980; Badhwar, 1984; Thenkabail et al., 2009; Pena-Barragan et al., 2011; Z. Wu et al., 2014; Zhang et al., 2015). In the long run, an automated mapping algorithm could produce consistent cropland cover at very low cost (Lucas et al., 2007; Waldner et al., 2015). Although the long-term application is effective and economical, at the development stage, the automated algorithm requires substantial expert input and image analysis to isolate type-specific properties from inter-annual and inter-region variability. Applying automated algorithms to the identification of individual specific crop types remains a challenge. The difference in spectral signals between crops can be trivial or observable only at certain development stages (Reis and Taşdemir, 2011). The process is overwhelming to human labor especially when a large set of time series is used as the input to the classification algorithm, which is often the case for multi-temporal classification nowadays.

The use of phenology from multi-temporal images plays an important role in crop classification (Friedl et al., 1999; Knight et al., 2006; Evans and Geerken, 2006; Geerken, 2009). The identification of crop classes may benefit from phenological information either by employing phenological transitions to interpret multi-temporal profiles, study the seasonal dynamic of separability and select the optimum imaging date for classification, which is an essential step of many existing efforts (Lloyd, 1990; Simonneaux and Francois, 2003; Conrad et al., 2010; Son et al., 2014; Siachalou et al., 2015), or by explicitly deriving quantitative phenological metrics and building classification rules based on crop calendar and phase-dependent crop conditions (Zhong et al., 2011, 2012; Dong et al., 2015; Qin et al., 2015). Phenological metrics and have the advantage of bearing a physiological signification. Consequently, rules built based on phenological metrics are more robust than statistical methods, which are more dependent on specific datasets and are therefore harder to generalize (Simonneaux et al., 2008). Phenological stages derived from image series are useful in crop discrimination in two aspects. First, phenological stages could be directly used to separate crops with distinct crop calendars. Second, when phenological stages are available, it is possible to derive crop spectral properties at certain phenological stages (Zhong et al., 2014). The mere use of phenological stages may be subject to the inter-annual and inter-regional variation of crop calendar. Under certain circumstances, the dynamic of spectral properties offers more reliable measurements to identify crop types. The use of phenology enables the cross-year alignment of input data using phenological stages rather than original imaging dates, reducing the influence of inter-annual variability on the automated algorithm (Zhong et al., 2014). In the development of an automated classification algorithm, the comparison of crop spectral characteristics should be undertaken at specific phenological stages to minimize the impact of crop calendar variation. In general, the use of remote sensing based phenology information is capable of expediting the process of rule development and improving the stability of rules over time and space by utilizing interpretable metrics from image series.

The objective of this study was to present a robust automated classification approach to map soybean and corn repeatedly, consistently and efficiently at low cost. Building on previous work (Zhong et al., 2014), we used phenological metrics in crop classification for cross-year classifier extension, which applies a universal set of rules to multiple years.

2. Materials and method

2.1. Study area

The study area is the state of Paraná in the south of Brazil. Paraná occupies an area around 20 million hectares, with 399 administrative units at the municipal level (municipality). The major climate type is subtropical to the north and temperate to the south (Fig. 1).

Paraná was the main grain growing area of Brazil before the reclamation of Amazon in recent decades. At present Paraná is the second largest producer of soybean and corn in Brazil. In the crop year 2013, soybean production of Paraná was about 16 million tons, and corn production was about 17 million tons, accounting for about 20% and 22% of the national total, respectively.

In Paraná, soybean and corn are usually grown one or two seasons. Soybean is mostly cultivated in the main season between the spring and summer, approximately from October to April. A small portion of soybean is grown later as the second crop (safrinha) from February to June. Corn production occurs abundantly in both seasons, with the first season roughly from September to April, and the second season from February to August. The rotation between soybean and corn is very common in the main croplands. Other crops include wheat, sugarcane, dry beans, and so forth. Soybean and corn comprise over 75% of the total area of annual crops or over 90% of the total area in the summer.

2.2. MODIS

The primary input data were the MODerate-resolution Imaging Spectroradiometer (MODIS) product MCD43A4, which offers nadir