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

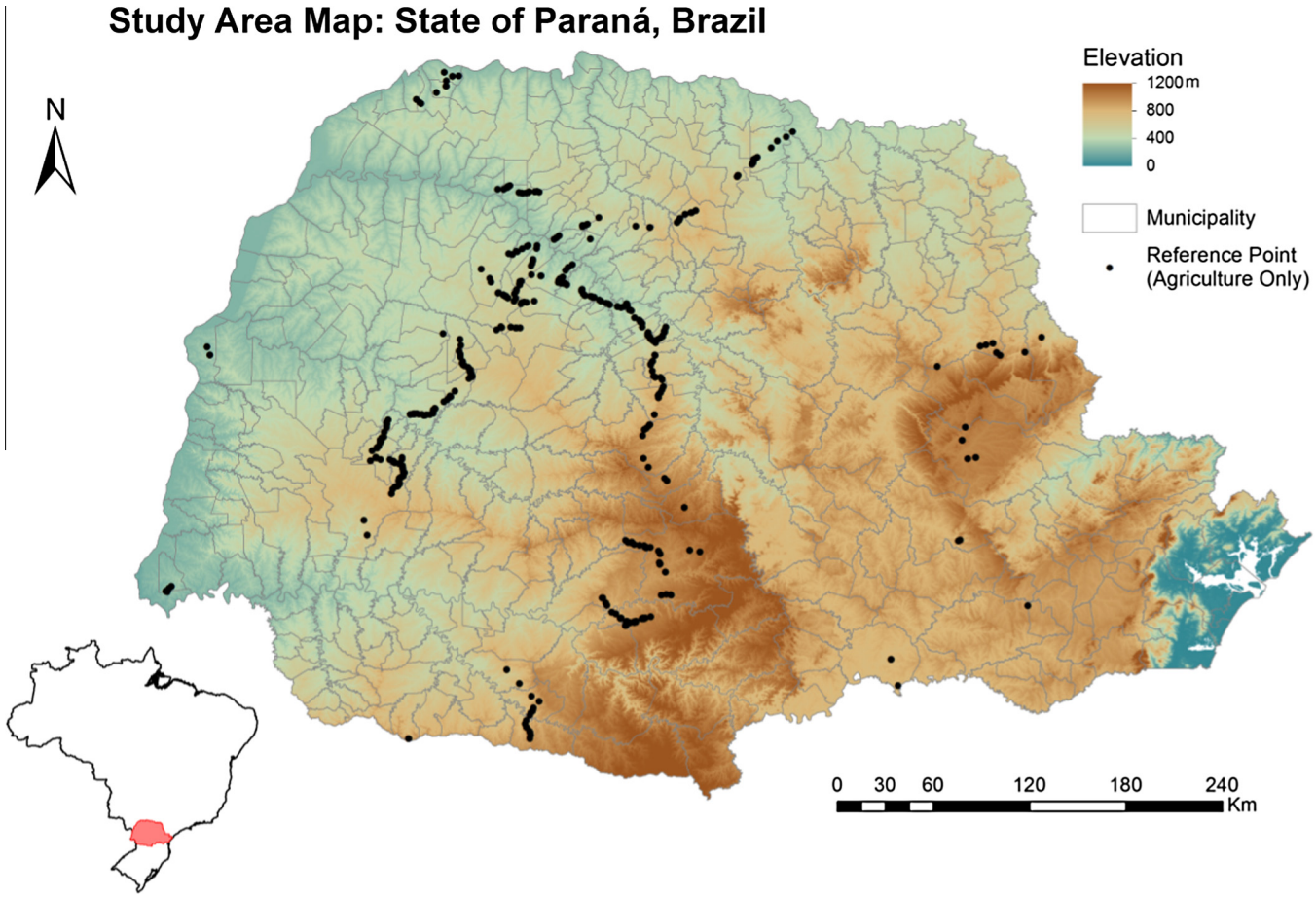

Fig. 1. The state of Paraná, Brazil. Topography and municipal boundaries are displayed. Black dots represent locations of ground reference data in the crop year 2012 which have higher density in cropped areas.

and bidirectional reflectance distribution function (BRDF) adjusted spectral reflectance bands (Lucht et al., 2000; Schaaf et al., 2002). MCD43A4 uses daily Terra and Aqua satellite overpasses to compute reflectance within each 16-day period by an algorithm based on the inversion of radiative transfer models. The algorithm adjusts reflectance at local solar noon and reduces the effects of view angle and anisotropic scattering. MCD43A4 reflectance is computed for MODIS spectral bands 1–7 at 500 m resolution every eight days. The reflectance product also comes with a detailed set of quality information product MCD43A2. Therefore, MCD43A4 is a relatively stable and consistent reflectance product which at the same time offers acceptable resolution and frequency. It is directly used as the primary input in the production of various land use maps like the global MODIS Land Cover Product MCD12Q1 (Friedl et al., 2010).

A crop year is defined by the harvest date and usually ranges from the spring of the previous calendar year to the winter of the current calendar year. For example, to include all crop growth cycles in the crop year 2012, a time window of July 1st, 2011 to November 1st, 2012 was used for image selection. For each crop year, all MCD43A4 images within the time window were processed into multi-spectral time series, and low-quality observations were excluded based on per-pixel quality flags.

2.3. Retrieval of phenological stages

Phenological stages were retrieved from time series of crop growth as the main input to the classification algorithm. Crop growth is characterized by the Enhanced Vegetation Index (EVI) which is an optimized index to enhance the vegetation signal by separating background signal and reducing atmospheric conditions (Huete et al., 1997, 2002). EVI was found to be a suitable index with sufficient sensitivity to crop growth in high biomass periods in earlier phenology based classification efforts (Zhong et al., 2014). Phenological stages were obtained by fitting a pre-defined curve to each growth cycle in the EVI series. Asymmetric double sigmoid function was selected as the curve function (Soudani et al., 2008):

$$V(t) = V_b + \frac{1}{2}V_a[\tanh(p(t - D_i)) - \tanh(q(t - D_d))] \quad (1)$$

where $V(t)$ is the EVI at time t and the unit of t is day of year (DOY). Parameters derived from curve-fitting and their uses are described in Table 1 and shown in Fig. 2. All parameters in Eq. (1) along with D_1 to D_4 (Table 1) formed a group of input variables named *phenological*.

In our tests, it was found that this curve-fitting method could generally determine the shape of EVI profile even with partial-season data when EVI just starts decreasing. In the middle of the growing season, transition dates in the increasing segment (D_1, D_2, and D_i) can be precisely derived, and other metrics can be reasonably estimated. Although curve-fitting results are better with full season data available, it is also possible to retrieve phenological stages before the growing season ends for the purpose of rapid mapping.

Table 1
Variables generated for phenology measurement and classification.

Variable	Description	Use in classification
V_b	"Background" EVI value corresponding to the non-growing season	Not used
V_a	The amplitude of EVI variation within the growing cycle	Field crops have higher V_a than natural vegetation. Average V_a of soybean is greater than corn
p, q	Changing rate parameters of increasing and decreasing segments in the cycle, respectively	Field crop cycles show fast EVI increase and decrease
D_i, D_d	Middle dates of segments when the increasing or decreasing rates (first derivative) are maximum	Stable indicators of the dates of rapid growth and harvest
D_1, D_2, D_3, D_4	Four dates when the second derivative of the curve reaches local maximum or minimum, representing starting and end dates of the two segments (Fig. 2)	D_1 and D_4 are used as starting and end dates of the whole growing season
G_{14}	The difference between D_4 and D_1, representing the length of growing season	The range of G_{14} for field crops is relatively consistent. Soybean has slightly shorter G_{14} than corn
R_{SWIR_Di}	MODIS band 6 (1628–1652 nm, shortwave infrared) reflectance at D_i	R_{SWIR_Di} of soybean is slightly higher than corn. Pixels with very high R_{SWIR_Di} should be soybean, and very low R_{SWIR_Di} should be corn, providing high-confidence pixels for training

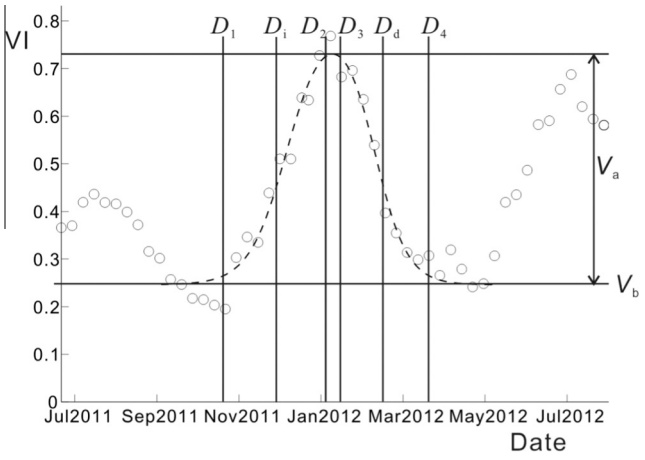

Fig. 2. An example of EVI time series (circle symbol) and the fitted curve (dashed line) for a soybean field in the study area. *Phenological* variables are labeled.

2.4. Spectral metrics at phenological stages

Spectral properties at certain phenological phases, which reveal the internal physiological characteristics of crops, are relatively consistent across regions and years. In this study, at each transition date (D_1 to D_4), values of multi-spectral bands were determined by linearly interpolating the time series. This group of variables, termed *pheno-spectral*, are phenology-specific and include the same bands as the MCD43A4 product.

2.5. Algorithm development

Our automated classification approach utilizes a set of decision rules to identify crop types by making use of the separation power of crop phenology and phenologically adjusted spectral properties. The algorithm is not built directly on any specific data set so that the mapping process can be performed rapidly year after year without considering the conditions of the year of original development. Compared to traditional methods which used data mining tools to generate image-specific rules, we manually built simplified rules by substantial expert input and intensive interpretation of images and multi-dimensional plots to enhance the repeatability of the algorithm and reduce the influence of noises. The process of finalizing classification rules was conducted according to the comparison between data and knowledge by iterations of trial and error. We started with interpreting plots of *phenological* and *pheno-spectral* variables for places that we are familiar with. After identifying the general distribution of corn and soybean in the multi-dimensional space of phenology-related variables, we repeated examining plots of other sub-areas with less available information or more crop mixture. By iterations of experiments, we chose the rule structure and threshold values which can be universally applied to the whole state in all years. Thousands of plots were reviewed for numerous combinations of input variables, sub-areas, and years, and it is impossible to include all details in this paper. An example is shown in Fig. 3, which demonstrates the process of rule development based on *phenological* variables D_i and D_d for some study years. For conciseness, additional plots of more municipalities in all study years are put in Appendix A. Municipalities that evenly spread in the agricultural area of the state are selected to represent soybean and corn under various crop conditions. Campo Mourão (Fig. 3) and Pato Branco (Fig. A.2) are known to be characterized by double cropping of soybean. São Jorge do Ivaí (Fig. 3), Ubiratã (Fig. A.4), São Miguel do Iguaçu (Fig. A.5) and Sertaneja (Fig. A.6) are dominated by the rotation of first soybean and second corn, and the timing of planting second corn is quite different (for example, second corn in Ubiratã is grown earlier than the other municipalities). Soybean and corn mainly in the first season and rarely in the second season in Guarapuava (Fig. 3) and Carambeí (Fig. A.8), etc. These municipalities are listed for the purpose of demonstration, and locations used in actual classifier development during the operational stage were far more than them. The use of phenologically meaningful and consistent variables makes these plots relatively easy to comprehend and facilitates the identification of clusters of the main local crops on each plot. Rules on D_i and D_d were finalized after the area of test gradually extended to the whole state. In a similar manner, plots of other variables were also used to develop rules and the decision tree structure.

2.6. Mapping algorithm

The automated mapping was conducted based on rules in a hierarchical manner (Fig. 4). *Phenological* variables, which represent the seasonal dynamic of EVI and stages of phenological transition during the growing period, are useful in the classification of summer field crop cycles based on crop calendar. Due to the inter-annual and inter-regional variability of crop calendar, we used less restrictive rules on phenology to improve the algorithm repeatability under different conditions. As a result, *phenological* variables alone may not be sufficient to separate soybean and corn cycles fully. To solve this problem, additional classification rules were created based on *pheno-spectral* variables to utilize spectral properties at specific stages to improve the separation. The hierarchical procedure is described in details in the following sections.

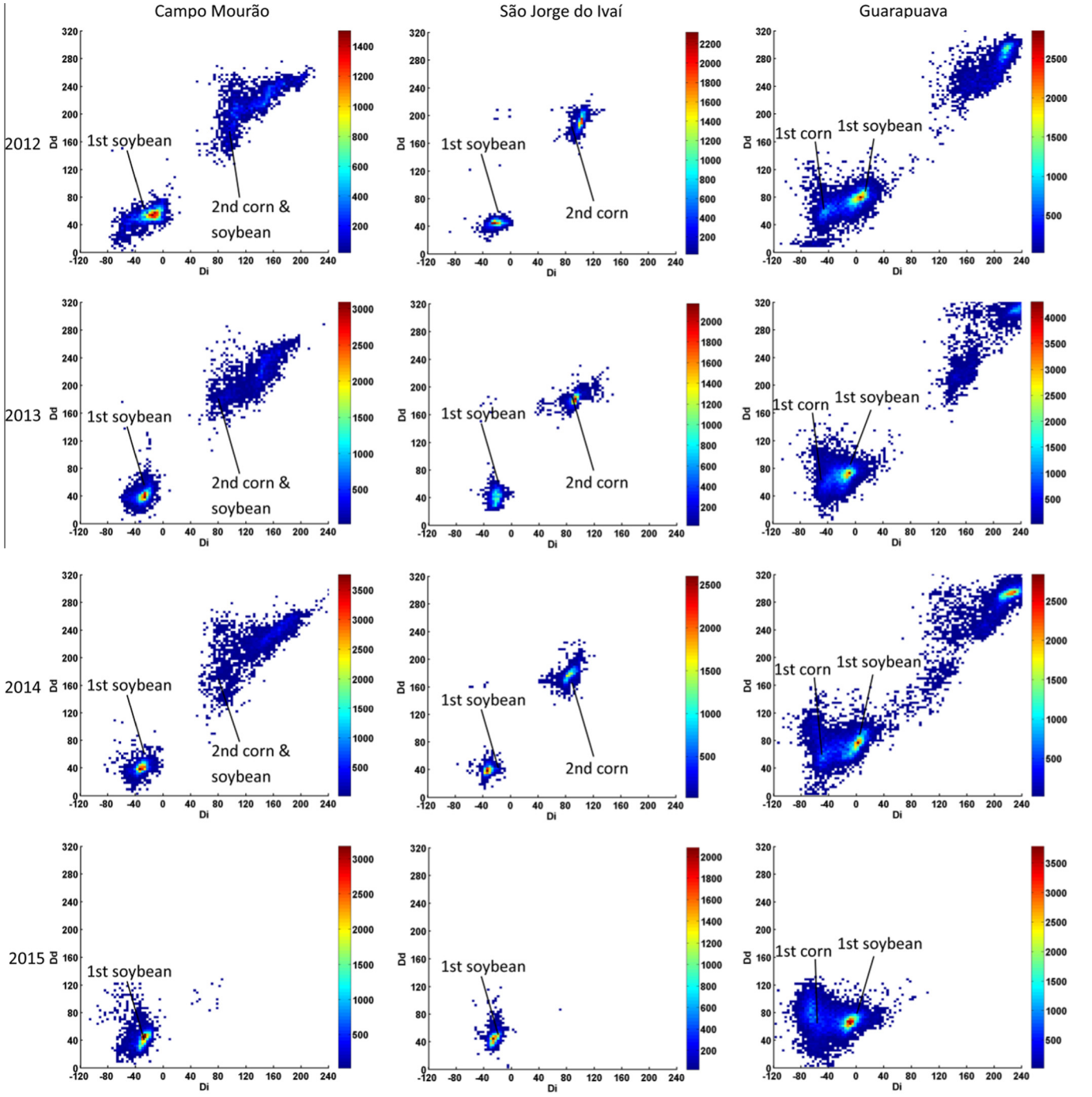

Fig. 3. Plots of variables show the process of developing classification rules that are consistent across regions and years. In this example, *phenological* variables D_i and D_d, which indicate period high biomass, are employed to create rules. Rows of plots are study years from 2012 to the first season of the crop year 2015. Columns are three municipalities selected to represent various cropping patterns. Clusters of soybean and corn are labeled on the plots according to the knowledge of local crop conditions. Color represents area in hectare. The complete set of plots in all municipalities and years is provided in Appendix A for conciseness. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.6.1. Growing cycles of field crops

Growing cycles consisting of increasing and decreasing slopes in EVI profiles are an indicator of the general vegetation phenology and type. Compared with forest and grassland, field crops like soybean, corn, wheat and barley have shorter growing cycles. For grassland, continuous rainfall may cause seasonal EVI peaks with similar length to field crops, but the magnitude of EVI variation is lower than cultivated land. Therefore, field crops were effectively extracted using *phenological* variables. Pixels of field crops are likely to have EVI cycles with higher p and q (sharper EVI slopes), higher V_a (greater EVI change in growing cycle), and lower G_{id} or G_{14} (shorter growing season) values. We used rules $p > 0.03$, $q > 0.02$, $V_a > 0.25$, and $70 < G_{14} < 180$ to identify field crop pixels.

In this study, we used loose rules based on *phenological* variables to reduce the influence of inter-annual or regional variability of growing season. Only threshold values that are stable for all years and all sub-areas were employed by the rules, making wide ranges of *phenological* variables. As a result, soybean and corn

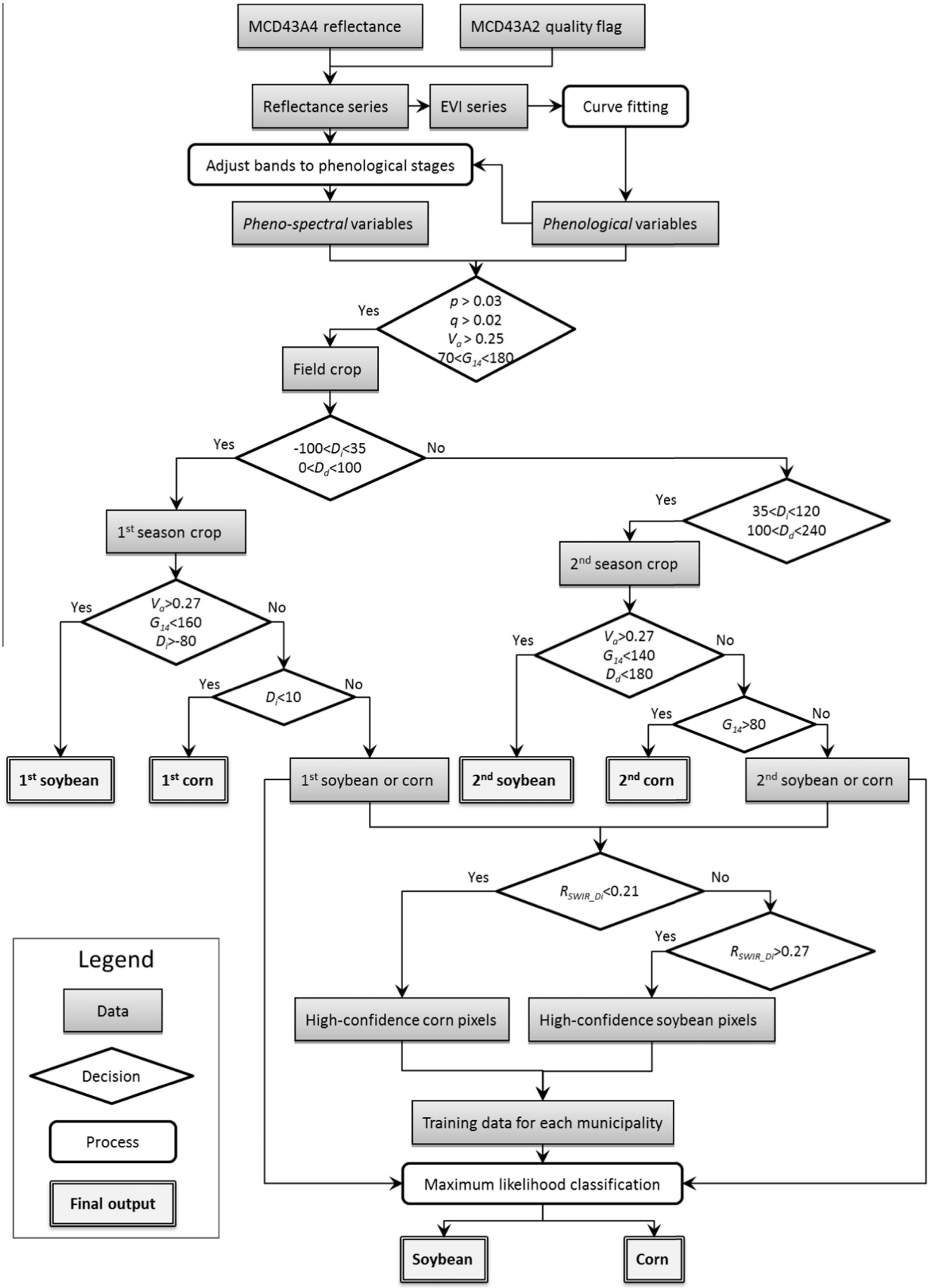

Fig. 4. The procedure of data processing and classification. MCD43A products were used to produce EVI time series, from which *phenological* variables were derived. *Pheno-spectral* variables were calculated by combining *phenological* variables and multi-spectral bands. *Phenological* variables were first used to develop decision rules to identify crop cycles and divide crop cycles into soybean, corn, and "soybean or corn". For the uncertain type "soybean or corn", high-confidence pixels were extracted based on rules on a *pheno-spectral* variable. These high-confidence pixels were employed as the training data of maximum likelihood classifiers for individual municipalities to separate soybean and corn further.

cannot be completely separated only based on phenology. In the study area, corn usually has earlier progress than soybean, but their phenology may not be separable. The difference in phenology between soybean and corn sometimes has less magnitude than the inter-annual or regional variability. If we applied very strict rules on *phenological* variables such as the dates of certain growing stages, these rules would possibly classify soybean and corn at a high accuracy for a year or a sub-area but yield poor results for another year or another sub-area with slightly different weather and crop progress. To develop an automated classifier, universal rules need to be established for all years by reviewing a large amount of data (Thenkabail et al., 2009; Baraldi, 2011; Dong et al., 2015), which drastically increase the difficulty compared to most other classification methods that train different classifiers for individual years.

2.6.2. Field crops in summer

In general, there are three cropping seasons of field crops in the study area. The first crop (first safra), or the main cropping season, is from about September to April of the next calendar year, which is the period with the best temperature and water conditions. The second crop (second safra, or safrinha) starts from mid-summer around February and could last until August. The third crop is mainly grown in winter, approximately from May to October. Single-, double- and triple-cropping are all observed in the study area, and multiple-cropping cycles tend to be shorter than single-cropping. Corn is only planted in the first and the second seasons around summer. Soybean is also mainly planted in the first and the second seasons, and third season soybean is very rare. Third season field crops are mostly winter grains like wheat, barley, and oat. Using *phenological* variables we focused on the first (primary) and the second crops to map soybean and corn, and winter crops were masked. We selected $-100 < D_i < 35$ and $0 < D_d < 100$ as the criteria for the first crop, and $35 < D_i < 120$ and $100 < D_d < 240$ for the second crop.

Soybean and corn are dominant field crops in the first and the second cropping seasons. Other crops like peanut and rice have very low coverage. In some areas, dry bean is planted in the first and the second cropping seasons and sometimes in the third season with similar phenology to soybean. Even in these areas soybean and corn still cover at least 90% of the cropland. Separating soybean and dry bean using an automated classifier is extremely challenging. We may confuse dry bean with soybean, but the influence is minimal because of the tiny area of dry bean.

2.6.3. First soybean

For the phenology-based rules that are selected in previous sections, threshold values were refined individually for each cropping season of soybean and corn. Within a cropping season, a field crop pixel might be identified as "soybean", "corn", or "soybean or corn". Because our rules based on *phenological* variables are quite loose with wide ranges, in this step a high proportion of pixels were identified as the undetermined type "soybean or corn", which were to be separated later.

Soybean is usually grown in large fields, and crop management is highly mechanized with great productivity. The peak EVI of soybean is generally higher than corn. For the first cropping season when weather conditions are relatively favorable, soybean pixels tend to have $V_a > 0.27$. Soybean is planted later than corn, and for soybean the lower limit of D_i was adjusted from -100 to -80. Compared to corn, the growing season of soybean is warmer and the length of the growing season is shorter, with $G_{14} < 160$. Therefore, among field crops in the first cropping season the following rules were used to further extract soybean pixels: $V_a > 0.27$, $G_{14} < 160$, and $D_i > -80$.

2.6.4. First corn

In the study area, the first corn is often planted as early as September or even August when the temperature is still relatively low. These early-planted corn fields are likely to have a long period of growing. By November corn planting is usually completed, and we used 10 as the upper limit of D_i for corn. For field crops in the first cropping season, the additional rule to identify first corn is: $D_i < 10$.

2.6.5. Second soybean

Soybean is mostly grown in the first season and second soybean is very rare, about 5% of the second crop area and less than 3% of the total soybean area. Second soybean cultivation is concentrated in the highly cropped areas of southwest and midwest municipalities. Similar to first soybean, second soybean is also unlikely to have V_a lower than 0.27. The harvest of second soybean is as early as June because soybean cultivation from June 15th to September 15th is prohibited to limit the spread of soybean rust. Hence, the growing season of second soybean is very short. Field crops in the second cropping season are possibly second soybean if: $V_a > 0.27$, $G_{14} < 140$, and $D_d < 180$.

2.6.6. Second corn

Corn does not have a "crop free" period as soybean, so the harvest of second corn could be as late as August or September. A growing cycle is unlikely to be second corn if its length is less than three months in a normal year. We used the rule of $G_{14} > 80$, which is a loose constraint, to extract pixels that are possibly second corn from second season field crops.

2.6.7. Separation of soybean and corn

The variability in planting and maturing dates within the area is substantial, depending on farm management factors such as market, weather, soil conditions, and seed availability. As a result, in the study area soybean and corn of the same season are not completely separable using universal rules on only phenological variables. In both the first and the second cropping seasons, a high proportion of cropping cycles would be assigned the undetermined type "soybean or corn" using the less restrictive rules from previous sections. This step is only applied for these cropping cycles which cannot be exactly identified merely based on phenology.

We used crop spectral properties at certain phenological stages, *pheno-spectral* metrics, to reduce the sensitivity of spectral signature to crop growth dynamic. Numerous combinations of spectral bands/indices and phenological stages were tested. We found that none of the combinations showed sufficient separability possibly because the 500 m resolution obscures the spectral difference between the two crops. The difference in the shortwave infrared band (MODIS band 6, 1628–1652 nm, referred to as R_{SWIR}) is the most stable over time and space. Shortwave infrared reflectance at 1400–1800 nm is very sensitive to canopy water content (Jacquemoud et al., 2009). Soybean generally has a lower water content in the growing season than corn according to previous measurement and simulation experiments (Anderson et al., 2004; Yilmaz et al., 2008). For the choice of phenological stages, we found that D_i was derived from the EVI time series by the curve-fitting method at a relatively high accuracy, because D_i is the date when the first derivative of the curve reaches its maximum. Also, at D_i the cropland EVI is already at a high level, usually around 0.4, so spectral signals of the field are mostly from crops rather than soil. Due to all these concerns above, we selected the MODIS band 6 shortwave infrared reflectance at D_i, referred to as R_{SWIR_Di}, as the *pheno-spectral* metric to improve the classification between soybean and corn in this study.

As soybean and corn green up, EVI and canopy water content increase rapidly while R_{SWIR} drops due to the strong absorption

of shortwave infrared radiation by water. In our multivariate plots, R_{SWIR_Di} of identified soybean clusters mostly varies from 0.21 to 0.35 depending on region and weather. Corn has higher water content, and R_{SWIR_Di} is about 0.15–0.27. Soybean and corn are not fully separable with R_{SWIR_Di}, but it is possible to extract some pure soybean and corn pixels at high confidence. For "soybean or corn" pixels, if R_{SWIR_Di} is less than 0.21 the corresponding pixels are unlikely to be soybean, and if R_{SWIR_Di} is greater than 0.27 the pixels are unlikely to be corn as far as we observed. In this way, high-confidence soybean and corn pixels were automatically retrieved for each municipality. We also have two assumptions on the statistical distribution: (i) *phenological* variables follow a normal distribution for each crop in a given municipality in a given year, and (ii) *phenological* and *pheno-spectral* variables are independent. The first assumption is likely to be valid because for a single municipality in a year the inter-annual and inter-regional variability of crop growth could be neglected. The second assumption is also acceptable because we isolated phenological and spectral measurements by adjusting image-specific spectral observations to a fixed phenological stage (D_i). Based on these assumptions, extreme pixels selected in the dimension of R_{SWIR_Di} still follow a normal distribution in the phenological dimensions. Thus, we performed maximum likelihood classification of soybean and corn for each municipality in each year. The training data are high-confidence soybean and corn pixels automatically extracted based on R_{SWIR_Di}, and variables used in the classification are *phenological* variables which reflect local crop phenology and general growing conditions. The process of automatic training data retrieval and classification was conducted for individual municipalities separately. The classification results of all municipalities were aggregated to produce the soybean and corn map of the entire state.

2.7. Validation

Validation was conducted at two levels of spatial resolution. First, the state of Paraná includes 399 administrative units at the municipal level with an average size of ~50,000 ha (Fig. 1). Statistics of these numerous units provide knowledge of the spatial distribution in the whole state as well as great temporal continuity from year to year, but the analysis was done for aggregated pixels (a municipality is about 45 by 45 pixels on average) at a coarser resolution. The second way of validation is at a finer resolution based on field level reference points. For the year of the reference data, accuracy assessment was performed based on a confusion matrix. This type of detailed validation is unavailable in other years due to the lack of reference data. Both sets of data for validation are independent of the classifier development process.

2.7.1. Comparison with official statistical data

Since 2002, the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística, IBGE) has been releasing agricultural production statistics at the municipal level every crop year (http://downloads.ibge.gov.br/downloads_estatisticas.htm). Statistics such as planted area, harvested area, yield, production and economic value are provided for individual crop commodities including soybean and corn. By our project time the latest available statistical data set is the crop year 2013 (checked in August 2015). We compared IBGE harvested area in crop years 2010–2013 to the municipal-level cropped area from our maps. Three sets of harvested area are available from IBGE data: soybean, first corn, and second corn. We also included the total area of soybean and corn (soybean + first corn + second corn) in the comparison. The coefficient of determination (R^2) was calculated for each item in each year to represent the goodness of the linear relationship between cropped areas from classification maps and statistical data. We also used Willmott's index of agreement (d) as a measurement of the agreement between the two sets of areas of individual municipalities:

$$d = 1 - \frac{\sum_{i=1}^{n}(A_{i,map} - A_{i,stat})^2}{\sum_{i=1}^{n}(|A_{i,map} - \overline{A_{stat}}| + |A_{i,stat} - \overline{A_{stat}}|)^2} \quad (2)$$

where $A_{i,map}$ and $A_{i,stat}$ are the cropped area of municipality i from classification maps and statistics, respectively; and $\overline{A_{stat}}$ is the mean value of statistics among all municipalities. The index of agreement measures how close the data points are to the 1:1 line, which varies between 0 (complete disagreement) and 1 (perfect agreement) (Willmott, 1981).

We also calculated the Root Mean Squared Error (RMSE) to express the magnitude of the error. Because municipalities differ in size, the relative measure of cropped percentage was used rather than the absolute cropped area. The cropped percentage is the ratio of cropped area over the total area of each municipality. The equation of RMSE is:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{i,map} - P_{i,stat})^2} \quad (3)$$

where $P_{i,map}$ and $P_{i,stat}$ are the cropped percentage of municipality i from classification maps and statistics, respectively. Classification error represented by RMSE has the same unit as the cropped percentage.

2.7.2. Accuracy assessment with a confusion matrix

Reference data were collected during the crop year 2012. Reference data points were mostly along main roads, and almost the entire state was covered. In total, 500 points from more than 100 municipalities were collected, and the land use type (soybean, corn, or other) was recorded (Fig. 1). The majority (304) of 384 soybean and corn points were from the first season, and 80 points were obtained during the early second season. The type "other" refers to all land cover types other than soybean and corn. In general, phenology-based rules easily identified uncultivated lands like natural vegetation and urban areas as shown in maps in the result section. To focus accuracy assessment on cropland and keep the numbers of the three types in the validation data set balanced, during the data collection along main roads we skipped lands that are unlikely to be cultivated. A confusion matrix was created to evaluate the goodness of the classification with overall accuracy, producer's accuracy, user's accuracy, etc.

3. Result

3.1. Classification maps

Crop maps in the study years were produced for the first and the second seasons respectively, shown in Fig. 5. The crop year 2015 only has the map of the first season. The spatial distribution of crops in the maps is consistent with the general patterns of land use in the state of Paraná. First season soybean followed by second season corn is the dominant double-cropping practice. Most large soybean and corn fields are concentrated in the plain area from the west to the north of the state, and fields are patchier in the south where the terrain is more complex. The distribution is coherent across study years, as a result of the combinations of climate, soil type, and water availability in the state.

3.2. Comparison with official statistical data

The coefficient of determination (R^2), index of agreement (d) and RMSE were used to indicate the agreement between the cropped area or the cropped percentage from classification maps and IBGE statistics (Table 2). The agreement for soybean and second season corn was high with R^2 greater than or around 0.9 and

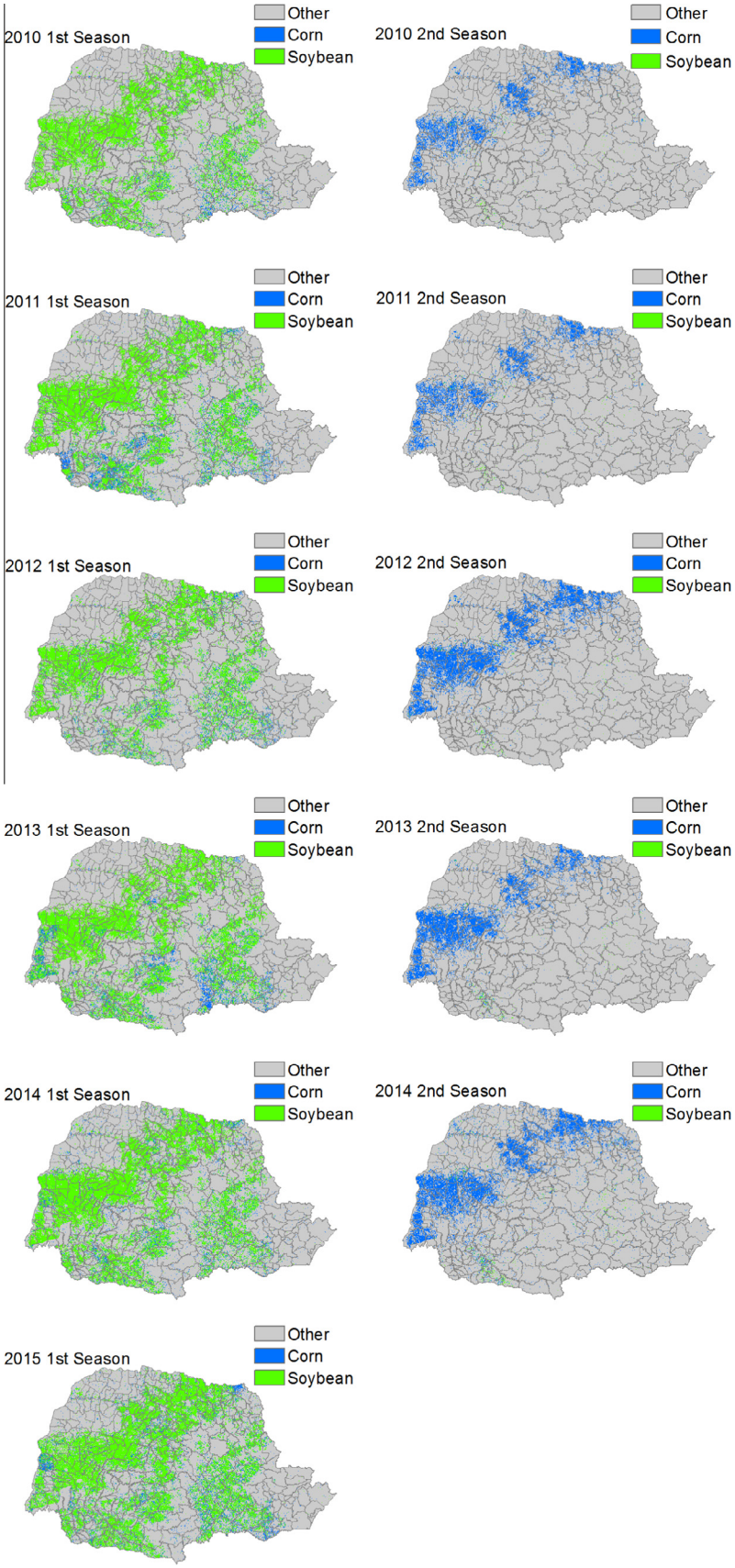

Fig. 5. Crop maps of soybean and corn in the study years produced by the automated classifier.

Table 2
Coefficients of determination (R^2) between the areas from classification maps and IBGE statistics.

Year		Soybean	First corn	Second corn	All
2010	R^2	0.958	0.695	0.930	0.954
	d	0.977	0.861	0.977	0.979
	RMSE	7.8%	4.7%	5.7%	12.3%
2011	R^2	0.946	0.530	0.947	0.956
	d	0.969	0.753	0.971	0.969
	RMSE	9.9%	7.3%	6.7%	14.3%
2012	R^2	0.938	0.687	0.965	0.954
	d	0.972	0.863	0.983	0.976
	RMSE	10.7%	5.0%	6.6%	16.0%
2013	R^2	0.947	0.504	0.891	0.951
	d	0.968	0.737	0.964	0.975
	RMSE	9.4%	7.6%	8.8%	16.2%

d greater than 0.96. The R^2 of first season corn was lower, from ~0.5 to ~0.7, indicating moderate agreement. The R^2 and d values were high for the total area of soybean and corn, because classification errors mostly occurred between soybean and corn, and the first season corn, the most problematic class, only has a relatively small area. RMSE values suggest that at the municipal level soybean area might be mapped with 7.8–10.7% error. The relative error is smaller for first corn (4.7–7.6%) and second corn (5.7–8.8%) because the cropped percentage of corn is generally lower than soybean in the state.

Total areas of the three crops were compared between official statistics and maps (Fig. 6). Soybean was mostly over-estimated. This is consistent with our observations that in the vast, flat, and highly mechanized soybean fields, non-crop lands at field edges tend to be classified as soybean. Examples will be given in the discussion section. 2012 is an exception possibly because crop growth was impacted by the drought of La Nina. Some soybean fields were not able to complete the growth cycle normally, resulting in irregular time series that were not identified by the automated algorithm. 2011 was another La Nina year, but the influence on seasonal precipitation was mild.

3.3. Confusion matrix

Accuracy assessment of soybean and corn maps was conducted using the reference data, and the results indicated that the 2012 map had a high accuracy of 87.2% and Kappa coefficient of 0.804 (Table 3). The accuracy suggests that the classification map agrees well with the reference data, and maps in other years from the same automated classifier are likely to be similarly reliable. It should be noticed that the accuracy mostly represents the classification result of large cropland parcels because these parcels have more chances to be observed and recorded along major roads. Small and patchy cropland parcels in complex terrain present a challenge to cropland mapping at the 500 m resolution, for which we have relatively less knowledge. In the reference dataset, many small soybean and corn fields were not successfully identified because of mixed pixels, which is one of the main reasons for the confusion. Another main source of classification error was corn fields that were misclassified as soybean. Such confusions possibly resulted from the relatively high shortwave infrared reflectance of these corn fields caused by water stress or neighboring fields.

4. Discussion

4.1. Automated classification rules

The study is the first to apply an automated classification method to the mapping of soybean and corn in subtropical and temperate South America. The greatest advantage of the automated algorithm is the ability to map croplands rapidly at low cost on independent data sources year by year. When completed, the automated algorithm could be run repeatedly over years without further need for re-training or re-calibration, which makes the routine mapping process very economical. The algorithm in this paper is ready for immediate implementation, requiring no additional data or efforts. The data processing time is very short when passing all pixels through the simple-structure decision tree and other steps. In addition, there is no need to collect training data every year, making it very timely and labor-efficient for a large area. When transferred to a new region, classification rules need to be changed, but the existing decision tree structure may still be useful and rules may only need some slight and comprehensible adjustments especially when the two areas have similar crop phenology.

In the development stage, the automated method requires a long time and intensive efforts spent on building classification rules. Expert knowledge and experience on local agricultural practices are employed to isolate the special physical characteristics of

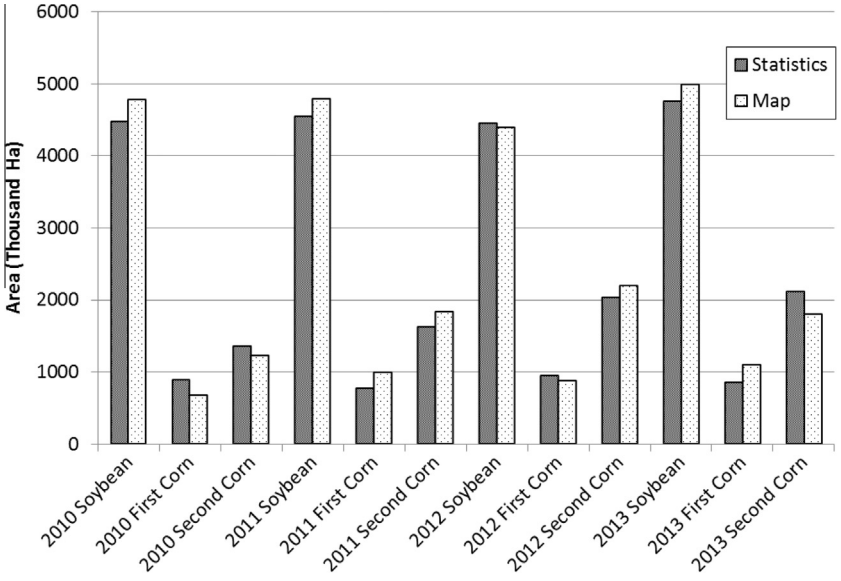

Fig. 6. The total area of soybean, first corn and second corn from official statistics and maps produced by the automated approach.

Table 3
The confusion matrix of the 2012 crop map validated by reference data.

Reference class	Map class			Total	Producer's accuracy
	Other	Corn	Soybean		
Other	111	2	3	116	95.7%
Corn	11	136	27	174	78.2%
Soybean	15	6	189	210	90.0%
Total	137	144	219	500	
User's accuracy	81.0%	94.4%	86.3%		
			Overall accuracy		87.2%
			Kappa coefficient		0.804

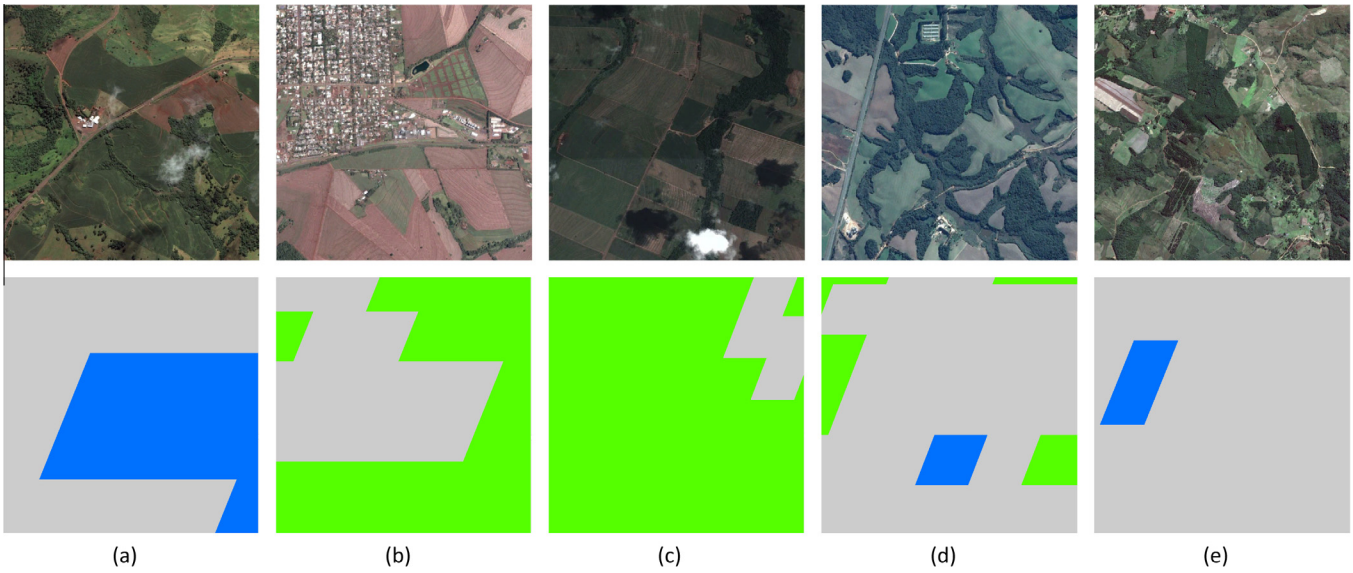

Fig. 7. Examples of mixed pixels and classification results. The upper row is high-resolution images from Google Earth and the lower row includes the corresponding classification maps. The green color is soybean, blue is corn, and gray is other. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

land features from the inter-annual and inter-regional variability of spectral signals. To improve the efficiency of algorithm development, we derived phenology-related variables from original image time series. Phenological properties from multi-temporal remote sensing are closely connected to local crop calendars and can be utilized to build rules that are consistent across years. Spectral properties at certain phenological stages are independent of imaging dates and can be used to reveal the inherent optical nature of a specific crop for effective identification.

Independent of the year, the automated method can construct a long and consistent historical record of cropland use based on continuous missions like MODIS, Landsat, etc. Traditionally the accuracy of a land use classification series is likely to be variable and affected by (i) classifier, (ii) reference data, and (iii) availability of high-quality data. The automated classification method eliminates the variability from different classifiers and reference data sets. The advanced methods of image preprocessing and quality assessment, such as the model inversion method by MCD43 and the LEPDAPS algorithm for Landsat imagery (Masek et al., 2006), have greatly reduced the temporal inconsistency of the image time series. The rule-based automated classifier, built mainly based on knowledge rather than machine-learning methods, only uses variables with known physical meaning, and thus reduces the influence of overfitting or a biased training set. As new platforms and new image records keep emerging today, the automated classification method is advantageous for the production of a consistent classification map series, which facilitates the quantitative analysis of cropland dynamics.

4.2. Issues involving classification accuracy

At the 500 m spatial resolution of MCD43A4, the effect of mixed pixels plays an important role in classification accuracy. Based on the comparison between classification maps and reference data, we found that classification bias (overestimation or underestimation) highly depended on the diverse patterns of pixel mixing across regions. In heavily-cropped regions, many mixed pixels are dominated by crops and were identified as cropland due to the strong vegetation signals of crops, causing over-estimation of crop area. For example, buildings and trees in the center of Fig. 7 (a), residential areas and grassland on the lower-right part of Fig. 7(b), and the forest along a creek on the upper-left part of Fig. 7(c), were all included in the continuous cropland in the corresponding classification maps. This type of confusion mostly occurs on the flat land in the north and west of the state, where soybean and corn fields are large and the environment is suitable for mechanized agricultural production and transport. By contrast, in hilly areas soybean and corn fields tend to be small and scattered, surrounded by dense natural vegetation like forests or tall bushes with a much longer leafy season. Croplands adjacent to natural vegetation will show strong vegetation signals during the non-growing season due to mixed pixel effect, resulting in flatter

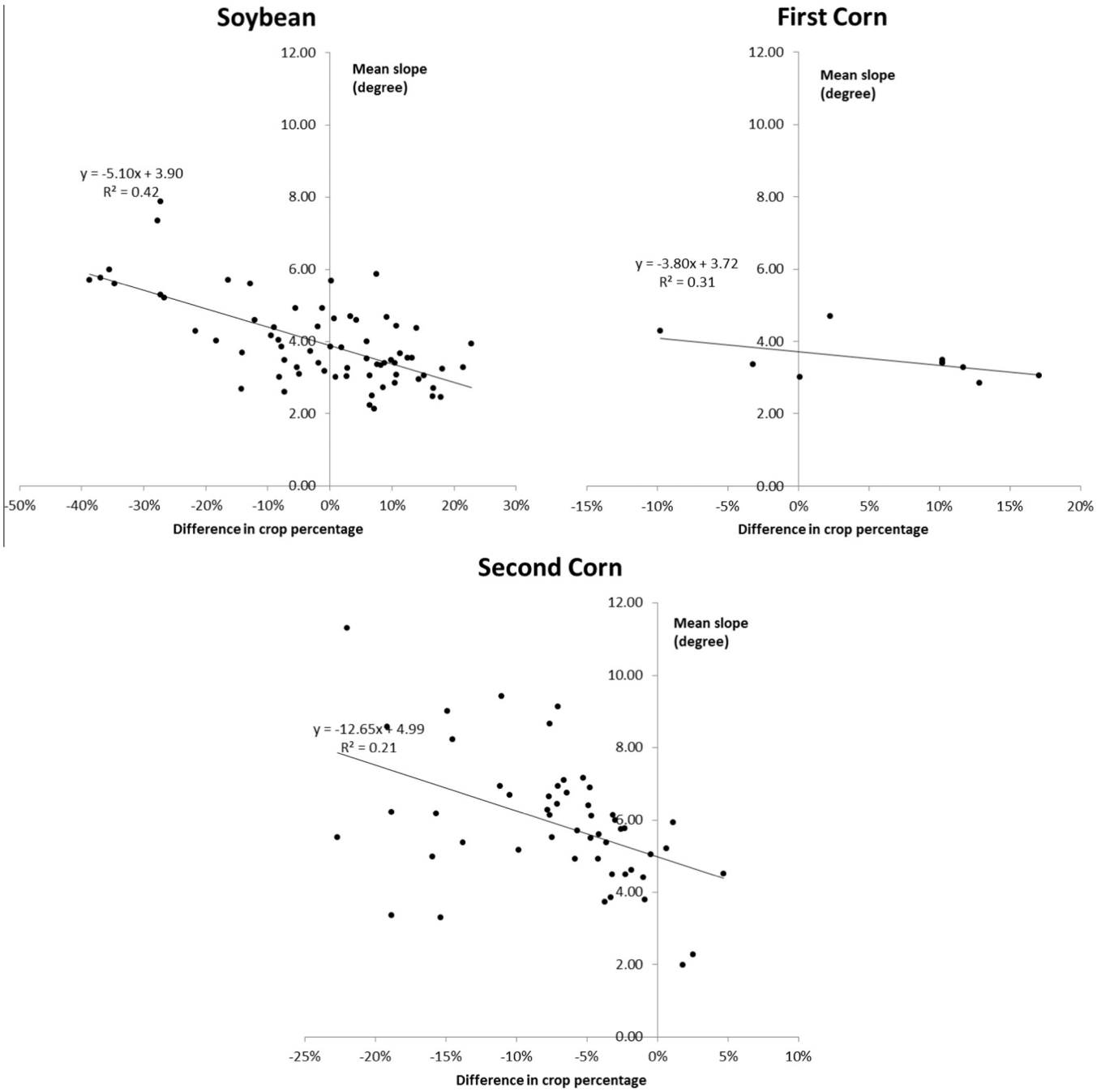

Fig. 8. The relationship between classification bias and the mean slope of individual municipalities. Classification bias is quantified as the crop cover percentage from maps minus that from statistics. Only municipalities with a relatively high cover percentage of the corresponding crop type are included. In general, the higher the mean slope (more complex terrain), the higher chance to under-estimate the crop percentage, and *vice versa*.

temporal profiles of EVI which would not be identified as crops by the automated algorithm. In Fig. 7(d), the high proportion of forests obscures the temporal and spectral signal of small soybean fields, making them unclassified or misclassified as corn (the pixel in the lower-middle part). Similarly, in Fig. 7(e), corn fields that scatter among forests were mostly unidentified except a few larger ones on the left side of the image.

Because of the 500 resolution and the scale of homogeneous land cover in the study area, most pixels are mixed to some extent. Usually the temporal and spectral properties of the mixed pixel are determined by the dominant component, but the pattern of mixing is unclear when components have similar proportions, calling for further quantitative analysis. When applying our automated classification approach, we observed that the contribution of dense natural vegetation to mixed pixels outweighs that of crops, in particular for areas with complex terrain where crop production is not optimal. In hilly areas dense natural vegetation has relatively strong vegetation signals, and sometimes even a small proportion of natural vegetation will completely alter the temporal profiles of crop EVI. This is similar to the phenomenon found by Bolton and Friedl (2013), that the small presence of sub-pixel natural vegetation would affect the detection of soybean and corn phenology in Wisconsin, US. This also possibly explains why first corn has a lower agreement between mapping results and official statistics

than soybean and second corn. In the first season when the crop productivity is the highest, preference is given to soybean cultivation because of the higher economic profit. Corn then becomes a good choice in the second season to prevent soybean rust, making soybean-corn the most common rotation in the state. As a result, in the first season corn is mostly grown in more hilly areas with less favorable conditions, generating many small and patchy corn fields which are affected more by mixed pixels than other crops. Therefore, in areas with scattered cropland, our classification maps are able to show the distribution of crops while the total cropland area should be used with caution.

Because of the difference in pixel mixing between plain and hilly areas, classification bias is affected by terrain complexity. In our analysis shown in Fig. 8, terrain complexity is represented by the mean slope of each municipality. Classification bias was calculated as the difference in crop cover percentage between maps and official statistics. Municipalities with a very low crop cover were excluded from the analysis to avoid the influence of other factors. These plots show a moderate relationship that over-estimation is more likely to occur in flatter areas, which is consistent with our observations.

In Brazil, the state of Paraná and other states in the south have a relatively long history of reclamation and agricultural production. Soybean and corn production is concentrated in flat areas as almost continuous cropland. In places under rapid agricultural land conversion like the Amazon, the pattern of pixel mixing may differ. According to the law, 65% of the native land is allowed to be converted to field crops in the cerrado (tropical savanna) ecosystem and 20% in the Amazon ecosystem, and the remaining land should be preserved. Hence, in newly-converted croplands, crops and natural vegetation are likely to be mixed. The effect of mixed pixels may play a different role in the mapping of tropical states like Mato Grosso.

5. Conclusion

In this study, we have presented a new automated approach to map soybean and corn in the state of Paraná, which is the second largest production state of Brazil and a good representative of temperate and subtropical cropland over the broad area near the boundary of Brazil, Argentina, Uruguay, and Paraguay. The most attractive feature of the automated approach is that it reduces the need to collect additional reference data or re-train the algorithm, greatly lowering the cost of cropland mapping. This is especially important for South America where reference data are often too scarce to routinely apply the supervised classification method. Variables related to phenology and spectral features at certain phenological stages were utilized as measurements that reflect the nature of crop types and remain stable over time and space. A simple and robust classification algorithm was developed to take into account prior knowledge on crop calendar and spectral properties of crop canopy. Resultant crop maps presented consistent patterns for the spatial distribution of soybean and corn across multiple years. The cropped area from the maps has a higher agreement for soybean than for corn when compared to the official statistics.

The success of the initial application in South America using the automated approach is encouraging, given the potential of extending the algorithm to other crop types and other remotely sensed data. To identify more crop types in other areas, expert knowledge on local agricultural practices, crop growth modeling, and crop spectral monitoring and simulation (Jacquemoud et al., 2009) are the main factors to consider when defining classification rules. For regions with smaller crop parcels, the extended automated approach may improve the spatial resolution by incorporating images from new platforms. Sensors onboard existing and planned satellites like Landsat and Sentinel-2 offer similar optical bands to MODIS but at a much higher spatial resolution. Because the automated algorithm is not image-specific, it is scalable to utilize these datasets with minimal revisions. Although the image time series from a single platform are usually not dense enough to retrieve phenology, the combined use of images from multiple satellites may increase the imaging frequency and improve the spatial resolution of cropland products generated by the automated method. We are confident that the automated method will have more potential applications as the technique of data retrieval and processing advances.

Acknowledgements

This research was partially supported by the National Natural Science Funds of China (grant number: 41301445) and a research grant from Tsinghua University (grant number: 20151080351).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.isprsjprs.2016.05.014.

References

Anderson, M.C., Neale, C.M.U., Li, F., Norman, J.M., Kustas, W.P., Jayanthi, H., et al., 2004. Upscaling ground observations of vegetation water content, canopy height, and leaf area index during SMEX02 using aircraft and Landsat imagery. Remote Sens. Environ. 92, 447–464.

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., Weber, J., 2008. GLOBCOVER: The Most Detailed Portrait of Earth. Esa Bulletin-European Space Agency, pp. 24–31.

Arvor, D., Meirelles, M., Dubreuil, V., Bégué, A., Shimabukuro, Y.E., 2012. Analyzing the agricultural transition in Mato Grosso, Brazil, using satellite-derived indices. Appl. Geogr. 32, 702–713.

Arvor, D., Jonathan, M., Meirelles, M.S.P., Dubreuil, V., Durieux, L., 2011. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. Int. J. Remote Sens. 32, 7847–7871.

Badhwar, G.D., 1984. Automatic corn soybean classification using landsat MSS data. 1. Near-harvest crop proportion estimation. Remote Sens. Environ. 14, 15–29.

Baraldi, A., 2011. Satellite image automatic Mapper™ (SIAM™) – a Turnkey software executable for automatic near real-time multi-sensor multi-resolution spectral rule-based preliminary classification of spaceborne multi-spectral images. Recent Pat. Space Technol. 1, 81–106.

Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 173, 74–84.

Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. Geocarto Int. 26, 341–358.

Brown, J.C., Kastens, J.H., Coutinho, A.C., Victoria, D.d.C., Bishop, C.R., 2013. Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data. Remote Sens. Environ. 130, 39–50.

Chavas, J.P., Holt, M.T., 1996. Economic behavior under uncertainty: a joint analysis of risk preferences and technology. Rev. Econ. Stat. 78, 329–335.

Conrad, C., Fritsch, S., Zeidler, J., Rücker, G., Dech, S., 2010. Per-field irrigated crop classification in arid central Asia using SPOT and ASTER data. Remote Sens. 2, 1035–1056.

DeFries, R.S., Foley, J.A., Asner, G.P., 2004. Land-use choices: balancing human needs and ecosystem function. Front. Ecol. Environ. 2, 249–257.

Dong, J., Xiao, X., Kou, W., Qin, Y., Zhang, G., Li, L., et al., 2015. Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms. Remote Sens. Environ. 160, 99–113.

dos Santos, J.S., Fontana, D.C., Silva, T.S.F., Rudorff, B.F.T., 2014. Identification of the spatial and temporal dynamics for estimating soybean crop area from MODIS images in the Rio Grande do Sul, Brazil. Rev. Bras. Engenharia Agric. E Ambiental 18, 54–63.

Drinkwater, L.E., Wagoner, P., Sarrantonio, M., 1998. Legume-based cropping systems have reduced carbon and nitrogen losses. Nature 396, 262–265.

Epiphanio, R.D.V., Formaggio, A.R., Rudorff, B.F.T., Maeda, E.E., Luiz, A.J.B., 2010. Estimating soybean crop areas using spectral-temporal surfaces derived from MODIS images in Mato Grosso, Brazil. Pesqui. Agropecu. Bras. 45, 72–80.

Evans, J.P., Geerken, R., 2006. Classifying rangeland vegetation type and coverage using a Fourier component based similarity measure. Remote Sens. Environ. 105, 1–8.

Friedl, M.A., Brodley, C.E., Strahler, A.H., 1999. Maximizing land cover classification accuracies produced by decision trees at continental to global scales. IEEE Trans. Geosci. Remote Sens. 37, 969–977.
Friedl, M., McIver, D., Hodges, J., Zhang, X., Muchoney, D., Strahler, A., et al., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sens. Environ. 83, 287–302.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., et al., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182.
Geerken, R.A., 2009. An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change. ISPRS J. Photogramm. Remote Sens. 64, 422–431.
Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., et al., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34, 2607–2654.
Gusso, A., Formaggio, A.R., Rizzi, R., Adami, M., Rudorff, B.F.T., 2012. Soybean crop area estimation by Modis/Evi data. Pesq. Agropecu. Bras. 47, 425–435.
Hill, J., Nelson, E., Tilman, D., Polasky, S., Tiffany, D., 2006. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. Proc. Natl. Acad. Sci. USA 103, 11206–11210.
Howard, D.M., Wylie, B.K., 2014. Annual crop type classification of the US Great Plains for 2000 to 2011. Photogramm. Eng. Remote Sens. 80, 537–549.
Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., Ferreira, L., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.
Huete, A.R., Liu, H.Q., Batchily, K., van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sens. Environ. 59, 440–451.
Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., et al., 2009. PROSPECT + SAIL models: a review of use for vegetation characterization. Remote Sens. Environ. 113 (Suppl. 1), S56–S66.
Johnson, D.M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens. Environ. 141, 116–128.
Knight, J.F., Lunetta, R.S., Ediriwickrema, J., Khorram, S., 2006. Regional scale land cover characterization using MODIS-NDVI 250 m multi-temporal imagery: a phenology-based approach. GISci. Remote Sens. 43, 1–23.
Lambin, E.F., Meyfroidt, P., 2011. Global land use change, economic globalization, and the looming land scarcity. Proc. Natl. Acad. Sci. USA 108, 3465–3472.
Lloyd, D., 1990. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. Int. J. Remote Sens. 11, 2269–2279.
Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens. Environ. 164, 324–333.
Lucas, R., Rowlands, A., Brown, A., Keyworth, S., Bunting, P., 2007. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. ISPRS J. Photogramm. Remote Sens. 62, 165–185.
Lucht, W., Schaaf, C.B., Strahler, A.H., 2000. An algorithm for the retrieval of albedo from space using semiempirical BRDF models. IEEE Trans. Geosci. Remote Sens. 38, 977–998.
Macdonald, R.B., Hall, F.G., 1980. Global crop forecasting. Science (New York, N.Y.) 208, 670–679.
Masek, J., Vermote, E., Saleous, N., Wolfe, R., Hall, F., Huemmrich, K., et al., 2006. A Landsat surface reflectance dataset for North America, 1990–2000 RID F-3944-2010. IEEE Geosci. Remote Sens. Lett. 3, 68–72.
Mercante, E., de Lima, L.E.P., Justina, D.D.D., Uribe-Opazo, M.A., Lamparelli, R.A.C., 2012. Detection of soybean planted areas through orbital images based on culture spectral dynamics. Engenharia Agricola 32, 919–930.
Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochem. Cycles 22, GB1022.
Mosier, A.R., Halvorson, A.D., Reule, C.A., Liu, X.J., 2006. Net global warming potential and greenhouse gas intensity in irrigated cropping systems in northeastern Colorado. J. Environ. Qual. 35, 1584–1598.
Ozdogan, M., Gutman, G., 2008. A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: an application example in the continental US. Remote Sens. Environ. 112, 3520–3537.
Pena-Barragan, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens. Environ. 115, 1301–1316.
Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V., Justice, C.O., 2010. Estimating global cropland extent with multi-year MODIS data. Remote Sens. 2, 1844–1863.
Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000? Global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. Global Biogeochem. Cycles 24, GB1011.
Qaim, M., Traxler, G., 2005. Roundup ready soybeans in Argentina: farm level and aggregate welfare effects. Agric. Econ. 32, 73–86.
Qin, Y., Xiao, X., Dong, J., Zhou, Y., Zhu, Z., Zhang, G., et al., 2015. Mapping paddy rice planting area in cold temperate climate region through analysis of time series Landsat 8 (OLI), Landsat 7 (ETM+) and MODIS imagery. ISPRS J. Photogramm. Remote Sens. 105, 220–233.
Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Global Biogeochem. Cycles 22, GB1003.
Reis, S., Taşdemir, K., 2011. Identification of hazelnut fields using spectral and Gabor textural features. ISPRS J. Photogramm. Remote Sens. 66, 652–661.
Rounsevell, M.D.A., Annetts, J.E., Audsley, E., Mayr, T., Reginster, I., 2003. Modelling the spatial distribution of agricultural land use at the regional scale. Agric. Ecosyst. Environ. 95, 465–479.
Sakamoto, T., Gitelson, A.A., Arkebauer, T.J., 2014. Near real-time prediction of U.S. corn yields based on time-series MODIS data. Remote Sens. Environ. 147, 219–231.
Sakamoto, T., Van Cao, Phung, Van Nguyen, Nhan, Kotera, A., Yokozawa, M., 2009. Agro-ecological interpretation of rice cropping systems in flood-prone areas using MODIS imagery. Photogramm. Eng. Remote Sens. 75, 413–424.
Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., et al., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83, 135–148.
Siachalou, S., Mallinis, G., Tsakiri-Strati, M., 2015. A hidden Markov models approach for crop classification: linking crop phenology to time series of multi-sensor remote sensing data. Remote Sens. 7.
Simonneaux, V., Francois, P., 2003. Identifying main crop classes in an irrigated area using high resolution image time series. Proceedings of the International IEEE Geoscience and Remote Sensing Symposium (IGARSS'03), 21–25 July, vol. 1. Toulouse, France, pp. 252–254.
Simonneaux, V., Duchemin, B., Helson, D., Er-Raki, S., Olioso, A., Chehbouni, A.G., 2008. The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. Int. J. Remote Sens. 29, 95–116.
Soares, B.S., Nepstad, D.C., Curran, L.M., Cerqueira, G.C., Garcia, R.A., Ramos, C.A., et al., 2006. Modelling conservation in the Amazon basin. Nature 440, 520–523.
Son, N., Chen, C., Chen, C., Duc, H., Chang, L., 2014. A phenology-based classification of time-series MODIS data for rice crop monitoring in Mekong Delta, Vietnam. Remote Sens. 6.
Soudani, K., le Maire, G., Dufrene, E., Francois, C., Delpierre, N., Ulrich, E., et al., 2008. Evaluation of the onset of green-up in temperate deciduous broadleaf forests derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. Remote Sens. Environ. 112, 2643–2655.
Sugawara, L.M., Theodor Rudorff, B.F., Adami, M., 2008. Feasibility of the use of Landsat imagery to map soybean crop areas in Parana, Brazil. Pesq. Agropecu. Bras. 43, 1777–1783.
Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., et al., 2009. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. Int. J. Remote Sens. 30, 3679–3733.
Turner II, B.L., Lambin, E.F., Reenberg, A., 2007. The emergence of land change science for global environmental change and sustainability. Proc. Natl. Acad. Sci. USA 104, 20666–20671.
Vieira, C., Mather, P., McCullagh, M., 2000. The Spectral-Temporal Response Surface and its use in the multi-sensor, multi-temporal classification of agricultural crops. Int. Arch. Photogramm. Remote Sens. 33, 582–589.
Wachholz de Souza, C.H., Mercante, E., Prudente, V.H.R., Justina, D.D.D., 2013. Methods of performance evaluation for the supervised classification of satellite imagery in determining land cover classes. Cien. Inv. Agr. 40, 419–428.
Waldner, F., Canto, G.S., Defourny, P., 2015. Automated annual cropland mapping using knowledge-based temporal features. ISPRS J. Photogramm. Remote Sens. 110, 1–13.
Wang, C., Zhong, C., Yang, Z., 2014. Assessing bioenergy-driven agricultural land use change and biomass quantities in the US Midwest with MODIS time series. J. Appl. Remote Sens. 8, 085198.
Wardlow, B.D., Egbert, S.L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: an assessment for the US Central Great Plains. Remote Sens. Environ. 112, 1096–1116.
Wardlow, B.D., Egbert, S.L., Kastens, J.H., 2007. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. Remote Sens. Environ. 108, 290–310.
Willmott, C.J., 1981. On the validation of models. Phys. Geogr. 2, 184–194.
Wu, B., Meng, J., Li, Q., Yan, N., Du, X., Zhang, M., 2014. Remote sensing-based global crop monitoring: experiences with China's CropWatch system. Int. J. Digit. Earth 7, 113–137.
Wu, Z., Thenkabail, P.S., Verdin, J.P., 2014. Automated cropland classification algorithm (ACCA) for California using multi-sensor remote sensing. Photogramm. Eng. Remote Sens. 80, 81–90.
Xin, Q., Gong, P., Yu, C., Yu, L., Broich, M., Suyker, A.E., et al., 2013. A production efficiency model-based method for satellite estimates of corn and soybean yields in the midwestern US. Remote Sens. 5, 5926–5943.
Yilmaz, M.T., Hunt Jr., E.R., Jackson, T.J., 2008. Remote sensing of vegetation water content from equivalent water thickness using satellite imagery. Remote Sens. Environ. 112, 2514–2522.
Yu, L., Wang, J., Clinton, N., Xin, Q., Zhong, L., Chen, Y., et al., 2013. FROM-GC: 30 m global cropland extent derived through multisource data integration. Int. J. Digit. Earth 6, 521–533.
Zhang, G., Xiao, X., Dong, J., Kou, W., Jin, C., Qin, Y., et al., 2015. Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data. ISPRS J. Photogramm. Remote Sens. 106, 157–171.
Zhong, L., Gong, P., Biging, G.S., 2014. Efficient corn and soybean mapping with temporal extendability: a multi-year experiment using Landsat imagery. Remote Sens. Environ. 140, 1–13.
Zhong, L., Gong, P., Biging, G.S., 2012. Phenology-based crop classification algorithm and its implications on agricultural water use assessments in California's Central Valley. Photogramm. Eng. Remote Sens. 78, 799–813.
Zhong, L., Hawkins, T., Biging, G., Gong, P., 2011. A phenology-based approach to map crop types in the San Joaquin Valley, California. Int. J. Remote Sens. 32, 7777–7804.