Geocarto International

ISSN: 1010-6049 (Print) 1752-0762 (Online) Journal homepage: https://www.tandfonline.com/loi/tgei20

Wheat planted area detection from the MODIS NDVI time series classification using the nearest neighbor method calculated by the Euclidean distance and cosine similarity measures

Miriam Rodrigues da Silva, Osmar Abílio de Carvalho Junior, Renato Fontes Guimarães, Roberto Arnaldo Trancoso Gomes & Cristiano Rosa Silva

To cite this article: Miriam Rodrigues da Silva, Osmar Abílio de Carvalho Junior, Renato Fontes Guimarães, Roberto Arnaldo Trancoso Gomes & Cristiano Rosa Silva (2019): Wheat planted area detection from the MODIS NDVI time series classification using the nearest neighbor method calculated by the Euclidean distance and cosine similarity measures, Geocarto International, DOI: 10.1080/10106049.2019.1581266

To link to this article: https://doi.org/10.1080/10106049.2019.1581266

Wheat planted area detection from the MODIS NDVI time series classification using the nearest neighbor method calculated by the Euclidean distance and cosine similarity measures

Miriam Rodrigues da Silva, Osmar Abílio de Carvalho Junior*, Renato Fontes Guimarães, Roberto Arnaldo Trancoso Gomes & Cristiano Rosa Silva

Department of Geography, University of Brasilia, Brasília, Brazil. Emails:

*corresponding author: Department of Geography, University of Brasilia, Campus Universitário Darcy Ribeiro, Brasília-DF, Postal Code: 70910-900 Brazil, email: osmarjr@unb.br

Wheat planted area detection from the MODIS NDVI time series classification using the nearest neighbor method calculated by the Euclidean distance and cosine similarity measures

This research aims to detect the wheat crop in the Northwest region of Rio Grande do Sul (Brazil) using MODIS NDVI time series. Detection of wheat crops presents two difficulties: (a) high variation of wheat phenological curves due to climatic fluctuations during the crop cycle; and (b) the plantations are in an environment with different types of rural and urban targets. In solving these problems, we propose a classification based on the nearest neighbor (a specific case of the K-NN method) from the similarity and distance metrics combined with the determination of the best threshold value to individualize the wheat mask. The nearest neighbor classification using minimum distance (Kappa of 0.75) obtained a result equivalent to that of cosine similarity (Kappa of 0.74) as attested by the McNemar test. The planted area result was comparable to official statistics from the Brazilian Institute of Geography and Statistics obtained through direct interviews.

Keywords: Agriculture, K-Nearest Neighbors, Optimal Threshold, Phenology, Brazil

Introduction

Wheat is the most cultivated cereal in the world, occupying an annual planted area of more than 220 million hectares in diverse climatic conditions and geographic regions (Shiferaw et al. 2013). Also, wheat has high relevance in the world's nutrition due to its quality and quantity of proteins and variety of derived products (Sun et al. 2012), being responsible for about 20% of the total dietary calories and proteins worldwide (Shiferaw et al. 2013). Atchison et al. (2010) examined 10,235 supermarket items in Australia and found that 20.7 percent contained wheat, with 1977 items in food and 144 items in non-food items, while considering only the food products the presence of wheat increased to about 29.5%. This product leads world trade in agricultural commodities and is the leading food import product of developing nations (Dixon et al. 2009). During the last century, wheat productivity increased from the improvement of several factors: wheat genetics, mechanization, pest and disease control, and agricultural management practices (Dixon et al. 2006). However, global demand for wheat tends to increase significantly due to the growth of urban populations that consume more animal products and choose wheat-based products that are easy to prepare (Pingali 2007; Lobell and Burke 2010). Moreover, a future concern is that global warming may represent a fall in world wheat production by 6% for each degree Celsius of temperature rise (Asseng et al. 2014). Consequently, future food security depends on the continuity of productivity gains, addressing the threats of climate change, and establishing a monitoring system of production and demand.

In Brazil, the main winter crop is wheat with an annual production of more than six million tons distributed in the South, Southeast, and Central regions. However, Brazil is one of the world's largest importers of wheat, importing about half of its consumption. Traditionally, the primary supplier of wheat to Brazil is Argentina. The

core wheat-growing region in Southern Brazil, concentrating 92% of the national production, where the State of Paraná is the most abundant producer (51% of the national production) followed by the Rio Grande do Sul (37% of the national production) (IBGE, 2016). The concentration of wheat cultivation in Southern Brazil is due to the cold temperature and the development of cultivars adapted to this region (Canziani and Guimarães 2009).

The provision of reliable information on agricultural production is crucial for the world's policy-makers and food security at the national and global level. The traditional acquisition of information on agricultural production is carried out through the national census. However, the agricultural census data from household surveys (in situ) is expensive and requires a great deal of fieldwork and time. Furthermore, this approach generates data that contain aggregated information about plantation types and their areas, not allowing a detailed description of the spatial and temporal distribution of crops. Hence, there is a consensus that the use of remote sensing techniques enables an improvement in the acquisition of agricultural data, establishing an accurate spatiotemporal distribution of the crops, faster data acquisition, and lower cost. These conclusions have been demonstrated by researchers in different countries, such as India (Biggs et al. 2006), Pakistan (Dempewolf et al. 2014), China (Frolking et al. 1999), South and Southeast Asia (Xiao et al. 2006), Brazil (Junges and Fontana 2009), among others. Thus, several surveys using remote sensing have been developed to support a wide range of users (government agencies, researchers, and farmers) (Lobell and Asner 2004). The estimation of wheat production has many applications, such as the formulation of supply policies and the grain industry.

Remote sensing techniques for the agricultural crop mapping focus on time series processing, which identifies the cultivation types by their phenological phases,

impossible from data of only one date (Pan et al. 2015; Hao et al. 2016). In this context, high temporal resolution-satellite data bring new possibilities for solving problems of crop phenology detection. The MODIS (Moderate Resolution Imaging Spectroradiometer) sensor has been widely used in the temporal analysis because of its high temporal resolution (quasi-daily repetition) and moderate spatial resolution (250m) (Justice et al. 1998). Therefore, MODIS images and time series processing techniques have been widely used for different wheat crop studies: area estimation and spatial distribution (e.g., Ozdogan 2010; Pan et al. 2012); monitoring (e.g., Wenbo et al. 2007; Feng et al. 2009); wheat leaf area index retrieval (e.g., Yi et al. 2008; Dong et al. 2016); gross primary productivity (e.g., Yan et al. 2009); forecasting wheat yields (e.g., Wit et al. 2012; Kouadio et al. 2012).

The main strategies used to classify wheat planting by remote sensing time series are methods based on specific phenological metrics (Pan et al. 2012; Chu et al. 2016) and supervised classifiers such as maximum likelihood (Murthy et al. 2003), artificial neural networks (Murthy et al. 2003), support vector machine (Low et al. 2013), and tree decision (Wardlow and Egbert 2008). However, the spatial, temporal, and ecological complexity of the planting areas and their surrounding environment is a challenge for establishing a simple method to describe them. These classifications at annual time scales need to be flexible enough to detect variations within the phenological phases due to management practices and environmental factors (climate, soil, and water supply). Thus, the definition of specific metrics phenological becomes more complex, considering all the different planting cycles and its distinctive features concerning neighboring targets, which in turn also vary. Among the supervised classification methods, decision tree techniques show better performance for wheat crop mapping with considerable intra-class variability attributable to regional variations,

considering different algorithms: (a) C45 (Wardlow and Egbert 2008); (b) random forest (Hao et al. 2015); and (c) gradient boosting (Lawrence et al. 2004). However, these techniques depend on extensive sampling of the various targets present in the scene.

In Southern Brazil, wheat mapping using times-series images accentuate two problems: (a) temporal wheat signatures have high variability due to the distinct regional climatic conditions; and (b) wheat planting focuses on small farmers, with many interspersed targets presents (natural vegetation, other crops, built areas, among others). In this context, this research aims to develop a wheat planting detection method in the Northwest region of Rio Grande do Sul (Brazil) from the MODIS time-series, considering its internal variability and does not require sampling of the other targets present in the scene. The proposed procedure conjugates the nearest neighbor method, a specific case of the K-NN method, with the determination of the best threshold value to adjust the wheat mask. Complementarily, the research seeks to test and compare the metrics of cosine similarity and Euclidean distance in the calculation of the nearest neighbor.

Study Area

The study area is the Northwest region of Rio Grande do Sul, which concentrates wheat production within the state (Camponogara et al. 2015) (Figure 1). According to Köppen's climatic classification, the region is in the Temperate Climatic Zone (C), humid fundamental climate (f), subtropical specific variety (Cfa), with well-distributed precipitation during the year and average temperature of the warmer month at 22ºC (Junges and Fontana, 2011). The region is under the Plateau North-Rio-Grandense, formed by basaltic rocks of the Mesozoic Era. Altitudes decrease in the East-West direction, varying from 1300 meters to 700 meters (Becker and Nunes,

2012). The vegetation types are fields, Subtropical Forest, and Araucária Forest (the name refers to the predominant species, Araucária Angustifolia) (Moreira 2007). This region has a land use configuration formed by a large concentration of small grain-producing properties.

<<Figure 1>>

Material and Methods

MODIS data

The research used the MODIS Surface Reflectance product (MOD09Q1), with a spatial resolution of 250 meters. This product is an 8-day composition, which selects the best pixels with low view angle, the absence of clouds or cloud shadow, and aerosol loading (Gumma et al. 2015). The MODIS Reprojection Tool (MRT) was used to convert the images in the sinusoidal projection and Hierarchical Data Format (HDF) format to the geographic coordinate system (WGS84 datum). The images correspond to the vegetative cycle of the wheat in the State of Rio Grande do Sul, when the plantation occurs during May, June and July and the harvest during October, November, December.

Elaboration and filtering of the NDVI temporal cube

The Normalized Difference Vegetation Index (NDVI) calculation uses the red (ρRED) (620-670 nm) and near-infrared (ρNIR) (841-875 nm) bands in the following equation (Rouse et al. 1973):

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{RED}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{RED}}} \quad (1)$$

Subsequently, we group the time series images into a cube, where the "x" and "y" axes correspond to the geographic coordinates and the "z" axis to the increasingly

ordered NDVI temporal signatures. The time series are susceptible to noise from data acquisition errors, the presence of clouds and shadows, and atmospheric inferences (Hird and McDermid 2009; Carvalho Júnior et al. 2013). In this research, we use the Savitzky-Golay (S-G) smoothing method to minimize noise (Savitzky and Golay 1964). This convolution technique fits an array of "n" equally-spaced points to a polynomial, which is a subgroup of complete temporal series. The S-G method tends to preserve peak attributes such as height, width, asymmetry, maximum and minimum point (Schefer 2011). This peak shape conservation property by the S-G filters is desirable in the remote sensing time-series processing for vegetation analysis, which uses these attributes to distinguish the phenological phases (Chen, et al. 2004; Abade et al. 2015). Figure 2 demonstrates the effect of the S-G filtering for a wheat temporal profile present in the study area.

<<Figure 2>>

Selection and classification of temporal signatures

The selection of NDVI temporal signatures of wheat considered a spatially distributed sampling, depicting the different sowing dates and vegetation growth. Thus, 13 temporal signatures exemplified the different phenological patterns of wheat in the region. The proposed method is based on two steps: (a) establishing a metric that highlights the area of interest; and (b) apply a mask from the threshold value. In this approach, searches for several targets have been developed from the threshold value determination for spectral indices: (a) land-cover change detection from NDVI differencing (Lunetta et al. 2006); (b) burned area detection using NDVI and Normalized Burned Ratio (NBR) differencing (Escuin et al. 2008), and (c) detection of built-up areas using the built-up index-based index (IBI) (Xu et al. 2008).

In this research, we adopted the nearest neighbor metrics from the k-NN algorithm to evidence wheat growing areas, instead of taking a spectral index. The k-NN algorithm is a non-parametric multivariate method and one of the simplest algorithms among machine learning algorithms, but yields highly competitive results (Cover and Hart, 1967). The classification criterion of a point is the majority class of its k neighbors, where k is a user-defined constant consisting of a positive integer. If k is equal to 1, the method is called the "nearest neighbor". The adoption of this variant of the k-NN method was due to the quality of the selected samples and the variation of their signatures. This method allows "learning by example" since it uses a known response close to the given vector.

Typically, the metric used to identify nearest neighbors is the Euclidean distance. In addition to this metric, we evaluated the measure of cosine similarity (Salton and Mcgill, 1983), which is widely used in spectral classification from the Spectral Angle Mapper algorithm (Kruse et al. 1993). This metric determines the spectral similarity between two vectors, calculating the angle between them in a space with dimensionality equal to the number of bands. The value similarity is maximum when the vectors are in the same direction (angle equal to 0º) and becomes minimum when perpendicular (angle equal to 90º). The difference between the two metrics is the ability to recognize forms (similarity metric) or establish differences of magnitude (distance metric), generating additional information for classification (Carvalho Júnior et al. 2011). The inclusion of cosine similarity in the calculation of the k-NN method was performed in the *Abilius* program (Carvalho Júnior et al. 2011) using C ++ language. The k-NN method is widely used to classify remote sensing data, mainly in forestry studies (e.g., Chirici et al. 2016; Magnussen et al. 2009).

The detection of the best threshold value between the wheat and non-wheat areas in the nearest neighbor metric images adopted the method proposed by Carvalho Júnior et al. (2015). Initially, the use of this algorithm was to detect the best threshold value between burned and unburned areas in NBR images. The method considers two input data: (a) reference map; and (b) metric image. The technique automatically applies a series of threshold values and compares the resulting classifications with a reference map from the Kappa index. The highest Kappa index is the best threshold value. The elaboration of the reference map used the visual interpretation of Landsat 8 images in an area corresponding to 5% of the total area studied.

Accuracy analysis

In the accuracy analysis, we used the confusion matrix, Overall coefficient, Kappa index and errors of commission and omission (Congalton and Green, 1993). The accuracy analysis compared the classified images and the ground truth image produced from the visual interpretation of Landsat-8 images and field data, corresponding to 10% of the total area mapped. The location of this ground truth map was different from the reference map for the detection of the best threshold value in the wheat classification.

McNemar's test assessed the statistical significance between the two nearest neighbor classifications (McNemar, 1947). This test is a non-parametric method that establishes a binary distinction between correct and incorrect classes from the same set of samples, assuming a confusion matrix with a size of 2 by 2 (Foody, 2004). The chi-square ($\chi 2$) statistic with one degree of freedom is computed from confusion matrix by the following equation:

$$x^2 = \frac{(f_{12} - f_{21})^2}{f_{12} - f_{21}} \quad (2)$$

Where f_{12} is the number of wrong classifications by method 1 but correct by method 2, and f_{21} is the number of correct classifications by method 1 but wrongly classified by method 2. This accuracy comparison based on related samples are deep-rooted in the literature, containing broad application (e.g., Leeuw et al. 2006, Manandhar et al. 2009). In the McNemar's test, we used 1000 samples.

Moreover, we compared the planted area by time series classification with the official statistics from the Brazilian Institute of Geography and Statistics (IBGE) obtained by direct interviews. In this approach, several studies compare agricultural census statistics and maps from digital image processing (eg., Biradar and Xiao, 2011; Frolking et al. 1999, Junges and Fontana, 2009; Qiu et al. 2003).

Results

Analysis of Temporal Signatures of Wheat Crop

Figure 3 shows the variation of wheat temporal signatures with different dates to start planting and maximum NDVI position since each region has specific differences of the agricultural calendar. The minimum values of NDVI occurred in April in the period before sowing, and maximum values occurred from July during the vegetative development stage of the plant. The decrease of the NDVI value from September is a function of leaf senescence when the plants began to reach maturity. Junges and Fontana (2009) found similar results for the temporal signatures of winter crops in the Rio Grande do Sul.

<<Figure 3>>

Results of the Ratings of the Time Series

Figures 4 represents the Kappa index curves in the detection of the best threshold value

for wheat crop classification, considering the nearest neighbor metrics of the Euclidean distance and cosine similarity. The optimal threshold for Euclidean distance was 0.710 with Kappa index of 78.15 (Figure 4a), while for cosine similarity was 0.203 degrees with Kappa index of 75.60 (Figure 4b).

Figure 5 shows the wheat planting masks from the two metrics. The classification using the Euclidean distance metric showed an overall accuracy of 90.01%, Kappa index of 0.75, an error of omission of 21.17%, and an error of commission of 13.06% (Table 1). This performance was slightly higher than that of the cosine similarity metric with an overall accuracy of 89.32%, Kappa index of 0.74, an error of omission of 21.50%, and an error of commission of 14.93%.

<<Figure 4 and 5>>

The McNemar's test certified the resemblance of the nearest neighbor classifications using Euclidian distance and cosine similarity (Table 2). The calculated χ^2 was 1.125 so the test provides evidence to accept the null hypothesis of marginal homogeneity states or equal classifier performance at the 5% level of significance ($p = 0.05$). In addition, the 2x2 contingency table demonstrated 96.8% of equal pairs (correct vs. correct and incorrect vs. incorrect).

Discussion

The results demonstrate a significant variation of the wheat temporal curves. Different factors caused the phenological variation of the cultivated areas: genetic varieties, latitude positioning, temperature, precipitation, photoperiod, soil altimetry, soil, plant age, pest presence, among others. In addition, the temporal behavior of the crop is influenced by farmer decisions and management practices (Wardlow and Egbert 2008). The results may also present changes due to space-time resolution, smooth procedures,

and methods for extracting phenological features (Chu et al. 2016). Therefore, the variability of temporal signatures can be high even within the same culture.

Specifically, in the analysis of wheat phenological profiles using remotely-sensed temporal series, several studies have shown spatiotemporal variations (Lobell et al. 2013, Vyas et al. 2013). Huang et al. (2016) found relationships between the different growth stages of winter wheat and meteorological factors. Franch et al. (2015) verified temporal variations of the NDVI peak position due to climatic variations. Lu et al. (2014) evidenced spatial changes in wheat phenology in relation to the degree of latitude in the Northern China Plain. Manfron et al. (2017) demonstrated a marked interannual variability in sowing times for winter wheat mainly due to the type of the previous crop and the climatic conditions during the summer period before to planting. These variations explain the limitation of establishing a single threshold value on a vegetation index to detect vegetation types and crops (White et al. 1997; Fisher et al. 2006).

Further research indicates that the climate fluctuations during the crop cycle have a strong influence on the development of the wheat plantation in Southern Brazil, conditioning its yield, economic viability, and industrial quality characteristics. Cunha et al. (2001) generated maps of the most suitable wheat-sowing time for the State of the Rio Grande do Sul, evidencing a significant delay between the regions. In the northwest of the State, sowing time begins earlier due to its warmer temperatures (May 1), while the northeastern part with higher altitudes and lower temperatures may occur with a lag time of three months (July 31). Silva et al. (2011) identified these variations also in the neighboring state (Paraná), where two regions have different sowing seasons to maximize yield of wheat grains, being in the Guarapuava region in July and Palotino region in April. Therefore, wheat crop phenology has a pronounced variation in South

Brazil, since the weather oscillations and plant development stages are not uniform in the spatial-temporal scale.

The proposed method has two crucial qualities for wheat mapping in the study area. The first quality is the high flexibility for significant differences in wheat phenological behavior (sowing dates or amplitude variations) without the need for pre-processing or selection of specific features. This characteristic demonstrates an advantage over the methods based on specific phenological metrics (Pan et al. 2012; Chu et al. 2016), which need to identify and include detailed aspects of regional variations. The second quality is not having to sample the other targets of the scene, consisting of an advantage over supervised methods (Murthy et al. 2003; Murthy et al. 2003; Low et al. 2013; Wardlow and Egbert 2008). The sampling of the different targets intensifies in the study area due to many small farms interspersed with other types of targets.

The results demonstrated a statistical equivalence between the nearest neighbor methods using the Euclidean distance and cosine similarity. This equivalence differs from other classification types using the two metrics such as spectral analysis (Sohn and Rebello 2002) or change detection (Carvalho Júnior et al., 2011). The statistically equivalent result can be induced by the nearest neighbor method that selects the sample within the training data of the shortest distance or angle, favoring correct classification and convergence of the results between both metrics, despite their differences.

The wheat mapping showed accuracy values compatible with other surveys with the MODIS sensor and obtained a planted area close to the estimated values by traditional agricultural surveys. In the analysis of accuracy, an error source frequently reported in agricultural crop mappings using MODIS NDVI time series is the use of higher resolution ground truth maps (e.g., Li and Tiam, 2011; Junges et al. 2013;

Dempewolf et al. 2014). This problem is minimized in extensive and continuous plantation mappings (Wardlow et al. 2006; Wardlow and Egbert, 2008), while the errors are aggravated in small areas of cultivation, for example, small wheat farms in the region Yellow River Delta, China (Chu et al. 2016). In the study area, wheat cultivation occurs in small farms, resulting in an overestimation of the planted area and a high error concentration at the property boundaries. Therefore, part of the error is inherent in the accuracy analysis that compares data with different resolution. The use of new sensors with high spatial and temporal resolution allows minimizing these limitations in the precision of the crop areas.

Conclusions

This research developed a method to classify wheat crops using MODIS NDVI time series from the threshold values of distance and similarity metrics of the nearest neighbor. This approach allows similar results for different behaviors of wheat phenology, which presents a significant variation for the study area. Besides, the method does not require the extraction of phenological metrics or the sampling of other targets present in the scene. One proposed innovation was the use of the cosine similarity metric for the calculation of the nearest neighbor. This similarity metric is used in the SAM classifier, widely applied in spectral classification. In the wheat detection, the nearest neighbor methods using Euclidean distance and cosine similarity obtained statistically equivalent results according to McNemar's test. The results of wheat planted area are compatible with the agricultural census using traditional field surveys in the same year of analysis. Government departments committed to the dissemination of statistics on agricultural production may use the methodology to assist in the work of calculating planted areas.

References

Abade NA, de Carvalho Júnior OA, Guimarães RF, Oliveira SN. 2015. Comparative analysis of MODIS time-series classification using support vector machines and methods based upon distance and similarity measures in the Brazilian Cerrado-Caatinga boundary. Remote Sensing, 7(9):12160-12191.

Asseng S, Ewert F, Martre P, Rötter R, Lobell D, Cammarano D, Kimball BA, Ottman M, Wall G, White J, et al. 2014. Rising temperatures reduce global wheat production. Nat. Clim. Change 5:143–147.

Atchison J, Head L, Gates A. 2010. Wheat as food, wheat as industrial substance; comparative geographies of transformation and mobility. Geoforum, 41(2):236-246.

Becker ELS, Nunes MP. 2012. Relevo do Rio Grande do Sul, e sua representação em maquete [Relief of Rio Grande do Sul, Brazil, and their representation in mockup]. Revista Percurso 4:113-132,

Biggs TW, Thenkabail PS, Gumma MK., Scott CA, Parthasaradhi GR, Turral, HN. 2006. Irrigated area mapping in heterogeneous landscapes with MODIS time series, ground truth and census data, Krishna Basin, India. International Journal of Remote Sensing 27(19):4245-4266.

Biradar CM, Xiao X. 2011. Quantifying the area and spatial distribution of double-and triple-cropping croplands in India with multi-temporal MODIS imagery in 2005. International Journal of Remote Sensing 32(2):367-386.

Camponogara A, Gallio E, Borba WF, Georgin J. 2015. O atual contexto da produção de trigo no Rio Grande do Sul [The current context of wheat producing in Rio Grande do Sul]. Revista Eletrônica em Gestão, Educação e Tecnologia Ambiental 19(2):246-257.

Canziani JR, Guimarães VDA. 2009. O trigo no Brasil e no mundo: cadeia de produção, transformação e comercialização [Wheat in Brazil and in the world: production, processing and marketing chain]. In: Cunha GR, editor, Oficina sobre trigo no Brasil: bases para a construção de uma nova triticultura [Workshop on wheat in Brazil: bases for the construction of a new Brazilian wheat crop]. Passo Fundo, RS, Brazil: Embrapa Trigo, p. 29-72.

Carvalho Júnior OA, Guimarães RF, Gillespie AR, Silva NC, Gomes RAT. 2011. A New approach to change vector analysis using distance and similarity measures. Remote Sensing 3: 2473-2493.

Carvalho Júnior OA, Guimarães RF, Silva CR, Gomes RAT. 2015. Standardized time-series and interannual phenological deviation: new techniques for burned-area detection using long-term MODIS-NBR Dataset. Remote Sensing 7: 6950-6985.

Carvalho Júnior OA, Guimarães RF, Silva NC, Gillespie AR, Gomes RAT, Silva CR, de Carvalho APF. 2013. Radiometric normalization of temporal images combining automatic detection of pseudo-invariant features from the distance and similarity spectral measures, density scatterplot analysis, and robust regression. Remote Sensing, 5(6):2763-2794.

Chen J, Jönsson P, Tamura M, Gu Z, Matsushita B, Eklundh L. 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote Sensing of Environment 91: 332–344.

Chirici G, Mura M, Mcinerney D, Py N, Tomppo EO, Waser LT, Travaglini D, Mcroberts RE. 2016. A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. Remote Sensing of Environment 176: 282-294.

Chu L, Liu QS, Huang C, Liu GH. 2016. Monitoring of winter wheat distribution and phenological phases based on MODIS time-series: A case study in the Yellow River Delta, China. Journal of Integrative Agriculture 15: 2403-2416.

Congalton RG, Green K. 1993. A Practical look at the sources of confusion in error matrix generation. Photogrammetric Engineering & Remote Sensing 59(5):641-644.

Cover T, Hart P. 1967. Nearest neighbor pattern classification. IEEE Transactions on Information Theory 13(1):21-27.

Cunha GR, Haas JC, Maluf JRT, Caramori PH, Assad ED, Braga HJ, Zullo Júnior J, Lazzarotto C, Gonçalves S, Wrege M, Brunetta D, Dotto SR, Pinto HS, Brunini O, Thomé VMR, Zampieri SL, Pasinato A, Pimentel MBM, Pandolfo C. 2001. Agricultural zoning and sowing dates for wheat in Brazil. Revista Brasileira de Agrometeorologia, 9(3):446-459.

Dempewolf J, Adusei B, Becker-Reshef I, Hansen M, Potapov P, Khan A, Barker B. 2014. Wheat yield forecasting for Punjab Province from vegetation index time series and historic crop statistics. Remote Sensing 6(10):9653-9675.

Dixon J, Braun HJ, Crouch J. 2009. Overview: transitioning wheat research to meet the future demands of the developing world. In: Dixon J, Braun HJ, Kosina P, Crouch J., editors. Wheat facts and futures 2009. Mexico, DF: International Maize and Wheat Improvement Center (CIMMYT). Chapter 1, p. 1-25.

Dixon J, Nalley L, Kosina P, La Rovere R, Hellin J, Aquino P. 2006. Adoption and economic impact of improved wheat varieties in the developing world. Journal of Agricultural Science, 144, 489–502.

Dong T, Liu J, Qian B, Zhao T, Jing Q, Geng X, Wang J, Huffman T, Shang J. 2016. Estimating winter wheat biomass by assimilating leaf area index derived from fusion of Landsat-8 and MODIS data. International Journal of Applied Earth Observation and Geoinformation, 49, 63-74.

Escuin S, Navarro R, Fernandez P. 2008. Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. International Journal of Remote Sensing 29(4):1053-1073.

Feng MC, Yang WD, Cao LL, Ding GW. 2009. Monitoring winter wheat freeze injury using multi-temporal MODIS data. Agricultural Sciences in China, 8(9):1053-1062.

Fisher JI, Mustard JF, Vadeboncouer MA. 2006. Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. Remote Sensing of Environment 100: 265−279.

Foody GM. 2004. Thematic map comparison. Photogrammetric Engineering & Remote Sensing, 70(5):627-633.

Franch B, Vermote EF, Becker-Reshef I, Claverie M, Huang J, Zhang J, Justice C, Sobrinho JA. 2015. Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. Remote Sensing of Environment 161: 131–148.

Franke J, Menz G. 2007. Multi-temporal wheat disease detection by multi-spectral remote sensing. Precision Agriculture, 8(3): 161-172.

Frolking S, Xiao X, Zhuang Y, Salas W, Li C. 1999. Agricultural land-use in China: A comparison of area estimates from ground-based census and satellite-borne remote sensing. Global Ecology and Biogeography 8(5):407-416.

Gumma MK, Mohanty S, Andrew N, Rala A, Irshad AM, Das SR. 2015. Remote sensing based change analysis of rice environments in Odisha, India. Journal of Environmental Management 148: 31-41.

Hao P, Wang L, Zhan Y, Niu Z. Using Moderate-Resolution Temporal NDVI profiles for high-resolution crop mapping in years of absent ground reference data: a case study of Bole and Manas Counties in Xinjiang, China. ISPRS International Journal of Geo-Information 5(67):1-23, 2016.

Hao P, Zhan Y, Wang L, Niu Z, Shakir M. 2015. Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA. Remote Sensing, 7(5): 5347-5369.

Hird JN, McDermid GJ. 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. Remote Sensing of Environment, 113(1):248-258.

Huang Q, Wang LM, Chen ZX, Hang L. 2016. Effects of meteorological factors on different grades of winter wheat growth in the Huang-Huai-Hai Plain, China. Journal of Integrative Agriculture 15(11):2647-2657.

[IBGE] Instituto Brasileiro de Geografia e Estatística. 2016. Levantamento Sistemático da produção Agrícola: pesquisa mensal de previsão e acompanhamento das safras agrícolas no ano civil [Systematic survey of agricultural production: monthly survey of the forecast and monitoring of agricultural crops in the calendar year]. Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística, 29(5):1-79.

Junges AH, Fontana DC, Pinto DG. 2013. Identification of croplands of winter cereals in Rio Grande do Sul state, Brazil, through unsupervised classification of normalized difference vegetation index images. Engenharia Agrícola 33(4):883-895.

Junges AH, Fontana DC. 2009. Desenvolvimento das culturas de cereais de inverno no Rio Grande do Sul por meio de perfis temporais do índice de vegetação por diferença normalizada [Evaluation of winter cereal crop development in Rio Grande do Sul, Brazil, throughout the temporal profiles of normalized difference vegetation index]. Ciência Rural 39: 1349-1355,

Junges AH, Fontana DC. 2011. Modelo agrometeorológico-espectral de estimativa de rendimento de grãos de trigo no Rio Grande do Sul. Revista Ceres 58: 9-16,

Justice CO, Vermote E, Townshend JR, Defries R, Roy DP, Hall DK, Salomonson VV, Privette JL, Riggs G, Strahler A, et al. 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. IEEE Transactions on Geoscience and Remote Sensing, 36(4):1228-1249.

Kouadio L, Duveillerb G, Djabya B, Jarroudia ME, Defournyb B. 2012. Estimating regional wheat yield from the shape of decreasing curves of green area index temporal profiles retrieved from MODIS data. International Journal of Applied Earth Observation and Geoinformation 18: 111–118.

Kruse FA, Lefkoff AB, Boardman JB, Heidebrecht KB, Shapiro AT, Barloon PJ, Goetz AFH. 1993. The Spectral Image Processing System (SIPS) - interactive visualization and analysis of imaging spectrometer Data. Remote Sensing of the Environment 44: 145–163.

Lawrence R, Bunn A, Powell S, Zambron M. 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. Remote Sensing of Environment, 90: 331−336.

Leeuw J, Jia H, Yang L, Liu X, Schmidt K, Skidmore AK. 2006. Comparing accuracy assessments to infer superiority of image classification methods. International Journal of Remote Sensing, 27(1):223-232.

Li S, Tian Q. 2011. Winter wheat area extraction and estimation based on MODIS-VI time series and multi-temporal HJ CCD images in Jiangsu province, China. In: Proceedings of the 2011 19th International Conference on Geoinformatics; Jun 26-24; Shanghai; IEEE. p. 1-6.

Lobell D, Burke M. 2010. Economic impacts of climate change in agriculture. In: Reynolds MP, editor, Climate change and crop production. Wallingford: CABI.

Lobell DB, Asner, GP. 2004. Cropland distributions from temporal unmixing of MODIS data. Remote Sensing of Environment, 93(3):412-422.

Lobell DB, Ortiz-Monasterio JI, Sibley AM, Sohu VS. 2013. Satellite detection of earlier wheat sowing in India and implications for yield trends. Agricultural Systems 115: 137–143.

Low F, Michel U, Dech S, Conrad C. 2013. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines. ISPRS Journal of Photogrammetry and Remote Sensing, 85: 102–119.

Lu L, Wang C, Guo H, Li Q. 2014. Detecting winter wheat phenology with SPOT-VEGETATION data in the North China Plain. Geocarto International 29:244–255.

Lunetta RS, Knight JF, Ediriwickrema, J, Lyon, JG, Worthy, LD. 2006. Land-cover change detection using multi-temporal MODIS NDVI data. Remote Sensing of Environment 105(2):142-154.

Magnussen S, Mcroberts RE, Tomppo EO. 2009. Model-based mean square error estimators for k-nearest neighbor predictions and applications using remotely sensed data for forest inventories. Remote Sensing of Environment 113: 476–488.

Manandhar R, Odeh IO, Ancev T. 2009. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. Remote Sensing 1(3):330-344.

Manfron G, Delmotte S, Busetto L, Hossard L, Ranghetti L, Brivio PA, Boschetti M. 2017. Estimating inter-annual variability in winter wheat sowing dates from satellite time series in Camargue, France. International Journal of Applied Earth Observation and Geoinformation 57: 190-201.

McNemar Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika, 12(2):153–157.

Moreira I. O espaço rio-grandense [The space of the rio grande do sul]. São Paulo: Editora Ática, 2007. p. 96.

Murthy CS, Raju PV, Badrinath KVS. 2003. Classification of wheat crop with multi-temporal images: performance of maximum likelihood and artificial neural networks. International Journal of Remote Sensing, 24(23): 4871-4890.

Ozdogan M. 2010. The spatial distribution of crop types from MODIS data: temporal unmixing using independent component analysis. Remote Sensing of Environment 114: 1190-1204.

Pan Y, Li L, Zhang J, Liang S, Zhu X, Sulla-Menashe D. 2012.Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. Remote Sensing Environmental 119: 232–242.

Pan Z, Huang J, Zhou Q, Wang L, Cheng Y, Zhang H, Blackburn GA, Yan J., Liu J. 2015. Mapping crop phenology using NDVI time-series derived from HJ-1A/B data. International Journal of Applied Earth Observation and Geoinformation 34: 188-197.

Pingali P. 2007. Westernization of Asian diets and the transformation of food systems: implications for research and policy. Food Policy, 32, 281–298.

Qiu J, Tang H, Frolking S, Boles S, Li C, Xiao X, Liu J, Zhuang Y, Qin X. 2003. Mapping Single-, Double-, and Triple-crop Agriculture in China at 0.5°× 0.5° by Combining County-scale Census Data with a Remote Sensing-derived Land Cover Map. *Geocarto International* 18(2):3-13.

Rouse JW, Haas RH, Schell JA. Deering DW. 1973. Monitoring vegetation systems in the Great Plains with ERTS. In: Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, 1973, Proceedings NASA SP-351, p. 301–317.

Salton G, Mcgill MJ. 1983. Introduction to modern information. Philadelphia, PA: American Association for Artificial Intelligence Retrieval.

Savitzky A, Golay MJE. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry 36(8):1627-1639.

Schefer RW. 2011. What is a Savitzky-Golay filter? IEEE Signal Process. Mag. 28, 111–117.

Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M, Muricho G. 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Security, 5(3):291-317.

Silva RR, Benin G, da Silva GO, Marchioro VS, de Almeida, JL, Matei G. 2011. Adaptabilidade e estabilidade de cultivares de trigo em diferentes épocas de semeadura, no Paraná [Adaptability and stability of wheat cultivars at different sowing dates in the state of Paraná, Brazil]. Pesquisa Agropecuária Brasileira, 46(11):1439-1447.

Sohn Y., Rebello NS. 2002. Supervised and unsupervised spectral angle classifiers. Photogrammetric engineering and remote sensing, 68(12): 1271-1282.

Sun H, Xu A, Lin H, Zhang L. 2012.Winter wheat mapping using temporal signatures of MODIS vegetation index data. International Journal of Remote Sensing 33(16):5026–5042.

Vyas S, Nigam R, Patel NK, Panigrahy S. 2013. Extracting regional pattern of wheat sowing dates using multispectral and high temporal observations from indian geostationary satellite. J. Indian Soc. Remote Sens. 41: 855–864.

Wardlow BD, Egbert SL. 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. Remote Sensing of Environment 112: 1096–1116.

Wardlow BD, Egbert SL. 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the US Central Great Plains. Remote Sensing of Environment, 112(3): 1096-1116.

Wardlow BD, Kastens JH, Egbert SL. 2006. Using USDA crop progress data for the evaluation of greenup onset date calculated from MODIS 250-meter data. Photogrammetric Engineering & Remote Sensing 72(11):1225-1234.

Wenbo X, Guoping Z, Jinlong F, Yonglan, Q. 2007. Remote sensing monitoring of winter wheat areas using MODIS data. Transactions of the Chinese Society of Agricultural Engineering, 2007(12).

White MA, Thornton PE, Running SW. 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global Biogeochemical Cycles 11: 217–234.

Wit A, Duveiller G, Defourny P. 2012. Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations. Agricultural and Forest Meteorology, 164:39-52.

Xiao X, Boles S, Frolking S, Li C, Babu JY, Salas W, Moore III B. 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. Remote Sensing of Environment 100(1):95-113.

Xu H. 2008. A new index for delineating built-up land features in satellite imagery. International Journal of Remote Sensing 29(14):4269-4276.

Yan H, Fu Y, Xiao X, Huang HQ, He H, Ediger L. 2009. Modeling gross primary productivity for winter wheat–maize double cropping system using MODIS time series and CO 2 eddy flux tower data. Agriculture, Ecosystems & Environment, 129(4):391-400.

Yi Y, Yang D, Huang J, Chen D. 2008. Evaluation of MODIS surface reflectance products for wheat leaf area index (LAI) retrieval. ISPRS Journal of Photogrammetry and Remote Sensing, 63(6):661-677.

Table 1. Confusion Matrix of the classifications based on the nearest neighbor metric images: Euclidean distance and cosine similarity. Where "EO" is the error of omission and "EC" is the error of commission.

Euclidian Distance Overall Accuracy: 90.01 Kappa Coefficient: 0.75				Cosine similarity Overall Accuracy: 89.32 Kappa Coefficient: 0.74			
Classes	Wheat	No-Wheat	Total	Classes	Wheat	No-Wheat	Total
Wheat	78.83	5.14	27.43	Wheat	78.50	5.97	27.91
No-	21.17	94.86	72.57	No-	21.50	94.03	72.09

Wheat				Wheat			
Total	100	100	100	Total	100	100	100
EO	13.06	21.17	-	E O	14.93	21.50	-
EC	8.82	5.14	-	E C	9.02	5.97	-

Table 2. McNemar's test between the two classifications by the nearest neighbor using Euclidean distance and the cosine similarity.

Allocation	Euclidian Distance		
	Correct	Incorrect	Total
Cosine Similarity			
Correct	881	13	894
Incorrect	19	87	106
Total	900	100	1000
χ^2	1.125		

Figures

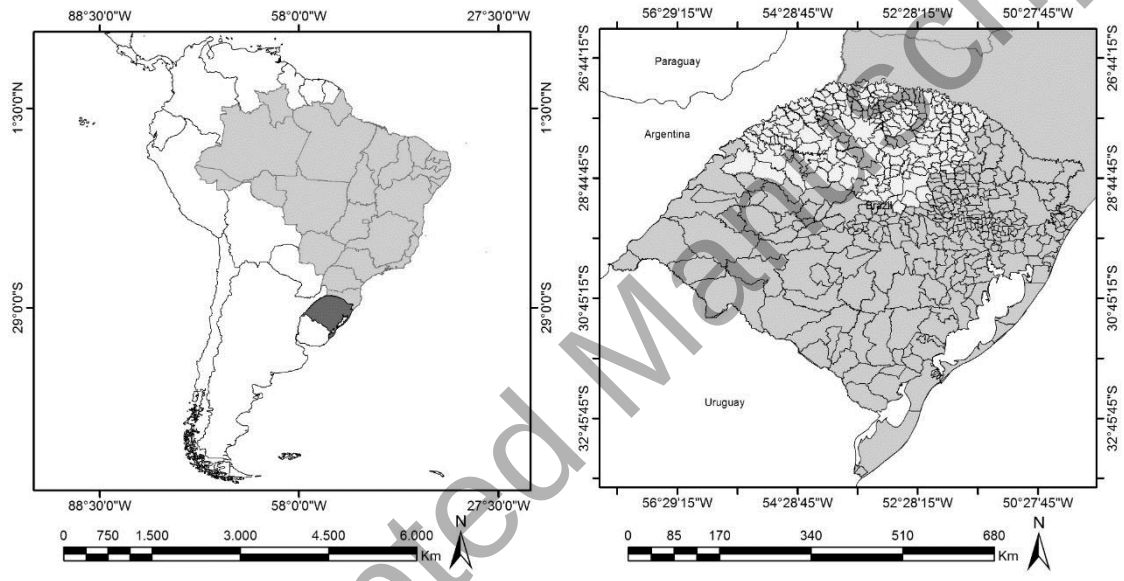

Figure 1. Location map of the study area.

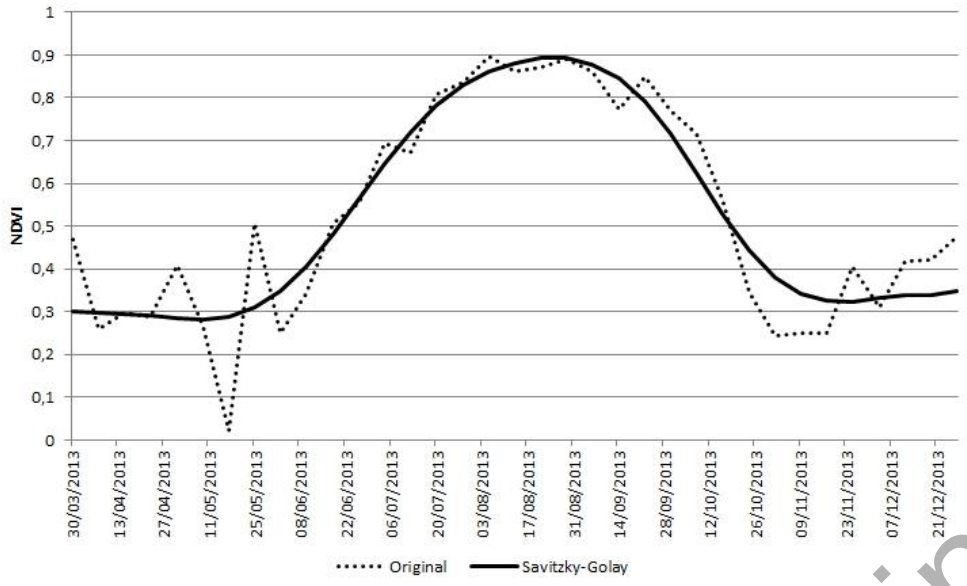

Figure 2. Filtering of the MODIS NDVI time series by the Savitzky-Golay method.

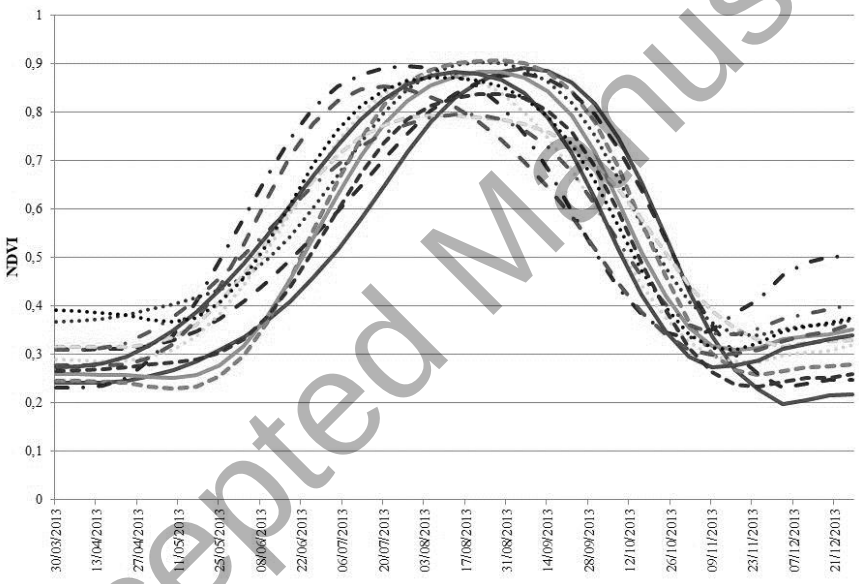

Figure 1. The thirteen MODIS NDVI temporal signatures for wheat crop in the Northwest Rio Grande do Sul.

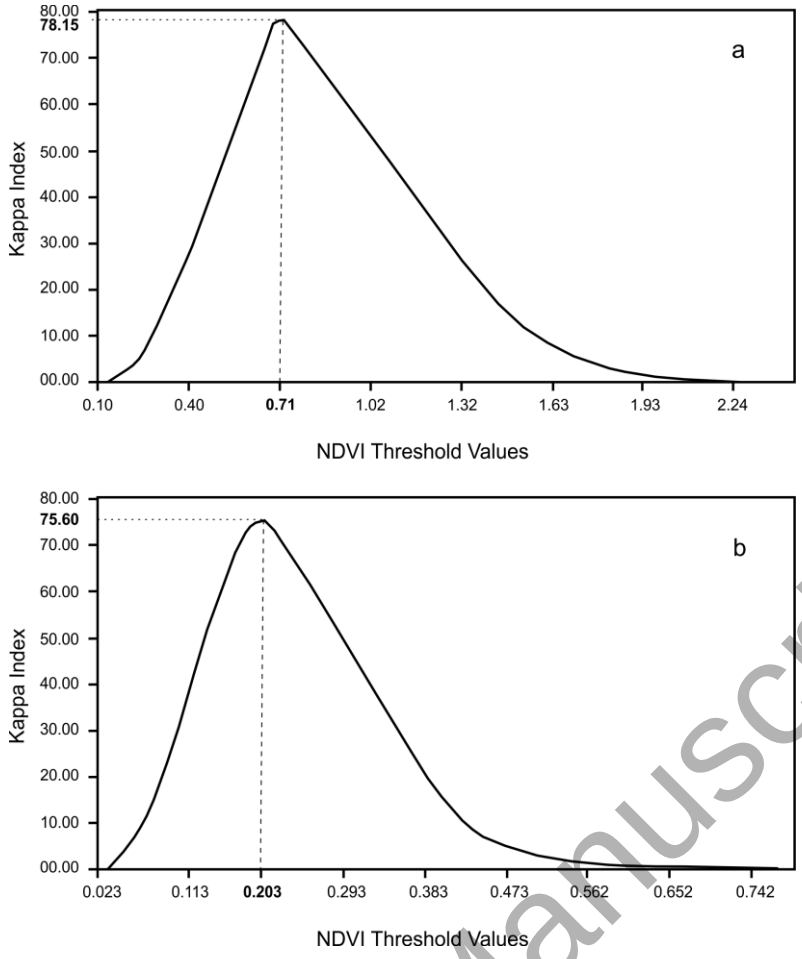

Figure 4. Kappa index curve for the definition of threshold values for the two metrics of the nearest neighbor: (a) Euclidean distance and (b) cosine similarity.

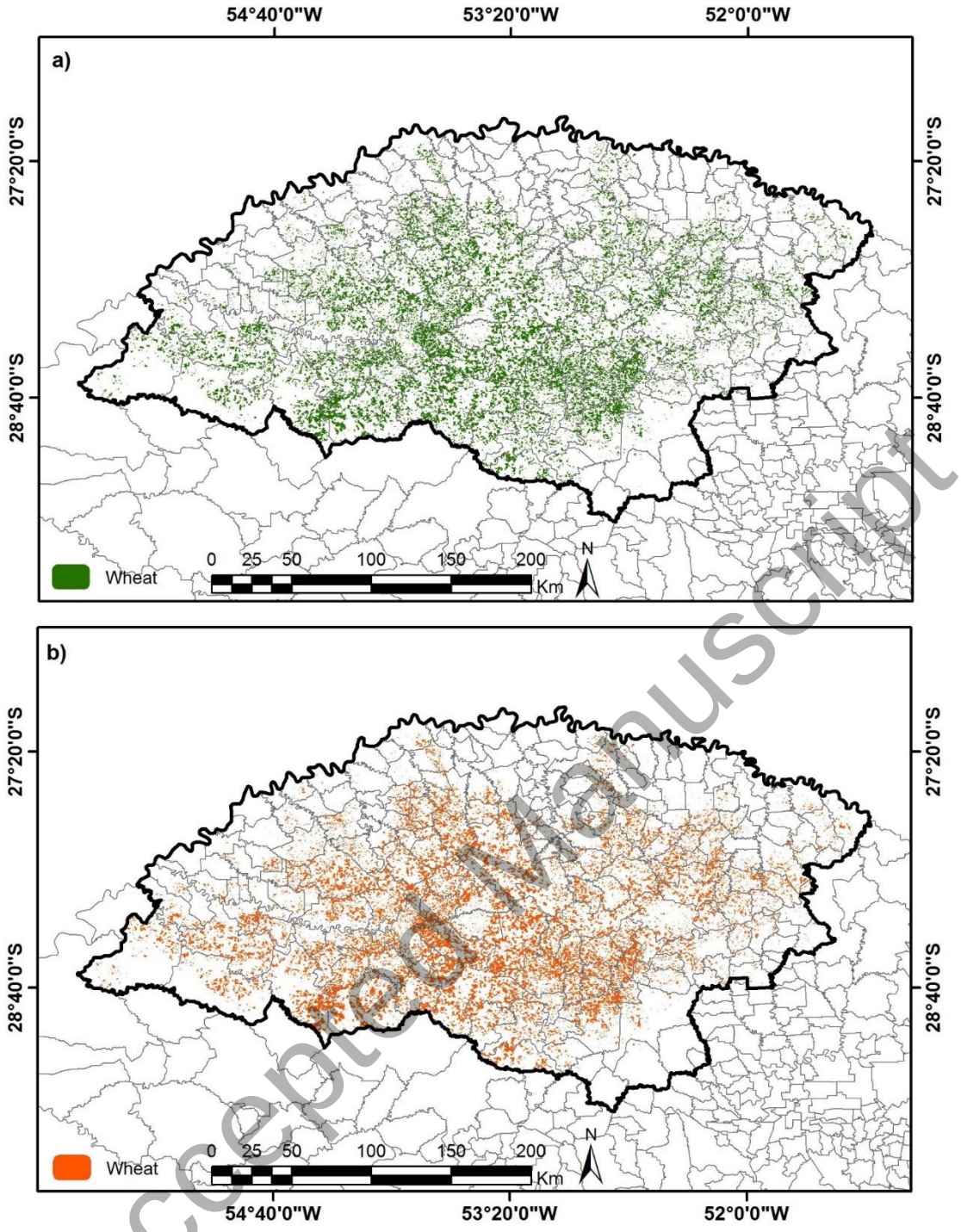

Figure 5. Wheat planting map by using the threshold value in the nearest neighbor metric images: (a) Euclidian distance; and (b) cosine similarity.