*Tarla Bitkileri Merkez Araştırma Enstitüsü Dergisi, 2017, 26 (1):11-23*
*Araştırma Makalesi (Research Article)*

# Mapping Wheat Growing Areas of Turkey by Integrating Multi-Temporal NDVI Data and Official Crop Statistics

*Ediz ÜNAL[1] C.A.J.M (Kees) DE BIE[2]

[1]Central Research Institute for Field Crops, Ankara, Türkiye
[2]University of Twente, Faculty of Geo-Information Science and Earth Observation, Enschede, Netherlands
*Corresponding author e-mail (Sorumlu yazar e-posta): ediz.unal@tarim.gov.tr

Geliş Tarihi (Received): 12.01.2017 Kabul Tarihi (Accepted): 15.05.2017

**Abstract**

Wheat is the most widely cultivated crop in the world providing critical food source of most countries. It exceeds most of the grain crops in acreage and production because of its ability to grow in wide range of climatic and geographic conditions. Timely and reliable information on wheat acreages is essential for government services in order to formulate their policies for planning of agricultural production and monitoring their food supply. Traditionally, agricultural statistics is considered as the main source of such information. Unfortunately, existing statistical data of wheat acreages of Turkey, mostly dependent on farmers' declarations, does not provide spatial information of where this crop specifically is grown. Satellite remote sensing technology can enable the acquisition of such information indirectly with the use of ancillary data of crop statistics. This study aims to determine wheat cultivation areas of Turkey as percentage per unit area in a crop map by integrating time series satellite NDVI imagery with the official crop statistics through regression analysis. The regression results indicated that satellite data explained 95.8% of the variability in official wheat crop statistics and actual wheat cropping areas were significantly related to NDVI-based wheat classes. Validation of the produced wheat map showed that there was good agreement between actual wheat fractions and estimated NDVI-based wheat fractions explaining approximately 69% (Adj. $R^2$) of the total variability between them. This study suggests use of the methodology employed here to governing bodies that need to identify and to map current wheat cropping areas.

**Keywords:** Map, NDVI, agricultural statistics, wheat, Türkiye

## Zaman Serisi NDVI Verileri ve Resmi Tarım İstatistikleri Kullanarak Türkiye Buğday Alanlarının Haritalandırılması

**Öz**

Buğday dünya genelinde tarımı en yaygın yapılan tarım ürünüdür ve birçok ülke için ana besin kaynağı olarak görülmektedir. Geniş iklimsel ve coğrafi koşullar altında yetişebilme özelliğinden dolayı, buğdayın üretim miktarı ve yetiştirme alanı diğer tahıl ürünlerinden daha fazladır. Buğday tarımı yapılan alanlarla ilgili olarak güncel ve güvenilir bilgiye erişim, ülkelerin tarımsal üretimlerini planlamaya ve üretim alanlarını gözlemlemeye yönelik politikaların geliştirilmesinde büyük önem arz etmektedir. Tarım istatistikleri geleneksel olarak bu tür bilgilerin ana kaynağı olarak öngörülse de, ülkemizde olduğu gibi çiftçi beyanına bağlı tarım istatistikleri maalesef hangi tarım ürününün hangi mekânsal konumda yetiştirildiği bilgisini sunmamaktadır. Uzaktan algılama teknolojisi, tarım istatistiklerini yardımcı veri kaynağı şeklinde kullanarak bu tür bilgiyi üretmemize imkân sağlamaktadır. Bu çalışmanın amacı, zaman serisi NDVI verileri ve resmi tarım istatistiklerini regresyon analizi ile entegre ederek Türkiye buğday alanlarını birim alanda yüzde değer olarak belirlemek ve haritalandırmaktır. Regresyon analizi sonuçlarına göre; NDVI uydu verisi, resmi buğday istatistiklerindeki değişkenliğin %95.8'ni açıklayabilmektedir ve gerçek buğday parselleri istatistiksel olarak NDVI verisinden üretilen buğday sınıfları ile önemli derecede ilişkilidir. Regresyon modeli ile elde edilen buğday haritasının doğruluk analizine göre, gerçek buğday alan yüzdeleri ile NDVI verisinden üretilen buğday alan yüzdeleri arasındaki ilişki %69 $R^2$ düzeyindedir. Bu çalışmada kullanılan yöntem, buğday üretimi yapılan parselleri belirlemek isteyen kurumlar için tavsiye edilebilir niteliktedir.

**Anahtar Kelimeler:** Harita, NDVI, tarım istatistikleri, buğday, Türkiye

DOI: 10.21566/tarbitderg.323560

## Introduction

Crop monitoring is important for agriculture-based countries to effectively and sustainably use their resources. Monitoring agriculture always requires up-to-date and reliable information on crop areas and production through which governments and organizations make management plans. In addition, spatial cropland and agricultural statistical data are also needed for improved productivity assessments and comparison of land use versus natural resources management. This rationale can be applied to most of the countries which cultivate crops especially in extensive areas such as Turkey which is in the top 10 cereal crop producers in the world (Anonymous, 2015). In this context, the knowledge on the areas devoted to various crops is mostly depend upon the recorded statistics, but which generally do not provide the spatial information of "which crop is grown where?". The information of crop locations can help governments to take responsive actions for managing agricultural policies and assist farmers to market their products as well.

In many countries, land use crop statistics are mostly logged at administrative unit in tabular form which provides acreage and annual production, but spatial location and distribution (Jansen and Di Gregorio, 2003). Even though these statistical data can be displayed as meaningful crop density maps at sub-national level by using simple GIS software, such maps do not have the ability to show location of crop producing areas. At the same time, preparation of detailed crop-specific land use maps is very labor-intensive, time consuming and costly, because it requires detailed ground surveys throughout the country. Land use maps are therefore infrequently prepared and are not readily available for many countries (Wood et al., 2000).

Recent technologies of remote sensing and geographic information system (GIS) provide novel approaches for monitoring and analyzing the spatial and temporal condition of land use and mapping of various croplands (Oetter et al., 2000; Brand and Malthus, 2004; Shao et al., 2010; Gumma et al., 2015). One of the most widely used remotely sensed data is the Normalized Difference Vegetation Index (NDVI), which is based on the normalized difference between the absorbed and reflected solar energy of red and near-infrared light by live vegetation, and which is assumed to be indicative of vegetative abundance and vigorousity (Goward and Huemmrich, 1992; Campbell, 1996). On the other hand, agricultural lands are active environments where vegetation conditions change temporally because of the phenological aspects of crops. Hyper-temporal NDVI data are suitable for monitoring such agro-ecosystems because of their daily global revisiting capability at continental scale and availability at any time. Hyper-temporal data also known as multi-temporal or time series data consists of daily, weekly, monthly stacked NDVI images for specific range of years. Many studies have shown that hyper-temporal NDVI data can be used effectively for observing both spatial as temporal differences and changes in land use (Groten and Ocatre, 2002; Hill and Donald, 2003; De Bie et al., 2011; Nguyen et al., 2012; Unal et al., 2014).

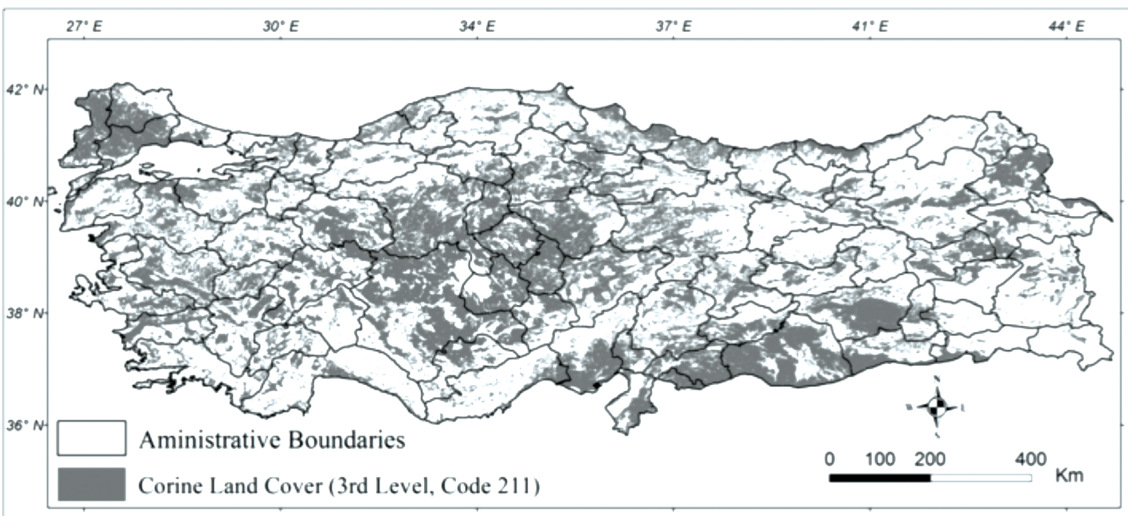


Figure 1. Non-irrigated arable agricultural areas (Corine land cover, 2000) of Turkey

*Şekil 1. Türkiye geneli sulanmayan-ekilebilir tarım alanları (Corine arazi örtüsü, 2000)*

This research explores a method of agricultural land use mapping with satellite earth observation and ground survey data. Specifically, the main goal of this study is to estimate wheat cropping areas of Turkey by integrating hyper-temporal MODIS NDVI imagery with the recorded statistics and then to map wheat areas as percentage per unit area. This is the preliminary study to delineate wheat farming zones of Turkey at national level by using remote sensing related application. The results are expected to contribute to the development of spatial and temporal land use mapping methods using remote sensing and geographic information based approaches.

## Material and Method

### Administrative Boundary Map

The national administrative map of Turkey was obtained from the Global Administrative Areas (GADM) web page (Anonymous 2016) which serves as a depository of spatial data of all world's administrative areas. The GADM database includes by country, various lower level subdivisions (boundaries of provinces, counties and municipalities) as ESRI shapefile format based on the WGS84 geographic coordinate system.

### CORINE Land Cover (CLC)

CORINE means Coordination of Information on the Environment which is a programme initiated by European Union aimed at gathering information relating environments such as air, water, soil, land cover etc. The CLC map describes land cover in 44 classes organized in three sub-levels and is presented as a cartographic product at a scale of 1:100,000. The CLC data was produced by visual interpretation of moderate and high resolution satellite imageries (Bossard et al., 2000) and has been gradually updated in 2000, 2006 and 2012. The CLC database is available in both vector and raster-based grid formats. In this study, the CLC 2000 map (Third level-Code: 211, Non-irrigated arable lands) was used to mask off agricultural areas, and thus to limit the crop mapping exercise to only those areas (Figure 1).

### Hyper-Temporal NDVI Data

NDVI hyper-temporal data contain time series of imagery, basically captured at a daily cycle; they contain information on the variability in vegetation condition in space and time. This study used a dataset of 180 MODIS Terra images spanning from 2004 to 2013. Each year dataset consisted of 18 two-week normalized differenced vegetation index (NDVI) composite based on 250 m MODIS data covering time leg of February to October which coincides the wheat growing season. NDVI is calculated from portions of the near-infrared and red wavelengths of the electromagnetic spectrum that differ in their absorptive and reflective characteristics by vegetation (Tucker and Sellers, 1986; Reed et al., 1994). For this reason, seasonal changes in vegetation conditions can be identified by NDVI time series dataset. The used dataset was downloaded from the Global Agriculture Monitoring (GLAM) Project archive (Anonymous, 2002) GLAM is a joint research project along with the partners of the US Department of Agriculture, the Foreign Agricultural Service (USDA FAS), the National Aeronautics and Space Administration (NASA), the University of Maryland (UMD) and the South Dakota State University (SDSU). Their initiative established a system aimed at crop-condition monitoring and production assessment throughout the world. The GLAM data archive provides 16-day composited NDVI time series of MODIS product generated from aerosol corrected land surface reflectance data and water masked using MOD44W (Carroll et al., 2009; Carroll et al., 2004) The resulting V2 WM2 geo-tiff images are greyscale images using the formula "ndvi_byte = (ndvi_raw × 200.0) + 50.0". Digital values below or equal 50 or higher than 250 indicate bad/missing data, while the value of 253 indicates water bodies. The remaining values represent the stretched NDVI values of current land covers.

The 10-year NDVI data was compiled into a multi-temporal NDVI stack image composed of sequentially ordered 180 NDVI layers. Though the NDVI data is corrected from some disturbances caused by cloud contamination, atmospheric variability and bi-directional effects, they still contain a considerable amount of noises that should be removed. A number of methods for reducing or removing the noises caused by haze and clouds could be found in literature (Viovy et al., 1992; Roerink et al., 2000; Swets et al., 1999). In this study, the stacked NDVI imagery was cleaned by employing the adaptive Savitzky-Golay smoothing filter through the TIMESAT package (Jönsson and Eklundh, 2004; Kim et al., 2014).

**Crop Statistics**

Agricultural statistics are compiled through official records of the Ministry of Food, Agriculture and Livestock (MFAL). Turkish Statistical Institute (TUIK) is yet the responsible body for the coordination of production and publication of official statistics and authorized body to publish and distribute the official statistics. All official statistical data is disseminated through internet as html pages, databases, or pdf files (http://www.turkstat.gov.tr) or in the form of publications or CD format. The agricultural statistics database in province basis and contains huge amount of data, numeric facts and figures such as reported crop areas, annual production, average yield and etc. of all crops grown throughout the country. This statistical data unfortunately does not provide the information of where exactly crops are grown, which prevents monitoring crop conditions and estimation of crop production. However, such data can be significant source of information for crop plans and rotations.

In this study, wheat acreages in hectares were extracted from database of TUIK’s agricultural statistics to build province-based matrix of a tabular data. The matrix data, as a dependent variable, was then correlated with area of NDVI classes (independent variable) through multiple linear regression analysis. It should be noted that the statistics are annual and cover the period of between 2004 and 2013 matching the time span of NDVI images.

**Reference Data for Validation**

To validate the produced wheat map, official records of the General Directorate of Farm Enterprises (TIGEM) were used. TIGEM is subsidiary establishment of Ministry of Food, Agriculture and Livestock of Turkey and responsible for producing certified seeds of various field crops (90% of wheat) in extensive areas across the country. Currently 17 TIGEM farms grow wheat on approximately 181000 hectares of arable land in every year.

Reference data used for this study included the list of wheat cultivated parcels, their locations and acreages in tabular form for 2013-2014 cropping season. Unfortunately reference data covered only 10 TIGEM farm lands due to lack of data (Figure 2). The reference data were obtained either by downloading from TIGEM’s homepage through the internet (www.tigem.gov.tr) or by hand directly from head offices of TIGEM farms. The tabular statistics unfortunately did not provide necessary spatial data to produce the GIS layer (as polygon) of wheat cropping parcels which are needed to calculate the fractions of wheat in the reference segments distributed throughout the lands of 10 TIGEM farms. These wheat fractions were then correlated with fractions of NDVI classes of wheat map for validation.

This study benefited from the MODIS NDVI image to produce GIS layer of wheat cropping parcels with series of process. According to recent practices and experts’ opinions, mid-May was maximum canopy cover period of wheat for TIGEM farm locations in study area. So, MODIS NDVI image of second week of May, 2014 coinciding the maximum canopy cover stage of wheat were applied to zonal statistics operation with feature zone data of available TIGEM farm borders to extract wheat representative pixels. It’s assumed that

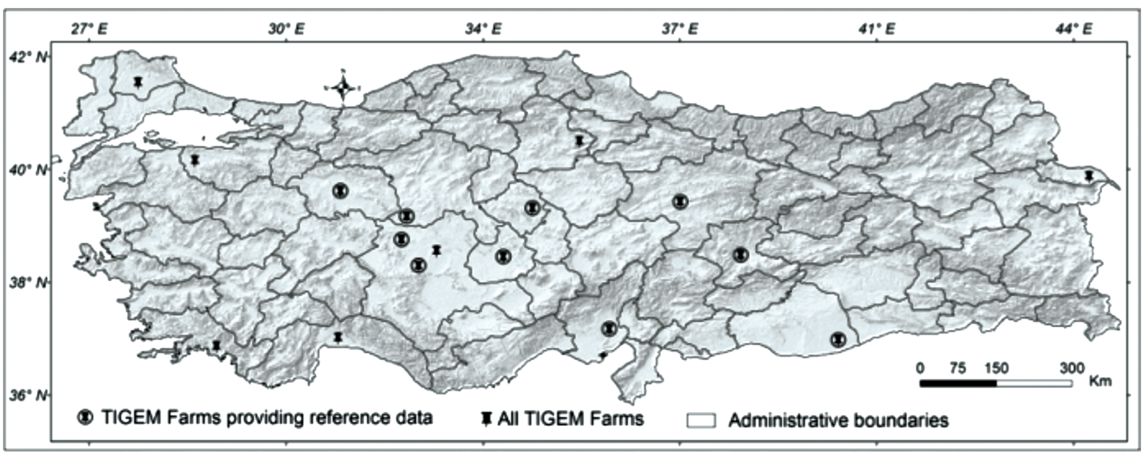


Figure 2. Locations of TIGEM farms

*Şekil 2. TİGEM çiftliklerinin konumları* even the pixel having minimum NDVI value in the boundary of recorded TIGEM parcels represent wheat cultivated spot because increasing canopy cover of wheat leads pixels having higher NDVI cell values. Applied zonal operation was forced to identify the minimum NDVI cell value among 10 TIGEM farm lands. NDVI value of 120 was the minimum one which indicates the extracted NDVI values of more than 120, the threshold value, represent wheat cropping parcels. The pixels with higher than the threshold NDVI value were converted to GIS polygons and then intersected with reference segments to calculate reference fractions of wheat.

The segments were constructed as rectangular grid (1 km$^2$) with Fishnet tool in GIS software and then applied to selection to get "reference segments" with the conditions that they should be completely within the border of TIGEM farmlands and each TIGEM Farm had at least 1 segment (Figure 3). In total, 150 segments were selected in 10 TIGEM farmlands to calculate fractions of existing wheat acreages and NDVI classes inside for validation purposes.

The method used in the study includes; unsupervised classification of multi-temporal NDVI data, overlay analysis, integrating crop

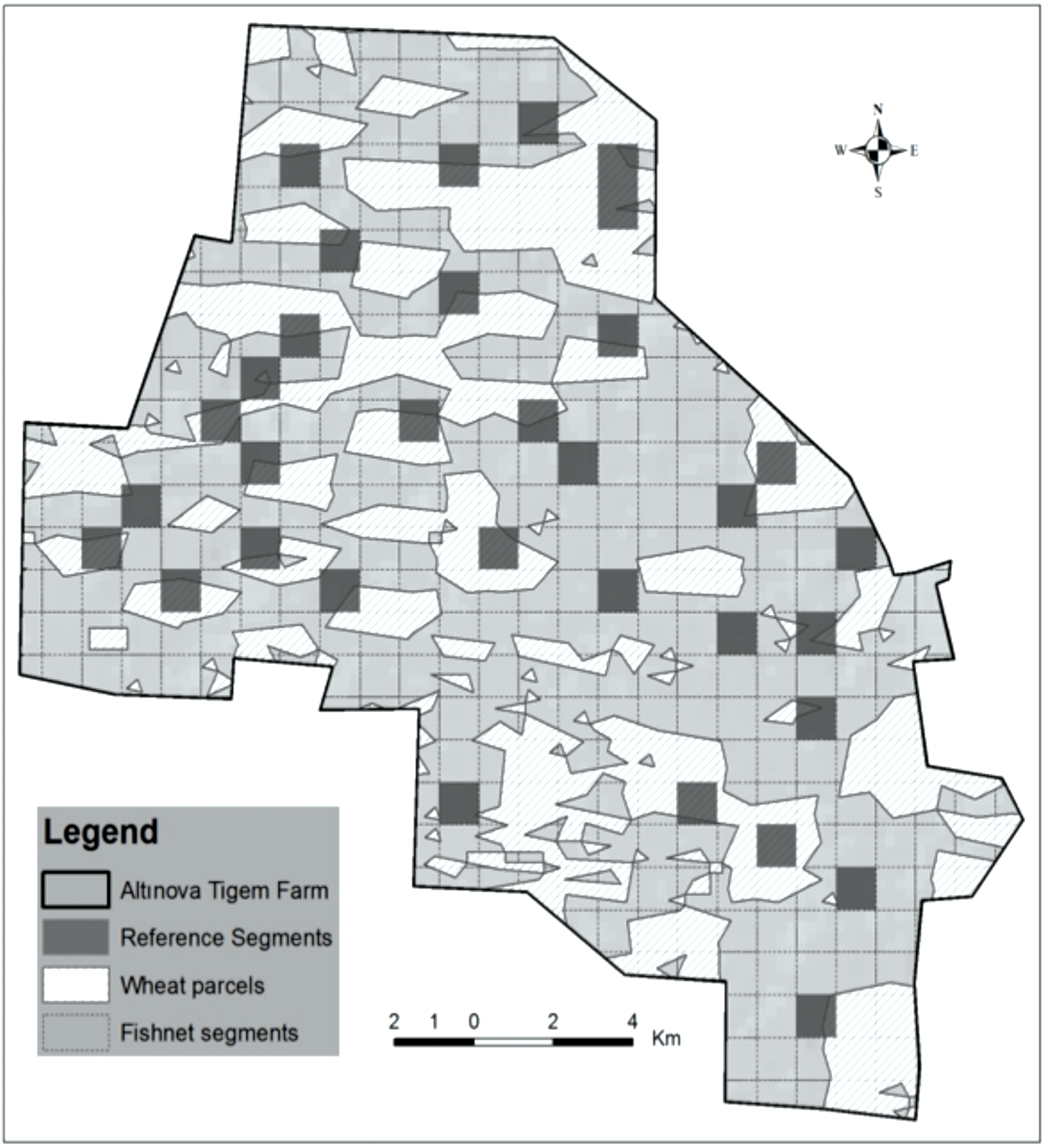


Figure 3. An example of selected reference segments for Altınova TIGEM Farm

*Şekil 3. Altınova TİGEM çiftliği için seçilmiş örnek referans segmentleri*

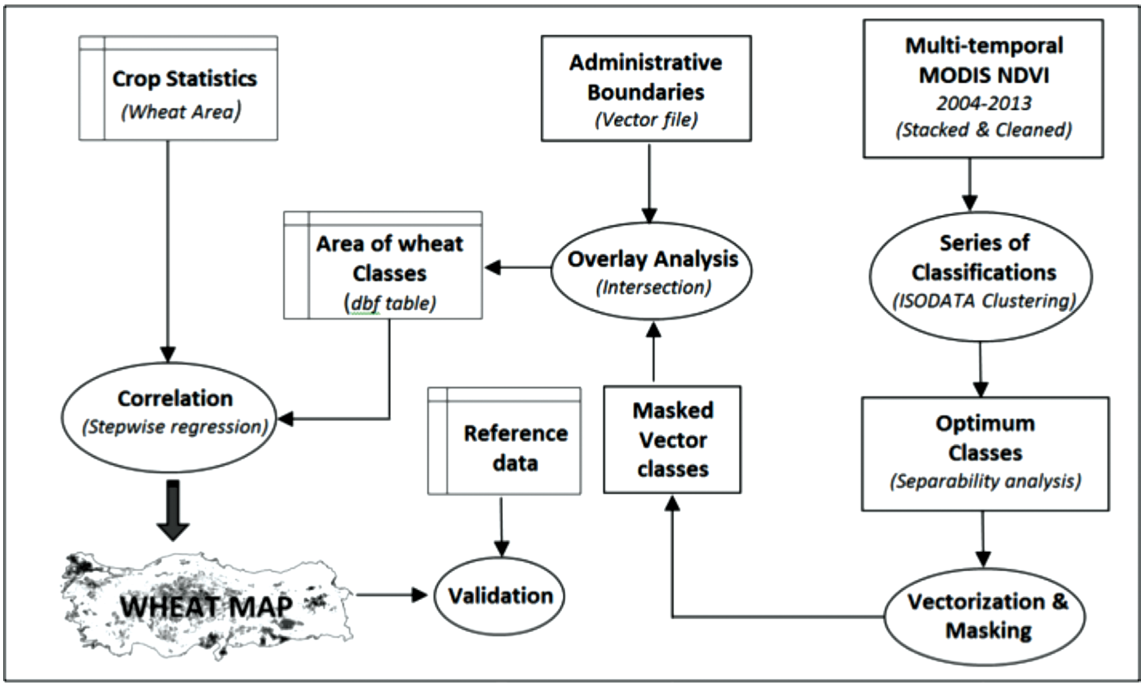


Figure 4. Graphical representation of methodology

*Şekil 4. Yöntemin grafiksel gösterimi*

statistics with tabular data of NDVI classes, producing final wheat map and validation of the map. The graphical representation of method is given in Figure 4.

**Classification**

Image classification is the process of clustering the pixels having similar spectral characteristics into same groups representing specific land covers. Identification of land cover types especially green covers by classifying original satellite imagery may be problematic due to continuum in vegetative development and phenologic aspects of crops. Using stacked time series NDVI data can solve this problem, because NDVI imagery has the ability to display the responses of growing vegetation throughout its phenological cycle. Classifying the hyper-temporal NDVI data (stacked NDVI) allows delineating green cover gradients, because the classification uses the space-time cube of NDVI-data to extract what it has to offer in differences, as classes.

Classification could be done by either supervised or unsupervised methods. Performing an unsupervised classification is simpler than a supervised classification, because the cluster signatures are automatically generated by the Iterative Self-Organizing Data Analysis (ISODATA) algorithm. The unsupervised classification uses not just 1 NDVI-image but series of images spaced at 10-day intervals and covering many years. Indeed, for a 1-date image, similar NDVI values will poorly differentiate land cover classes (and cropping systems), but time series NDVI data will be able to distinguish agro-ecological differences captured during wheat's vegetative phases such as emerging, development, flowering etc.

ERDAS software is used to run ISODATA classification algorithm. The maximum number of iterations was set to 50 and the convergence threshold was set to 1.0. Separate ISODATA runs were carried out to define 10–100 classes. In each run, the desired number of classes was produced and their separability statistics which indicated classes' uniqueness or divergence measure was calculated. The divergence measure of distance between cluster signatures was used to compare the various runs. Optimal run or the best number of representative NDVI classes was determined by looking at peak points in separability values. It should be noted that term of "the best number NDVI classes" is the number of classification run which includes the same number of NDVI classes as run's number denoted.

**Vectorization and Masking**

NDVI classes determined with the optimal run selection (explained in Section 4.1) were converted to polygons through "Raster to Polygon" tool of GIS software. The GIS polygons were then masked by using agricultural zones as defined by the CLC 2000 map in order to get major cultivation areas. Ultimately the masked map displays NDVI-classes as polygons of agricultural fields of whole country. The areas of the polygon NDVI-classes in each province were calculated through GIS analysis in order to correlate with those of official crop (wheat) area statistics of provinces.

**Overlay Analysis and Integration of NDVI Classes with Crop Statistics**

Overlay analysis which integrates spatial data with attribute data is one of the spatial GIS operations. Overlay analysis does this by combining information from one GIS layer with another GIS layer to derive an attribute for one of the layers. Here, intersect tool of overlay analysis methods was used. Administrative (province) boundaries and polygon NDVI-classes (by optimal run selection) were applied to intersection analysis so that NDVI classes were clipped by province boundaries. The resultant tabular data of clipped NDVI classes was in form of “dbase” table file containing class names as “grid code”, class area as hectare and province names. The area NDVI classes and recorded wheat area statistics of the provinces were merged into a matrix (Table 1) through which multiple stepwise regression was performed. It’s assumed that there is a direct relationship between area of NDVI classes representing agricultural fields in each province and official wheat areas statistics of same provinces. The relationship between these variables was estimated through the stepwise regression analysis

The matrix where the area of each NDVI-class per province and the wheat area statistics by same province were used as independent and dependent variables respectively was used in statistical software to estimate an linear function. This function below was explained through stepwise forward multiple regressions with no constant and coefficients constrained between 0.0 and 1.0 because the cropped area can neither be in negative nor more than 100% of the area (Khan et al., 2010).

$$Y=\sum_{i=1}^{n} bi^{*} Xi+\varepsilon i \qquad (1)$$

Where, Y is the average of a wheat crop area (ha) per province throughout 2004 - 2013, bi is the regression coefficient, xi is the average area (ha) of NDVI class i per province, n is the number of NDVI classes and εi is residual error. The regression coefficient in the equation is an estimate of the fractions (the percentage of wheat acreage per unit area) of each related NDVI classes which presumably represent wheat cropping areas of the country.

**Validation of Estimated Wheat Map**

Produced final wheat map was validated for the 2013 - 2014 growing season. Validation of map was done by correlating the reference fractions (RF) with the average estimated fractions (AEF) of segments (1x1 km) through linear regression.

The reference fractions are the area decimals of wheat in per segment and calculated (Equation 2) by dividing existing acreages of wheat parcels estimated from TIGEM statistics to total segment area (Figure 5).

$$RF_t=\frac{a_{wt}}{A_t} \qquad (2)$$

$RF_t$ is reference fraction of segment t. $a_{wt}$ is wheat area in *segment t*. $A_t$ is total area (1 km$^2$) of segment *t*.

The average estimated fractions of NDVI classes in the segments were calculated by as area weighted average of fractions. The formulation of AEF was below (Equation 3). Trial calculation for “*Segment t*” was given in Figure 5.

$$AEF_t=\sum \frac{a_{it}*f_{it}}{A_t}+\frac{a_{it}*f_{it}}{A_t}\dots \qquad (3)$$

Table 1. Matrix data of variables
*Çizelge 1. Değişkenlere ait matriks verisi*

| Dependent variable | | Independent variables (Area of NDVI Classes ha) | | | | | |
|---|---|---|---|---|---|---|---|
| Provinces | Official wheat acreages (ha) | Class1 | Class2 | Class3 | … | … | Class53 |
| A | 22876 | 4478 | 275 | 1155 | … | … | 54 |
| B | 15703 | 8822 | 3287 | 569 | … | … | 5789 |
| C | 23089 | 128 | 8714 | 189 | … | … | 1209 |
| … | … | … | … | … | … | … | … |
| … | … | … | … | … | … | … | … |

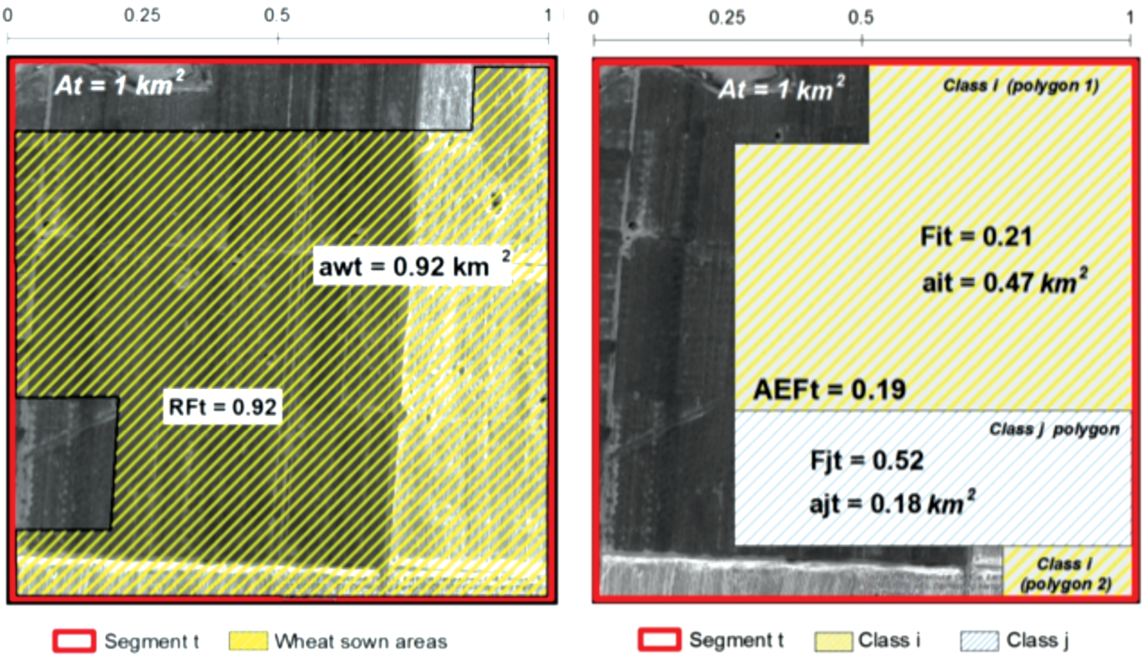


Figure 5. Sample calculation of reference fraction, RF (left) and average estimated fraction, AEF (right) for segment *t*

*Şekil 5. t segmenti için referans birim alan buğday oranı, RF (solda) ve ortalama tahmini birim alan buğday oranı, AEF (sağda) örnek hesaplaması*

$AEF_t$ is average estimated fraction of segment t. $a_{it}$ is wheat area of class *i* in segment t. $f_{it}$ is fraction of class *i* in segment *t*. $A_t$ is total area (1 km²) of segment *t*.

## Results and Discussion

### Optimal Run Selection

Minimum and average separability values obtained from series of (10-100) ISODATA classification runs were plotted to the number of classes generated by same runs (Figure 6). Horizontal axis of figure indicates the number of classification run, each of which generates same number of NDVI classes as its name referred. For instance, Run10 has 10 NDVI classes; Run53 has 53 classes so on. Figure 6 provides the choice of best number of run representing variability in multi-temporal NDVI data. It should be considered in determining optimal classification run that it included the lowest possible number of classes and should also offer the highest value of the minimum and average separability seen as unique peaks (De Bie et al., 2008).

To select the optimum run, Figure 6 was visually checked thoroughly to find the minimum and average separability values peaked in the same run. Even though there are many peaks on the minimum separability mean line, they are small local peaks and are not highly distinctive

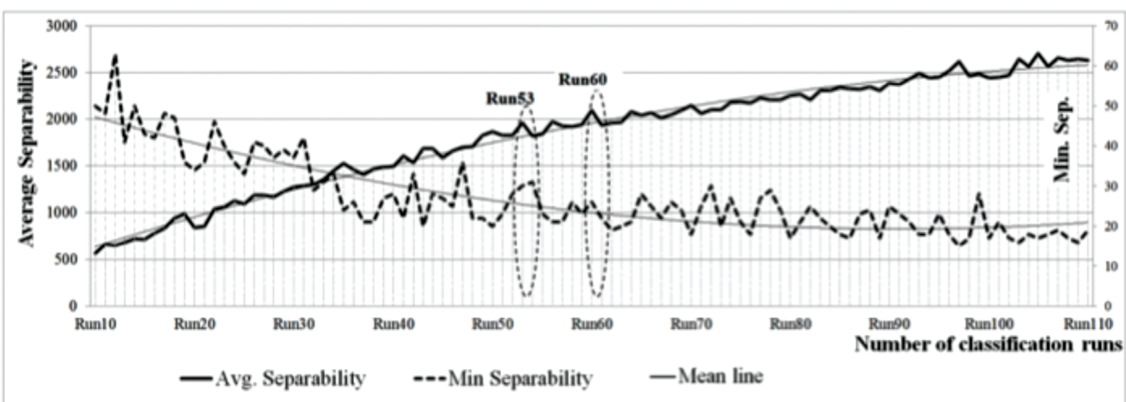


Figure 6. Average and minimum separability values from unsupervised classifications

*Şekil 6. Kontrolsüz sınıflama ile elde edilen ortalama ve en küçük ayrılabilme değerleri* making difficult to select optimal classification run. Those peaks do not show matching curves (peaks or bottoms) on the average separability mean line except two ones. According to graph below, Run53 and Run60 seem to be candidates for selection of optimal run, because only those runs have peaks on both average and minimum separability mean lines. Although Run60 has a peak at average separability line, it has not a clear peak for minimum separability. On the other hand, Run53 has distinctive peak on minimum separability line compared to Run60, which makes the Run53 the most reasonable choice. Assessing separability values reveals that the 53 NDVI-classes is statistically the best result in classification runs to represent the variations in multi-temporal NDVI data and to stratify the area into map units.

**Mapping**

The prepared matrix data was used in stepwise linear regression analysis where reported wheat area statistics (as dependent variable) and the area of 53 NDVI-classes (as independent variable) for provinces were correlated to find out the degree of relationship between variables. The stepwise regression run was continued until the adjusted $R^2$ did not increase more than 1% and coefficients were constrained between 0.0 and 1.0 in successive iterations. At the end of the 14$^{th}$ iteration, above conditions were met resulting in 12 classes with an adjusted $R^2$ of 95.8 and the coefficients ordered between 0.204 and 0.910 (Table 2).

The coefficients computed for 12 classes represent estimated fractions (EF) of wheat as percent per pixel and were used to produce wheat cultivated area map of the country (Figure 7).

The results of regression indicate that almost all variability was explained by NDVI classes with an adjusted $R^2$ of 95.8, which makes the NDVI data good predictor of final wheat map. The wheat map reflects quantitatively the wheat cropping area status from 2004 to 2013.

**Map Validation**

The reference fractions of wheat in segments were correlated with the fractions of NDVI-classes per segment through linear regression. The validation results showed that there was a good agreement between estimated fractions presented in NDVI based wheat map and reference fractions of wheat cultivated lands. The estimated wheat map explained approximately 69% of the total variability between sampled segments (Table 3). The variability between segments of a single NDVI class seems however to be high because defined units are internally heterogeneous. This is known as mixed parcels effects resulted from variability of landcover, which is quite applicable to this study. Even though the parcels in TIGEM farm lands are cultivated mostly with wheat, considerable amount of land is also reserved for barley cropping, which causes reference wheat parcels to have some barley pixels making the accuracy to decrease. Additional spatial information or higher resolution imagery could improve to capture this heterogeneity at the local level. Produced wheat map is therefore considered to have an amount of generalization and to be a small scale map.

Table 2. Results of stepwise linear regression analysis
*Çizelge 2. Aşamalı doğrusal regresyon analizi sonuçları*

| NDVI Class | Coefficients | t-value | Significance (%) |
|---|---|---|---|
| Class39 | 0.91 | 2.11 | 0.038 |
| Class30 | 0.86 | 4.53 | 0.000 |
| Class20 | 0.84 | 13.70 | 0.000 |
| Class46 | 0.58 | 2.31 | 0.024 |
| Class40 | 0.52 | 3.96 | 0.000 |
| Class27 | 0.49 | 3.21 | 0.002 |
| Class11 | 0.46 | 2.72 | 0.008 |
| Class18 | 0.35 | 4.47 | 0.000 |
| Class6 | 0.33 | 6.35 | 0.000 |
| Class50 | 0.32 | 2.22 | 0.030 |
| Class24 | 0.22 | 4.64 | 0.000 |
| Class42 | 0.20 | 3.52 | 0.001 |

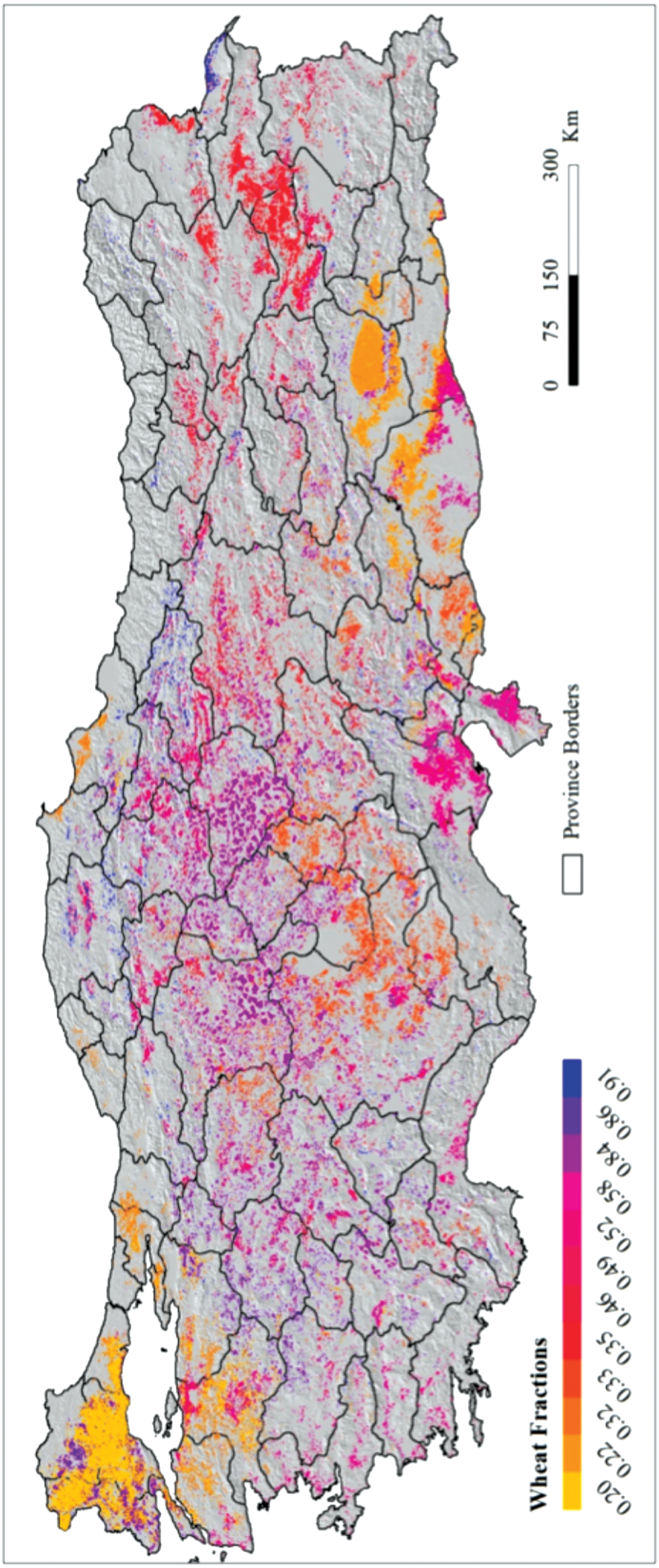


Figure 7. Estimated wheat growing areas by fractions per pixel
*Şekil 7. Buğday alanlarının birim alandaki yüzdelerinin hesaplanması için uygulanan aşamalı doğrusal regresyon analiz sonucu*

## Conclusions

Crop production and yield estimation are important and critical aims of monitoring agriculture to prepare effective management of policies. Before to reach these goals, there is a question of "where the crops are sown" that should be answered. Furthermore, timely and accurate estimation of cropped areas is also required for proper agricultural planning. In this context, this study focused on mapping wheat cultivation areas by using satellite imagery and crop statistics.

The method employed in this study suggests that recorded crop statistics can be used in combination with satellite derived NDVI data to identify and map the areas where wheat crop is grown. The method applied in this research makes use of analysis of NDVI imagery using unsupervised ISODATA clustering algorithm. The distinctive behaviors of crop can be distinguished from other vegetation types through analysis of their respective phenologies captured by temporal clustered hyper-temporal NDVI image (Guo et al. 2008). Indeed, the study results showed that NDVI time series has the ability to capture the crop phenologies and thus has a good relationship with the wheat cropping areas which represent interconnected NDVI classes having similar vegetation growth patterns during the cultivation season. The wheat map derived from regression analysis indicate that almost 96% of the variation in wheat cropping areas was explained by NDVI classes making MODIS NDVI data a good predictor. This good relationship also supports the notion that NDVI data comprises the combined effects of varying environmental attributes such as soil properties, temperature, rainfall, etc.). Similar studies also refer this relationship between NDVI and crops statistics data. Khan et al., 2010 and De Bie et al., 2008 found high correlations with adjusted $R^2$ of 98% and 98.8% respectively for winter wheat map produced by employing method of direct mapping with primary field data.

The validation of produced wheat maps, which is based on the average fractions of wheat in reference segments showed that almost 68.7% (adjusted $R^2$) of the variability was explained by NDVI class data. This indicates that the wheat map has large degree of generality as a result of low spatial resolution satellite imagery used, the pixel of which not only covers wheat cultivated areas but also other land cover types. When compared to study results (Verberien et al., 2008) where reported accuracy was found 39% with the method of linear mixture model, our validation accuracy is quite acceptable. In another study (Wardlow and Egbert, 2008) which used a decision tree classifier methodology to cluster NDVI time series data, reported classification accuracy was about 84% for the summer crop map. Even though the prepared wheat map wasn't so good due to moderate validation accuracy (69%), this work was basically a kind of prediction model at the end and so it could be incorporated into remote sensing based studies like crop acreage estimation, crop growth monitoring, crop pattern mapping and drought watch and early warning. It should be noted that accuracy of the prepared maps is also dependent on the quality of reported statistics. Such maps can be further improved and regularly updated by using higher spatial resolution hyper-temporal images and detailed spatial data.

It's concluded that NDVI data seems suitable and good predictor to map cropping areas in combination with ancillary data such as reported statistics. It's found the hyper-temporal NDVI data have a good correlation with spatial diversity of wheat crop grown indicating that NDVI alone explained a substantial portion of the variability in wheat areas. The employed method here can produce up-to-date crop maps with a cheap labor, less time consuming and reasonable degree of accuracy, which can help the policy makers to develop new strategies concerning management of natural resources, food supply and security issues.

## References

Anonymous, 2002. http://pekko.geog.umd. edu/usda/test. (Date of access 15.01.2016)

Anonymous, 2015. http://faostat3.fao.org. (Date of access, 14.02.2016)

Anonymous, 2016. http://www.gadm.org. (Date of access, 17.01.2016)

Bossard M., Feranec J., Otahel J., 2000. CORINE Land Cover technical guide-addendum 2000, Technical report No 40, European Environment Agency, Copenhagen

Brand S. and Malthus T.J., 2004. Evaluation of AVHRR NDVI for monitoring intra-annual and inter-annual vegetation dynamics in a cloudy environment (Scotland, UK). Proceedings of the XX$^{th}$ ISPRS Congress, Commission-II. Istanbul, Turkey, July 12–23, 2004 pp. 477-482

Carroll M., Townshend J., Dimiceli C., Noojipady P., Sohlberg R., 2009. A New Global Raster Water Mask at 250 Meter Resolution. International Journal of Digital Earth. (Volume 2 number 4)

Carroll M.L., Dimiceli C.M., Sohlberg R.A., Townshend J.R.G., 2004. 250 m MODIS Normalized Difference Vegetation Index, 250ndvi28920033435, Collection 4, University of Maryland, College Park, Maryland, Day 289, 2003

De Bie C.A.J.M., Khan M.R., Toxopeus A.G., Venus V., Skidmore A.K., 2008. Hypertemporal image analysis for crop mapping and change detection. Proceedings of the XXI Congress: Silk road for information from imagery: The International Society for Photogrammetry and Remote Sensing, 3-11 July, Beijing, China. Comm. VII, WG VII/5. Beijing: ISPRS, 2008. pp. 803-812

De Bie C.A.J.M., Khan M.R., Smakhtin V.U., Venus V., Weir M.J.C., Smaling E.M.A., 2011. Analysis of multi - temporal SPOT NDVI images for small - scale land - use mapping. International Journal of Remote Sensing, 32 (21):6673-6693

Campbell J.B. 1996. Introduction to Remote Sensing. 2nd edition. Guilford Press, New York, 622 p

Goward S.N. and Huemmrich K.F., 1992. Vegetation canopy PAR absorptance and the Normalized Difference Vegetation Index: an assessment using the SAIL model. Remote Sensing of Environment, 39: 119–140

Groten S.M.E. and Ocatre R., 2002. Monitoring the length of the growing season with NOAA. International Journal of Remote Sensing, 23(14): 1271-1318

Gumma M.K., Uppala D., Mohammed I.A., Whitbread A.M., Mohammed I.R., 2015. Mapping Direct Seeded Rice in Raichur District of Karnataka, India. Photogrammetric Engineering and Remote Sensing, 81(11):873-880

Guo W.Q., Yang T.B., Dai J.G., Shi L., Lu Z.Y., 2008. Vegetation cover changes and their relationship to climate variation in the source region of the Yellow River, China, 1990-2000. International Journal of Remote Sensing. 29: 2085-2103

Hill M.J. and Donald G.E., 2003. Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. Remote Sensing of Environment, 84: 367-384

Jansen L.J.M. and Di Gregorio A., 2003. Land-use data collection using the ‘‘land cover classification system’’ (LCCS): results from a case study in Kenya. Land Use Policy, 20 (2):131–148

Jönsson P. and Eklundh L., 2004. TIMESAT – a program for analyzing time-series of satellite sensor data. Computers & Geosciences, 30: 833–845

Kim S.R., Prasad A.K., El-Askary H., Lee W.K., Kwak D.A., Lee S.H., Kafatos M., 2014. Application of the Savitzky-Golay Filter to Land Cover Classification Using Temporal MODIS Vegetation Indices. Photogrammetric Engineering and Remote Sensing, 7(11):675-685

Khan M.R., De Bie C.A.J.M., Van Keulen H., Smaling E.M.A., Real R., 2010. Disaggregating and mapping crop statistics using hypertemporal remote sensing. International Journal of Applied Earth Observation and Geoinformation, 12: 36-46

Nguyen T.T.H., De Bie C.A.J.M., Amjad A., Smaling E.M., Chu T.H., 2012. Mapping the irrigated rice cropping patterns of the Mekong delta, Vietnam through hyper-temporal SPOT NDVI image analysis. International Journal of Remote Sensing, 33 (2):415–434

Oetter D.R., Cohen W.B., Berterretche M., Maiersperger T.K., Kennedy R.E., 2000. Land cover mapping in an agricultural setting using multi-seasonal Thematic Mapper data. Remote Sensing of Environment, 76: 139–155

Reed B.C., Brown J.F., Vander Zee D., Loveland T.R., Merchant J.W., Ohlen D.O., 1994. Measuring phenological variability from satellite imagery. Journal of Vegetation Science, 5(5):703–714

Roerink G.J., Menenti M., Verhoef W., 2000. Reconstructing cloud free NDVI composites using Fourier analysis of time series. International Journal of Remote Sensing, 21(9): 1911–1917

Shao Y., Lunetta R.S., Ediriwickrema J., Liames J.S., 2010. Mapping Cropland and Major Crop Types across the Great Lakes Basin using MODIS-NDVI Data. Photogrammetric Engineering and Remote Sensing, 76(1):73-84

Swets D.L., Reed B.C., Rowland J.R., Marko S.E., 1999. A weighted least-squares approach to temporal NDVI smoothing. Proceedings of the ASPRS Annual Conference, 21-27 May 1999; Portland, Oregon, USA, pp. 526-536

Tucker C.J. and Sellers P.J., 1986. Satellite remote sensing of primary production. International Journal of Remote Sensing, 7(11):1395–1416

Unal E., Mermer A., Yıldız H., 2014. Assessment of rangeland vegetation condition from time series NDVI data. Journal of Field Crops Central Research Institute, 23 (1):14-21

Wardlow B.D. and Egbert S.L., 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

Verberien S., Eerens H., Piccard I., Bauwens I. and Van Orshoven J., 2008. Sub-pixel classification of SPOT-VEGETATION time series for the assessment of regional crop areas in Belgium. International Journal of Applied Earth Observation and Geoinformation, 10: 486–497

Viovy N., Arino O. and Belward A., 1992. The best index slope extraction (BISE): A method for reducing noise in NDVI time-series. International Journal of Remote Sensing, 13(8):1585–1590

Wood S., Sebastian K. and Scherr S.J., 2000. Pilot Analysis of Global Ecosystems—Agroecosystems. World Resources Institute, Washington, DC, USA: International Food Policy Research Institute, pp. 125