ASSESSMENT OF CROPLAND USE IN THE AKKOL DISTRICT OF KAZAKHSTAN USING MODIS NDVI TIME-SERIES DATA

Aigul Bekbayeva (1), Farabi Yermekov (1), SholpanValiyeva (1), Jay Sagin (2)

[1]S. Seifullin Kazakh Agro Technical University, 010000, 62, Zhenis Avenue, Nur-Sultan, Kazakhstan
[2]Nazarbayev University, 010000, 53, Qabanbay Batyr Avenue, Nur-Sultan, Kazakhstan

Email: bekbaevaaigul@gmail.com, f.yermekov@gmail.com, sholpanvaliyeva@gmail.com, zhanay.sagintayev@nu.edu.kz

ABSTRACT

Monitoring land use efficiency of Kazakhstan territory, which has an area of about 3 million km,[2] is a complicated task. For the northernagricultural territories, we have applied and combined Remote Sensing technologies with field work investigationsfor calibration and verification purposes. Kazakhstan Land law allows the local farmers to use the land freely or fora very low fee, but these lands should be used for agricultural purposes. If the land is unused for crop planting over two consecutive years, the government can retrieve the land from the farmer. However, the local authorities have difficulty monitoring Kazakhstan's vast territory with proper mapping visualization. In this research, we analyze the cropland usage and map the used and unused crop fields through applying MODIS Normalized Differenced Vegetation Index (NDVI) time series.For 14 years, from 2004 to 2018, the combined remote sensing and field work data were processed. As well, the digitization of fields was conducted, and datasets were processed for the vegetation periods.Thesefour land cover groups were identified: croplands, unused, fallow, and hayfields. The calibration and validation of the remote sensing data was conducted through observations during site field visits. The overall accuracy was 92% and it was concluded that 87% of the territory was not used for agricultural purposes during that time period. The limitations of the research are related to the MODIS spatial resolution which only has a pixel size of 250 m. Moreover, because of Kazakhstan's large territory, climate is different in the different regions. Thus, data should be analyzed and reported on through a region to region basis instead of making an overall conclusion for land use and policy implementation.

KEY WORDS: MODIS, croplands, land use, NDVI

INTRODUCTION

Agriculture is one of the main sectors which contribute to a country's gross national product. According to the amount of arable land, Kazakhstan is in the top 20 countries in the world (Nationmaster.com, 2019). Farms in the country are divided into 3 groups: agricultural enterprises, private farms and subsidiary household plots. Private farms and agriculture enterprises generate outputs mainly for sale rather than for private consumption. The average size of enterprises is 3000 ha, private farms – less than 1000 ha, and household plots – 0.15 ha. Kazakhstan is an essential producer and exporter of high-quality wheat (Lindeman, 2010).

Monitoring of land use in Kazakhstan is very important. Article 94 of the country's Land Law states that, if the land plot intended for conducting farm or agricultural production, is not used for its intended purpose for two consecutive years from the date of initial detection of the fact of non-use, such land plot shall be subject to compulsory withdrawal (Information system PARAGRAPH, 2019).

Due to the large size of arable lands of enterprises and private farms it is difficult for government authorities to monitor the use of crop lands. Each year, the Kazakhstan farmers report on the use of their lands, but usually it is complicated to verify the reliabilityof the reported information. Remote sensing technologies may significantly improve the automatization of the monitoring process. However, at this moment, in Kazakhstan no technology to monitor the use of land has been developed. . Thus, the mapping of croplands using remote sensor data has significant potential.

The high-temporal resolution of the MODIS satellites allows for the recording of the vegetation dynamics at relatively high spatial resolution. Only a few studies have used these data to identify unused land - Central Great Plains of the United States (Wardlow& Egbert, 2008), and farmland abandonment and recultivation across Europe (Estel et al., 2015). As well, global fallow land was calculated with an integration of MODIS land-cover data to the MICRA 2000 model (Portmann, Siebert, &Döll, 2010; Siebert et al., 2010).

The aim of this research was to map and verify the possibility of using this freely available MODIS NDVI time series to identify unused lands and to apply a decision tree classification algorithm. Additionally, the simple

unsupervised classification algorithm was compared to the decision tree classification analysis. The possibility of applying simple k-means classification was assessed.

STUDY AREA

Kazakhstan consists of 14 administrative regions. About 75 percent of the country's wheat is produced in three regions in north-central Kazakhstan: Akmola, Kostanay and North Kazakhstan. The study area for this research is located in the Akmola region (Figures 1 and 2). The Akmola region has 17 administrative districts and Akkol is one.

The overall size of the study area is 5690^2 ha (Figure 3). Topographically, the study area is located on the Kokshetau upland. Its lowest height is 315 and the highest pointis 329 meters.The calculated average slope is 0.5 degree. The climate is designated as sharply continental. In the forest-steppe zone, an average of 300-400 mm of precipitation falls per year. The summer vegetation period runs from May to October. The winter is extremely cold, so vegetation only grows in the summer.

The type of soil in the study area is the southern chernozems. This type of soil has a high natural fertility and is widely used in agriculture. With the southern chernozems, it is possible to cultivate wheat, sugar beets, sunflowers, corn, and legumes. For the cultivation of a number of crops there is a lack of moisture, so snow retention, moisture-charging irrigation and other measures aimed at the accumulation and preservation of moisture in the soil are particularly important (Ecosystema.ru, 2019).

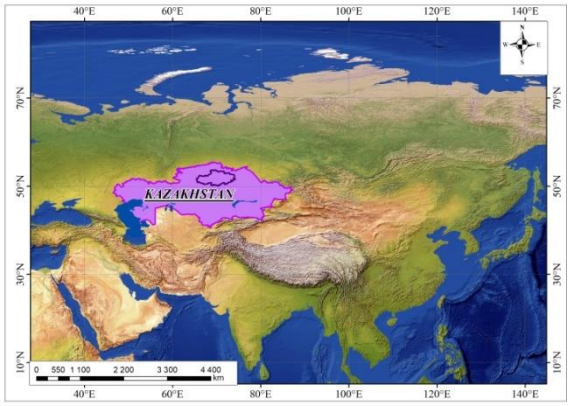


Figure1. Study area

Figure2. Study area

Figure 3. The fields of the study area.

DATA DESCRIPTION

Satellite images data

The NASA Moderate Resolution Imaging Spectroradiometer (MODIS) instrument is well suited to monitor agricultural systems. In our study, we used MODIS NDVI time series MOD13Q1 and MYD13Q1 of Terra and Aqua satellites from 2014 to 2018 in the vegetation period (from May to October) at a spatial resolution of about 250 m.

The NDVI is a simple indicator of an area with dense vegetation whose value tends to be positive (from 0.3 to 1) and clouds and snow fields will be characterized by negative values of this index. The value of index for open soil area is in the range of 0-0.2.

All MODIS date were downloaded from the USGS web-portal (https://earthexplorer.usgs.gov/)

The fields' datasets

The fields of the study area were digitalized using high-resolution PlanetScope satellite images from Planet web-site (Planet, 2019). The PlanetScope has 3 meter spatial resolution and it is possible to download images using a free 14 day trial account. Applying MODIS images for fields digitalizing will give higher inaccuracies in comparison to high spatial resolution images. However, there is an option to pay for the PlanetScope daily images. Additionally, the field borders were verified on the official web-site of the Department of Land use Resources of the Ministry of Agriculture of Kazakhstan (Aisgzk.kz, 2019).

METHODOLOGY

For the purpose of this research we defined 4 classesof agricultural land use (Figure 4):

- fallow land is usually included in crop rotation. The fallow field is maintained in a weed-free form with good soil conditions as a result of mechanical tillage and the use of herbicides.
- hayfield land is usually mowed at least once in the vegetation period. The hayfields include annual and perennial grasses. Annual herbs are grown during one growing season, perennials — for 2-3 years in field crop rotations, fodder for 5-7 years, and up to 10 years or more outside the crop rotation. Usually, during the vegetation season the hayfields are mowed between the middle of June to end of August.
- croplands are usually sown from May to the beginning of June depending on climate of the year. The sprout stage comes in approximately 2 weeks after sowing.The vegetation peak is usually reached by the end of July or the beginning of August. Then, the harvesting starts at the end of August.
- unused land was not sowed and mowed in the vegetation period. Also, any tillage and use of herbicide was not applied, so weeds are completely spread throughout the fields.

The NDVI value for fallow land during the vegetation season does not exceed 0.5. In general, it remains stable without any sharp increases and decreases. Also, NDVI value for fallow land should not increase. If this does happen, herbicides should be applied to prevent weed growth and to lower the NDVI.

Conversely, if fields have a sharp decrease in NDVI, this illustrates that the field vegetation has been moved or harvested. Consequently, it is possible to identify a hayfield. When the vegetation is at the peak, usually hayfields are mowed, which may happen twice a year in the region. If the hayfield is an annual herb in the beginning of the season, the NDVI value will be at a lower level, due to the fact that annual herbs are sowed and might be used in one season. If the hayfield is a perennial herb, the NDVI value may be already higher in the beginning.

In relation to crop land, the NDVI value grows gradually during the season and reaches its peak in August usually.

For unused land, vegetation is present during the season and any tillage work is not applied on the fields. In the beginning of the season, the NDVI usually already has a relatively high value. There are nosharp decreases in the NDVI due to the absence of mowing and harvesting

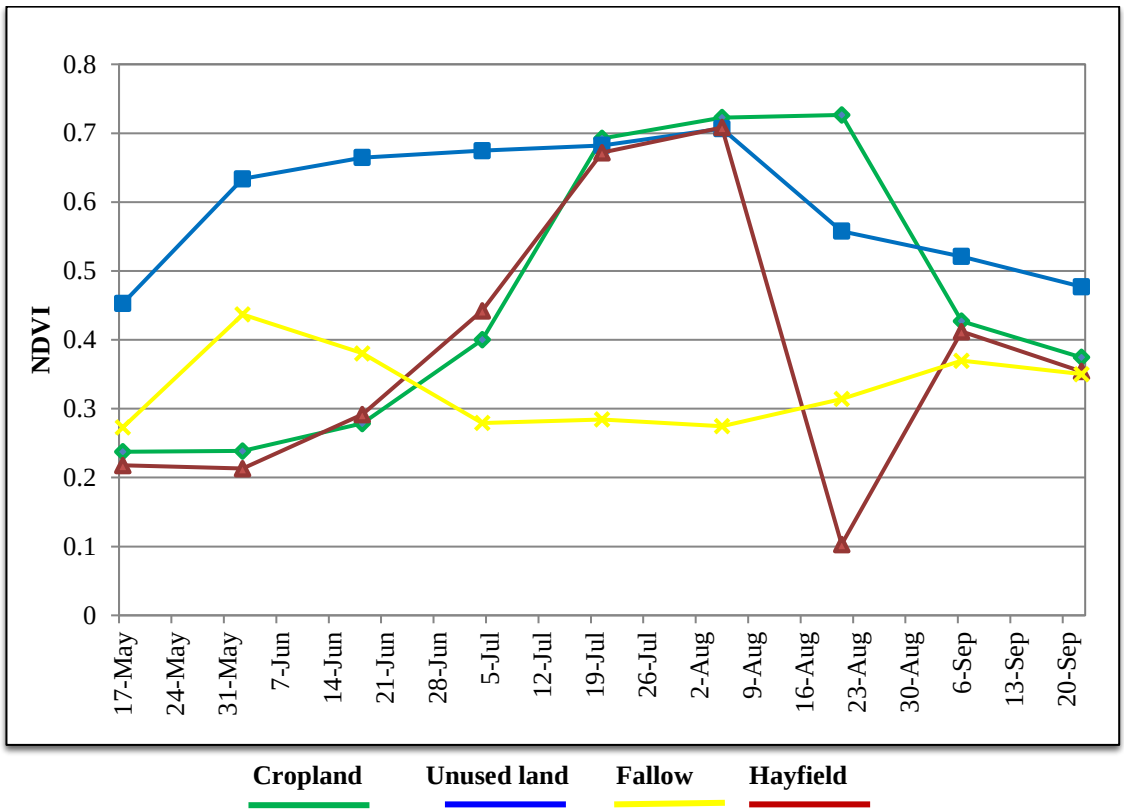


Figure 4. NDVI profiles of land use types (average trend for each year)

The general flowchart of our research is illustrated in Figure 5.

At the first stage of research, the datasets were obtained. Satellite images were downloaded (through web-sites https://www.planet.com/, https://earthexplorer.usgs.gov/) and the digitalization of the fields was accomplished using ArcGIS software.

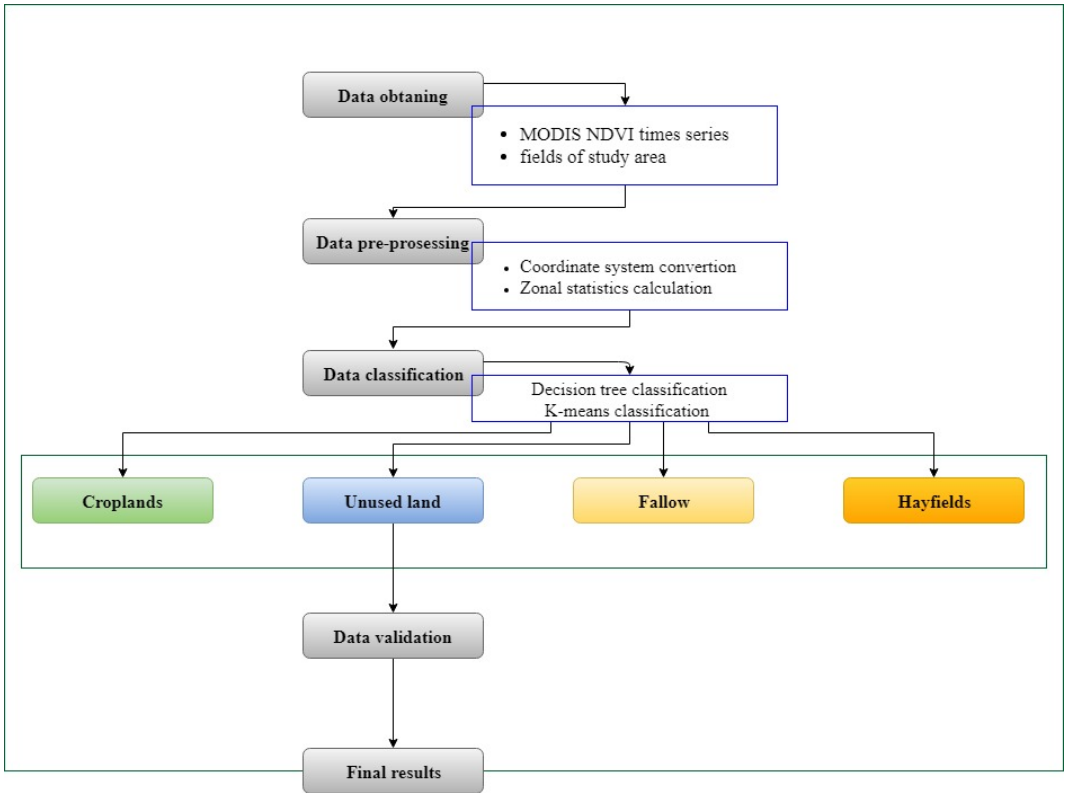


Figure 5. Research flowchart

Coordinate systems conversion

In order to handle the MODIS product, the MSTK toolkit was used with ENVI software. By applying this toolkit, it was possible to convert sinusoidal projection of HDF files into Geographic projection. The MODIS conversion toolkit was designed by D. White and it is free for downloading.

Zonal statistics

The zonal statistictool in ArcGIS software allows calculating statistics for each zone which is defined by a zone dataset based on values from another dataset (a value raster). One output value is calculated for each zone of the input zone dataset. The NDVI values for each field on the each date were calculated (Desktop.arcgis.com, 2019).

Decision tree classification

Decision tree analysis is a tree where each node represents an attribute, each link represents a decision and each leaf represents an outcome (categorical or continues value). The decision tree is a type of supervised classification, when there are some training datasets of labeled classes.

K-meansclassification

K-means classification is the simplest algorithm for classification.K-means algorithm identifies *k* number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible. K-means clustering is a methodologyoften used to automatically separate a data set into k groups (MacQueen, 1967). K-means classification is a type of unsupervised classification when there are no training datasets. This classification was tested for the possibility of simplifying the classification for monitoring significantly larger territories.

Validation and visual assessment

In order to validate the land use classification, the extensive validation dataset was collected: 1) ground based points; 2) available datasets of satellite images; and 3) field datasets from farms located in North Kazakhstan region. The most important part of the research was to validate the use of land in the study area. In 2017-2018, there were several verifications completed on the fields covering our entire study area. In addition, NDVI values were calculated using all the available satellite images, such as Landsat 7, 8, Sentinel 2, PlanetScope, andMODIS daily datasets. From the high resolution daily images the tillage of the fields was verified. A ploughed soil was identified from the multispectral images.In addition, the field's datasets with the type of crops were provided by the farm LLP "SXOS" located in the North Kazakhstan region.

RESULTS AND CONCLUSION

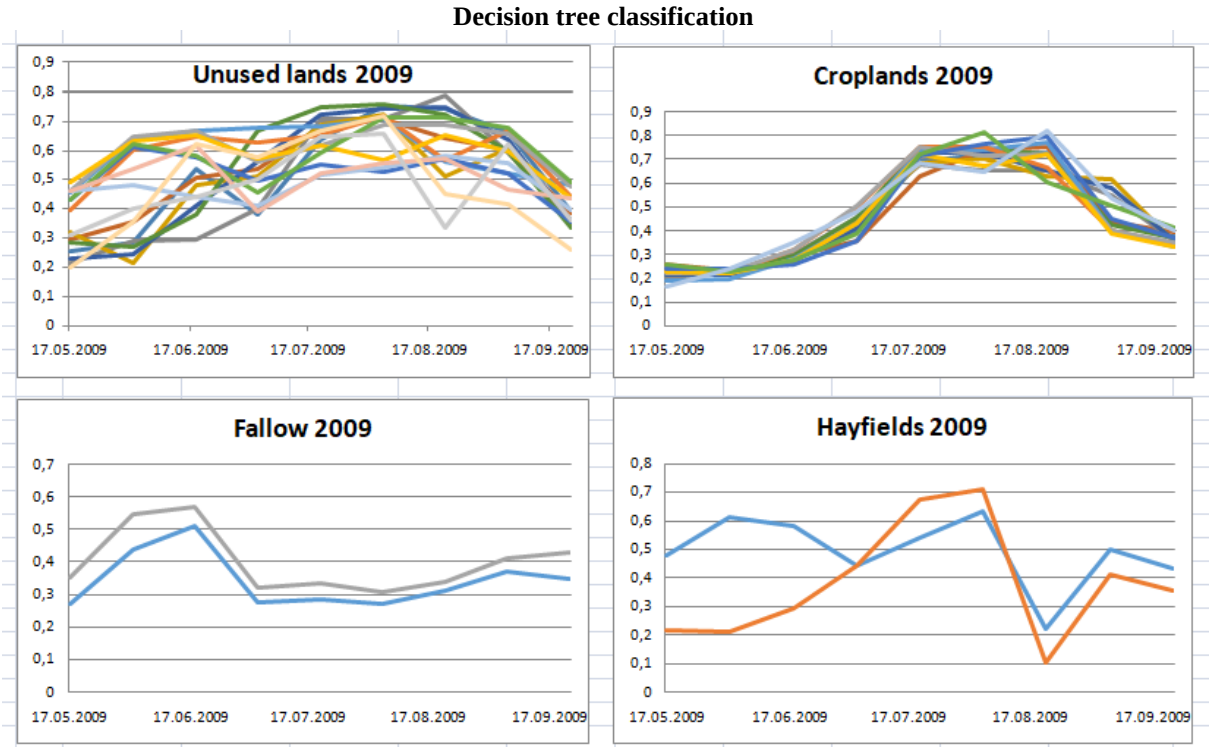
In total, as it can be seen from Table 1, 87% of the potential field area in the 16 years from 2004 to 2018 was not used. It is an extremely high percent of unused land in the agricultural area. Hayfields and fallow are included in the crop rotation process, so these lands belong to the category of used lands. However, about 60% of the used lands are hayfields which are probably perennial herbs. The perennial herbs might be sowed once in 5-10 years, which leads to the conclusion that the farm has been mowed infrequently. Therefore, the land was not used extensively.

N	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1	2	2	2	2	2	2	2	2	2	2	2	2	4	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	2	2	2	2	2	1	2	2	2	2	4	2	2	2	2
4	2	2	2	4	2	1	2	2	2	2	2	2	2	2	2
5	2	4	2	2	3	2	2	2	2	2	4	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2	4	2	2	2	2
8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	2	2	2	2	2	2	2	2	2	2	4	2	2	2	4
10	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
11	2	2	1	4	2	1	2	2	2	2	2	2	2	2	2
12	2	4	1	4	2	1	3	2	2	2	2	2	2	2	2
13	2	2	2	4	2	2	2	2	2	2	2	2	2	2	2
14	2	2	2	4	2	1	2	2	2	2	2	2	2	2	2
15	2	2	2	4	2	1	2	2	2	2	1	4	2	2	2
16	2	2	3	4	2	1	2	2	2	2	2	4	2	2	2
17	2	4	2	4	2	4	2	2	2	2	2	4	2	2	2

18	2	2	2	4	2	1	2	2	2	2	2	2	2	2	2
19	2	2	2	4	2	1	2	2	2	2	2	2	2	2	2
20	2	4	2	4	2	1	2	2	2	2	2	3	2	2	2
21	2	4	2	2	2	1	2	2	2	2	2	2	2	2	2
22	2	4	2	2	2	1	2	2	2	2	2	2	2	2	2
23	3	2	4	4	3	2	3	2	2	2	1	2	4	2	2
24	2	2	3	4	2	1	2	2	2	2	2	2	2	2	2
25	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
26	2	2	4	2	2	4	2	2	2	2	2	2	2	2	2
27	2	2	4	4	2	2	2	2	2	2	2	2	2	2	2
28	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
29	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
30	2	2	2	4	2	2	2	2	2	2	2	4	3	2	1
31	2	2	2	4	2	3	2	2	2	2	2	2	2	2	1
32	2	2	2	4	2	3	2	2	2	2	2	2	2	2	2
33	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

1	2	3	4
Cropland	Unused land	Fallow	Hayfield

Table 1. Overall land use classification of the study area



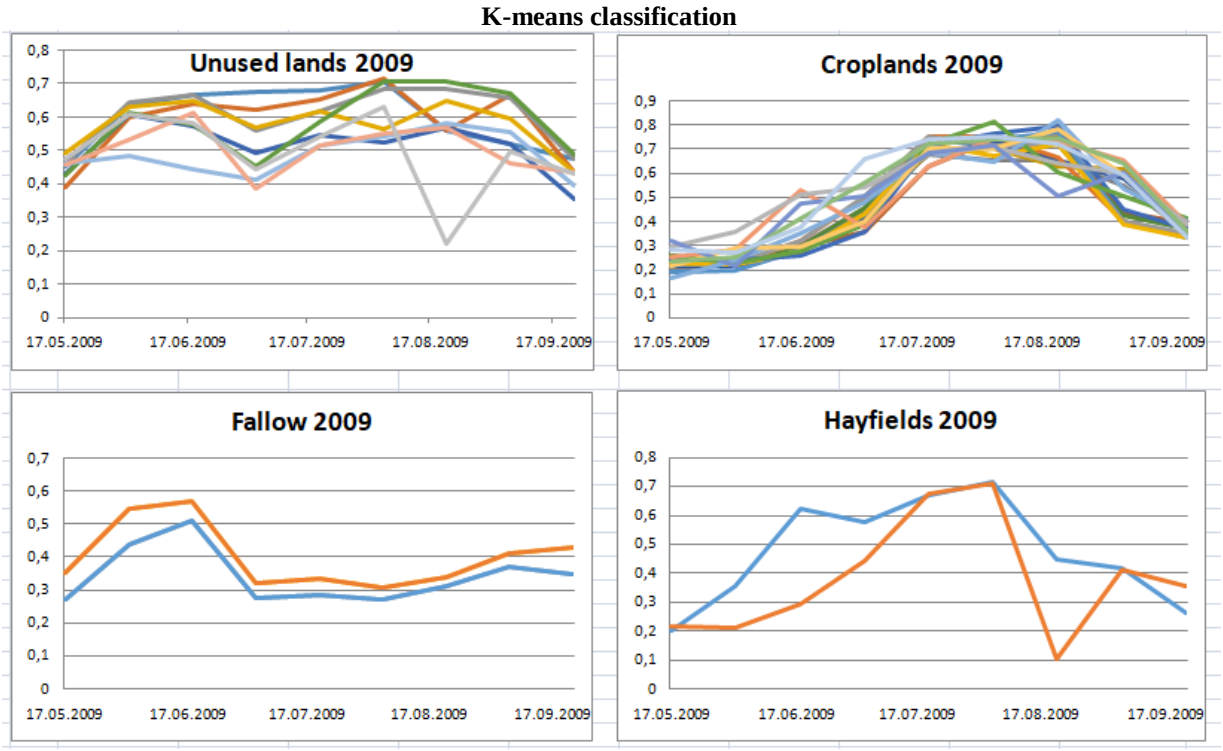


Figure 6. Results of classification methods comparison. Only classifications in 2009 are shown for clarify

As for the comparison of unsupervised and supervised methodologies, the decision tree performs slightly better. Its accuracy was 92%. It can be seen from Figure 6 that in K-means classification, the fields are classified differently. It seems that having training dataset and use of supervised classification produces more precise results.

The high percent of supervised classification accuracy confirms the possibility of applying MODIS NDVI time series to assess land use of agricultural fields. However, there is a limitation when using them on the fields that have an area that is smaller than the pixel size of images. In this case, high resolution images might be applied. Due to the large size of fields in Kazakhstan, the applied monitoring methodology of land use seems to be appropriate.

REFERENCES

Aisgzk.kz. (2019). *Департамент АИС ГЗК и информационной безопасности*. [online] Available at: http://www.aisgzk.kz/aisgzk/ru [Accessed 28 Aug. 2019].

Desktop.arcgis.com. (2019). *How Zonal Statistics works—Help | ArcGIS for Desktop*. [online] Available at: http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/h-how-zonal-statistics-works.htm [Accessed 29 Aug. 2019].

Ecosystema.ru. (2019). *Экологический центр Экосистема*. [online] Available at: http://ecosystema.ru/08nature/soil/098.htm [Accessed 27 Aug. 2019].

Estel, S., Kuemmerle, T., Alcántara, C., Levers, C., Prishchepov, A. and Hostert, P. (2015). Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. *Remote Sensing of Environment*, 163, pp.312-325.

InformationsystemPARAGRAPH. (2019). Земельный кодекс Республики Казахстан от 20 июня 2003 года № 442-II (с изменениями и дополнениями по состоянию на 03.04.2019 г.). [online] Available at: https://online.zakon.kz/Document/?doc_id=1040583#pos=2207;-28 [Accessed 27 Aug. 2019].

Lindeman, M. (2010). *Kazakhstan Agricultural Overview*. IPAD's Agricultural Production. [online] Washington: USDA. Available at: https://ipad.fas.usda.gov/highlights/2010/01/kaz_19jan2010/ [Accessed 27 Aug. 2019].

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Symposium on Math, Statistics, and Probability* (pp. 281–297). Berkeley, CA: University of California Press.

Nationmaster.com. (2019). *Countries Compared by Agriculture > Arable land > Hectares. International Statistics at NationMaster.com*. [online] Available at: https://www.nationmaster.com/country-info/stats/Agriculture/Arable-land/Hectares [Accessed 23 Aug. 2019].

Planet. (2019). *Daily Satellite Imagery and Insights*. [online] Available at: https://www.planet.com/ [Accessed 28 Aug. 2019].

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 Biogeochemical Cycles*, 24 GB1011.

Siebert, S., Portmann, F.T., &Döll, P. (2010). Global patterns of cropland use intensity. Remote Sensing, 2, 1625–1643.

Wardlow, B. and Egbert, S. (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(3), pp.1096-1116.