Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 26, 2014, Sofia

CROP AREA ESTIMATES BASED ON PER-PIXEL SUPERVISED CLASSIFICATION ON EO-1 ALI IMAGE FOR A TEST SITE IN NORTHEAST BULGARIA

Vassil Vassilev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: vassilev_vas@space.bas.bg

Abstract

The purpose of this article is to investigate the crop area estimates based on per-pixel classification for Zhiten test site situated in Northeast Bulgaria. The chosen satellite image is acquired from multispectral EO-1 ALI sensor on 09.07.2011. The methodology of this article includes the following working stages: 1) applying arable mask from CORINE 2006 land-cover database; 2) conducting per-pixel supervised classification using the maximum likelihood classifier (MLC) algorithm for crop identification; 3) applying accuracy assessment tool in ERDAS Imagine and deriving accuracy totals and kappa statistics; 4) calculating crop area estimated based on pixel-counting technique. The overall classification accuracy for the EO-1 ALI image is 96.24% and overall kappa statistics is 0.9397. The high overall accuracy can be accomplished after carefully chosing representative training samples. The beginning of July can be a perfect time to separate spring crops (sunflower and maize cultivars) and assess crop area estimated for all cultivated crops based on per-pixel classification with overall accuracy above 95%. The down side of this study is that at the time of image acquisition the winter crops are difficult to separate, since they are already with very low reflectance values.

1. Introduction[1]

Currently a major challenge in agricultural applications is forecasting crop production using low and coarse resolution satellite images,

[1] Abreviations used:
MARS – Monitoring Agriculture with Remote Sensing
LACIE – Large Area Crop Inventory Experiment
CITARS – Crop Identification Technology Assessment for Remote Sensing
NDVI – Normalized Difference Vegetation Index

while for high resolution (HR) satellite images one of the hottest topics is controlling area-based subsidies and applying precission agriculture practices amongst others. Satellite Remote Sensing (RS) provides synoptic, objective and relatively homogeneous data which can be geographically and temporally registered. Therefore, RS is an efficient tool for providing standard, high quality information on agriculture, evenly over broad-scale territories. The Monitoring Agriculture with Remote Sensing (MARS) project of the European Union was established in order to define and demonstrate how RS can be used operationally to supplement, interpret, and standartize agricultural statistical data provided by conventional techniques [1,2]. Satellite RS techniques have been proven to be effective and useful in broad-scale agricultural surveys such as: Large Area Crop Inventory Experiment (LACIE) project in the USA and MARS project in Europe [3]. Additonally, experiments from LACIE and Crop Identification Technology Assessment for Remote Sensing (CITARS) projects have also been conducted to demonstrate the capabilities of RS for crop inventory and forecasting [4.5].

Vegetation types can be characterized using their seasonal variations in the Normalized Difference Vegetation Index (NDVI) time-series, which include a series of images, acquired on weekly or decadal basis and showing the crop development dynamics. For example, the winter wheat phenophases like tillering and flowering as well as harvest, can be successfully identified using sensors with different spatial resolution in various band combinations and severe ground surveys, including collecting information for defining training samples for the supervised classification [6]. A number of different methods have been developed during the last two decades to discriminate crop types using data from NDVI and from the Advanced Very High-Resolution Radiometer (AVHRR). These methods employ a variety of different approaches including temporal profiles of crop phenology manifested in the NDVI [7,8], and classification of multi-temporal data [9,10], which can be applied on variously managed crop areas worldwide.

Crop identification during the growing season is a major challenge for forecasting crop production as well as for controlling area-based subsidies in the European Union member states [5]. The basis for separation one crop from another is the supposition that each crop species has a unique visual appearance and spectral signature on the image. However, separating these species may be difficult because of variations in soil properties, fertilization, pest conditions, irrigation practices, planting dates, as well as intercropping, and tillage practices [11], all of which can be adopted in precision farming using high quality satellite images. Thus, high-resolution satellite images are the key to the above mention difficulties.

The purpose of this case study is to investigate crop area estimates based on per-pixel classification for Zhiten test site situated in Northeast Bulgaria, and it includes the following tasks:

(1) Applying arable mask from CORINE 2006 land-cover database;

(2) Conducting per-pixel supervised classification using the maximum likelihood classifier (MLC) algorithm for crop identification;

(3) Applying accuracy assessment tool in ERDAS Imagine and deriving accuracy totals and kappa statistics;

(4) Calculating crop area estimated based on pixel-counting technique.

2. Materials, methods and data used

The study area – part of Zhiten test site is situated in North-East Bulgaria. The area represents intensively cultivated area sowed mostly with cereals and sunflower. This territory is one of the main agricultural regions of the country. The area is part of the European-continental climatic province of the temperate climatic belt. Climate is moderately warm with no distinctive dry season. Mean annual air temperature is 10.2°C. The main soil types are chernozems from the zonal ones and fluvisols from the azonal types.

The major cultivated winter crops (wheat and oilseed rape) and spring crops (sunflower and maize) were investigated in the present case study.

During the 2010-2011 agricultural season and in particular in the period between March–July 2011 four exhaustive field surveys were carried out and ground data was collected and organized in a GIS geodatabase. Field data was collected in the framework of a project financed by the Belgian Federal Science Policy Office (BELSPO) under the PROBA-V Preparatory Programme, with acronym – PROAGROBURO [12]. The ground-truth data consists of descriptions of the LU/LC types, phenological stages and vegetation cover of crops, GPS measurements, and photos. The collected ground data will contribute of selecting appropriate training samples for the supervised classification on the chosen satellite images. A multispectral EO-1 ALI image acquired on 09.07.2011 was used for the present investigation (Fig.1 and Fig. 2).

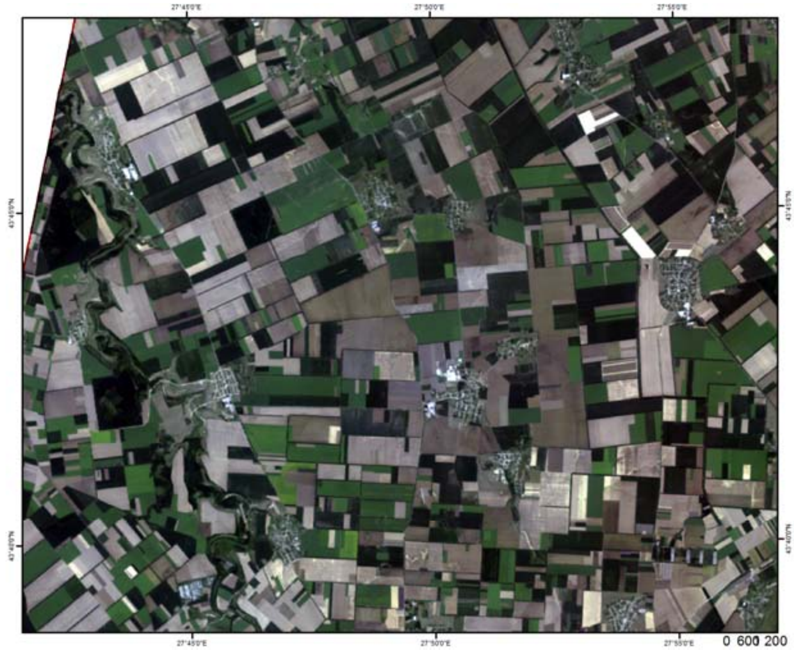

Fig. 1. Raw EO-1 ALI image True color composite

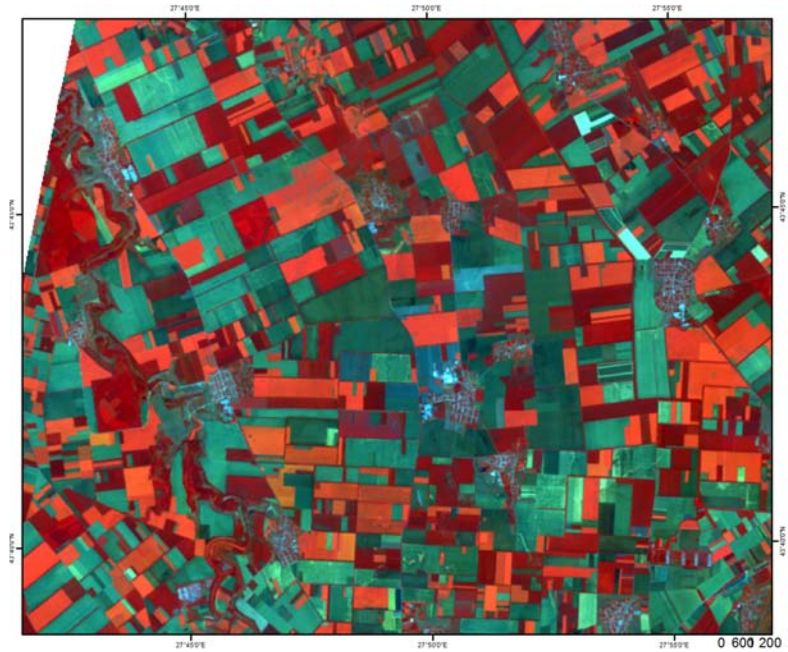

Fig. 2. Raw EO-1 ALI False color infrared composite

The spatial resolution of the image is 10m in the panchromatic band and 30m for the 9 multispectral bands: blue bands (0.43 – 0.45 μm; 0.45 – 0.51 μm); green band (0.52 – 0.60 μm), red band (0.63 – 0.69 μm), NIR bands (0.77 – 0.80 μm; 0.845 – 0.895 μm) and SWIR bands (1.20 – 1.30 μm; 1.550 – 1.750 μm; 2.080 – 2.350 μm). The temporal resolution of the satellite is 16 days, which makes it appropriate for monitoring agricultural applications.

An arable land mask using CORINE data was applied on the EO-1 ALI image in order to classify only the arable land and reduce the occurrence of mixed pixels with other non arable classes.

The *k*-mean and Iterative Self-Organizing Data Analysis (ISODATA) clustering algorithms are the most frequently used ones in RS. The ISODATA algorithm was selected in this study because it allows different number of clusters, while the *k*-mean algorithm assumes that the number of clusters is known a priori [13,14,15]. Unsupervised ISODATA cluster classification with four classes was applied to spectraly discriminate the crops and to collect the necessary information in order to delineate the training samples for the supervised per-pixel classification.

Per-pixel supervised classification using the Maximum Likelihood Classifier (MLC) algorithm was applied on the arable territories of the test site for crop identification purposes. In the MLC procedure, at least 10–15 independent training cases per class were used, so that its mean and variance can be estimated. Around 160-170 randomly distributed points were used for accuracy assessment for the classified image. The process continued with crop area estimates based on pixel counting technique. This step can be accomplished only if the pixel-based crop identification overall accuracy is higher than 95%.

3. Results and discussions

3.1. Applying arable mask from CORINE 2006 land-cover database

An arable land mask using CORINE 2006 land-cover data was applied on the EO-1 ALI image in order to classify only the arable land and reduce the occurrence of mixed pixels. In Fig.3 all the land cover classes present in the test area are shown and 2.1.1. Non-irrigated arable land class was used to build the mask layer.

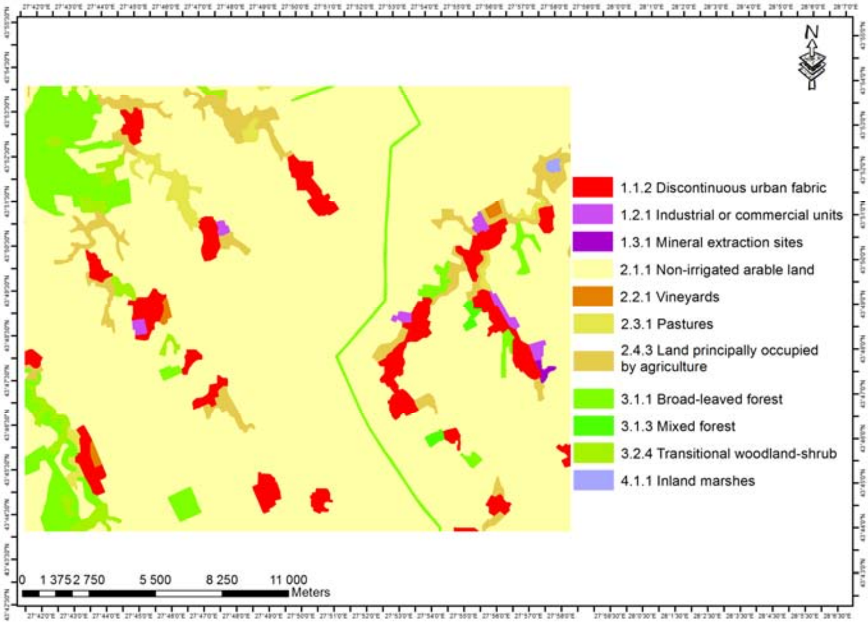

Fig. 3. CORINE land cover classes

3.2. Conducting per-pixel supervised classification using the maximum likelihood classifier algorithm for crop identification

The crop identification process was accomplished firstly by conducting unsupervised classification (using ISODATA algorithm) with 4-5 classes for the multispectral EO-1 ALI image (Fig.4.).

This spectral information was used together with the ground data as an indicator where to draw training samples for the supervised classification. The unsupervised classification is traditionally the first step and is accommodating the interpretation of the images. Supervised classification using the Maximum Likelihood Classifier (MLC) algorithm was applied to the arable land images. In the MLC procedure, a key concern is to collect a training set comprising of at least 10–30 independent training cases per class per discriminatory variable (e.g. band) to allow the formation of a representative description of the class, so that its mean and variance can be reasonably estimated [16].

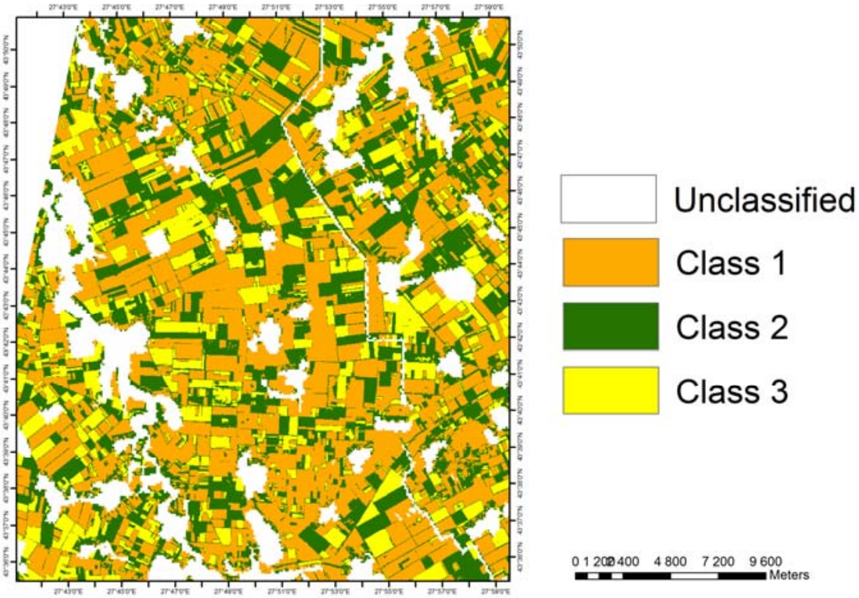

Fig. 4. Unsupervised classification on EO-1 ALI satellite image acquired on 09.07.2011

For example, the spectral response of an agricultural crop class in an image might vary as a function of variables such as: the crops growth stage, topographic position, density of vegetation cover, health, impact of management activities, substrate conditions and instrument view angle [17].

The gathered training set from the field data was good enough to make representative training samples for the arable land classes. The unsupervised classification in combination with the ground information helped to choose and delineate appropriate training samples for the supervised classification of the EO-1 ALI image. The investigated phenological stages based on the image acquisition date are: dough development for winter wheat and flowering phase for early sown hybrids and vegetative phase for late sown hybrids of maize cultivars. The identified classes and their distribution in percentage for the EO-1 ALI satellite image are: winter crops - 35.37 %; sunflower – 17.19%; maize – 20.39%; stubble fields – 0.17%; and class unclassified which includes the no data part and the applied mask of the image – 26.33% (Fig.5),

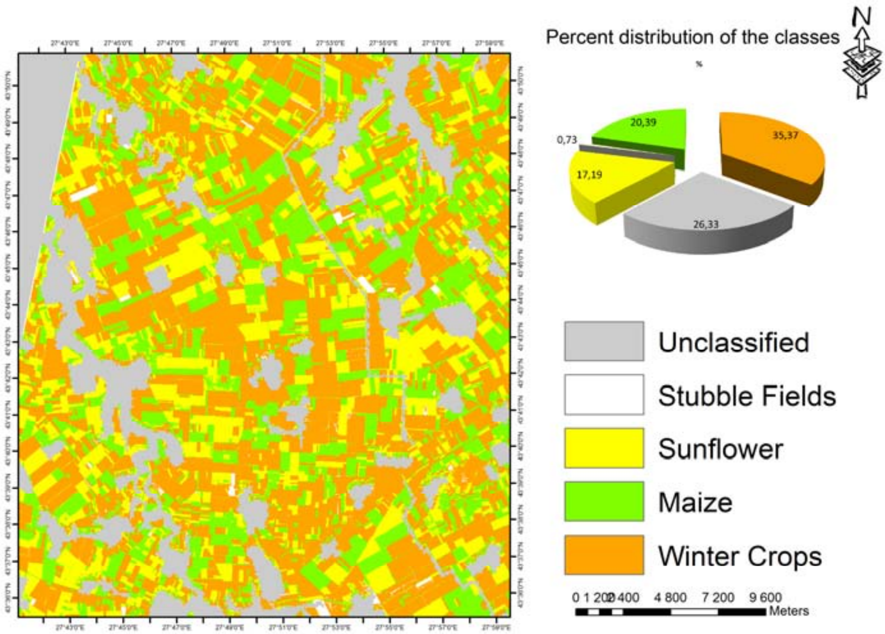

Fig. 5. Per-pixel supervised classification of EO-1 ALI satellite image acquired on 09.07.2011

3.3. Applying accuracy assessment tool in ERDAS Imagine and deriving accuracy totals and kappa statistics

Accuracy assessment tool in ERDAS Imagine software was utilized for assessing the accuracy of the per-pixel classified images of EO-1 ALI. Around 180-190 randomly distributed points were assessed for the classified image. Accuracy assessment was applied on the EO-1 ALI classified image for crop identification using its spectral resolution by applying visual interpretation on the panchromatic and both on the unsupervised and supervised classifications in combination with the ground data.

The achieved results on the overall classification accuracy for the EO-1 ALI image is 96.24% and overall kappa statistics is 0.9397 (Table 1).

The EO-1 ALI accuracy assessment shows that the class stubble fields can not be assessed because of the small part of the image occupied. The sunflower class has Producer’s accuracy of 89.47%, which is the only accuracy that is below 90%, since all the other cllasses have accuracies well above 90%. Considering the achieved result based on the per-pixel classification which is above 95%, this gives a possibility to apply pixel-counting method to calculate the area estimates, as it would not introduce much bias in the calculation.

Table 1. Accuracy totals for EO-1 ALI satellite image

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy (%)	User Accuracy (%)
Stubble fields	2	0	0	-	-
Sunflower	38	34	34	89.47	100.00
Maize	55	57	54	98.18	94.74
Winter wheat	92	96	92	100.00	95.83
Totals	**187**	**187**	**180**		

Overall accuracy 96.24% and overall kappa statistics is 0.9397

3.4. Calculating crop area estimated based on pixel-counting technique

The crop area estimates were calculated for the EO-1 ALI satellite image using the per-pixel supervised classifications. The crop area estimates are calculated using the followed formula: number of pixels for each class of the classified image multiplied by the area represented by each pixel in the chosen image based on [18]. This method was selected because the overall classification accuracy was high enough (above 95% overall accuracy) to apply that method and in the same time not to introduce bid bias. The whole territory occupies 234.5 km^2. The crop area estimated for the EO-1 ALI image classes show the following estimates: winter crops - 83 km^2; sunflower – 40.3 km^2, maize – 47.8 km^2 and stubble fields class occupies 1.7 km^2. The rest of the image is occupied with the class unclassified, which includes both the no data part of the image and the applied mask with 61.8 km^2.

4. Conclusions

The presented methodology provides opportunity to calculate crop area estrimates based on per-pixel supervised classification with higher than 95% overall accuracy. The results are encouraging and show that EO-1 ALI satellite image with both good spatial and spectral resolution and acquired in

July are beneficial for conducting crop area estimates on a highly cultivated crop area. The high overall accuracy can be accomplished after carefully chosing representative training samples. The beginning of July can be a perfect time to separate spring crops (sunflower and maize cultivars) and assess crop area estimated for all cultivated crops based on per-pixel classification with overall accuracy above 95%. The down side of this study is that at the time of image acquisition the winter crops are difficult to separate, since they are already with very low reflectance values.

References

1. M e y e r-R o u x, J., P. V o s s e n. 1994. The first phase of the MARS project, 1988–1993: overview, methods and results. In: Official Publications of the E.U., Luxembourg (Ed.), Report EUR 15599 EN. Conference on the MARS project: Overview and prospectives, Belgirade, 1994, 33-79.
2. D e W i n n e P. (2004). Les Besoins de la Direction Generale VI: Agriculture. In: Office for official publications of the E.U., Luxembourg (Ed.), Report EUR 15599 EN, 2004. Conference on the MARS project: Overview and prospectives, Villa Carlotta, Belgirate, Lake Maggiore, Italy, 17-22 (in French).
3.C o h e n, Y., M. S h o s h a n y. 2002. A national knowledge-based crop recognition in Mediterranean environment International Journal of Applied Earth Observation and Geoinformation, 4, 75- 87.
4. MacDonald R. 1984.. IEEE Transaction on Geoscience & Remote Sensing, GE-22, 473-481.
5. B l a e s, 2005. Efficiency of crop identification based on optical and SAR image time series. Remote Sensing of Environment, 96 352 – 365.
6. T o w n s h e n d, J., C. J u s t i c e, W.,L i, C. G u r n e y, J. M c M a n u s. 1991. Global land cover classification by remote sensing: present capabilities and future possibilities. Remote Sens Environ, 35, 243-255.
7. D e F r i e s, R., M. H a n s e n, J. T o w n s h e n d. (1995). Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. Remote Sens Environ, 64 (3), 209-222.
8. R e e d, B. C., J. F. B r o w n, D. V a n d e r Z e e, T. R. L o v e l a n d, J. W. M e r c h a n t, D. O. O h l e n. 1994. Measuring phenological variability from satellite imagery. Journal of Vegetation Science, 5, 703-714.
9. B r o w n, J. F., T. R. L o v e l a n d, J. W. M e r c h a n t, B. C. R e e d, D. O. O h l e n. 1993. Using multisource data in global land-cover characterization: Concepts, requirements, and methods. Photogramm Eng Remote Sens, 59(6), 977-987.
10. L o v e l a n d, T. R., J.W. M e r c h a n t, J. F. B r o w n, D. O. O h l e n, B. C. R e e d, P. O l s o n, Hutchinson J. 1995. Seasonal land-cover regions of the United States. Annals of the Association of American Geographers, 85, 339-355.

11. R y e r s o n, R. A., P. J. C u r r a n, P.R. S t e p h e n s. 1997. Agriculture. In: Philipson W.R. (Ed.) Manual of Photographic Interpretation. 2nd ed. Americal Society for Photogrammetry and Remote Sensing, Bethesda, MD, 365-397.
12. R o u m e n i n a, E., V. K a z a n d j i e v, P. D i m i t r o v, L. F i l c h e v, V. V a s s i l e v, G. J e l e v, V. G e o r g i e v a, H. L u k a r s k i. 2013. Validation of LAI and assessment of winter wheat status using spectral data and vegetation indices from SPOT VEGETATION and simulated PROBA-V images. Int J Remote Sens, 34(8), 2888-2904.
13. G r o o m, B.G., R. M. F u l l e r, A. R. J o n e s. 1996. Contextual correction: techniques for improving land-cover from remotely sensed images. Int J Remote Sens, 17, 69-89.
14. G a r c ı a-C o n s u e g r a, J., G. C i s n e r o s. 1999. Establishing spatially continuous in a non-supervised way. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'99), 1999, Hamburg, Germany.
15. Y a n g, H., S. K h o r r a m, X. L. D a i. 1999. Applications of simulated annealing minimization technique to unsupervised classification of remotely sensed data. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'99), 1999, Hamburg, Germany.
16. P i p e r, J. (1992). Variability and bias in experimentally measured classifier error rates. Pattern Recognition Letters, 13, 685–692
17. F o o d y, G. M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201.
18. G a l l e g o (2004): Remote sensing and land cover area estimation, International Journal of Remote Sensing, 25:15, 3019-3047.

ОЦЕНКА НА ЗЕМЕДЕЛСКИТЕ ПЛОЩИ ЧРЕЗ ПИКСЕЛНО-ОРИЕНТИРАНА КОНТРОЛИРАНА КЛАСИФИКАЦИЯ ВЪРХУ ИЗОБРАЖЕНИЕ НА ЕО-1 ALI ЗА ТЕСТОВИ УЧАСТЪК ЖИТЕН, РАЗПОЛОЖЕН В СЕВЕРОИЗТОЧНА БЪЛГАРИЯ

В. Василев

Резюме

Целта на настоящият доклад е да се определят засетите площи със земеделски култури на основата на пикселно - ориентирана класификация върху тестови участък Житен, разположен в Североизточна България. Избрано е спътниково изображение заснето от спътника Earth Observation-1 със сензор ALI на 09.07.2011. Методологията в настощето изследване включва следните етапи на работа: 1) прилагане на маска, включваща само обработваемите земи на територията за

изследване от базата данни на CORINE 2006 земно покритие; 2) провеждане на пикселно-ориентирана класификация по алгоритъма на максималното подобие с цел разпознаване на земеделските култури върху изображението; 3) прилагане на инструмента „оценка на точността“ в програмния продукт ERDAS Imagine и извличане на показателите обща точност и капа статистика от класифицираното изображение; 4) изчисляване на заетите площи по принципа на броя пиксели. Достигната обща точност на класификацията е 96.24 % и капа статистика от 0.9397. Високата обща точност е постигната чрез прецизен подбор на обучаващите множества включени в алгоритъма. Началото на месец юли може да се определи като идеален избор за разпознаване видовете пролетни култури (слънчоглед и царевица) и освен това за определяне на заетите площи за всички разпознати култури на основата на пикселно-ориентирана класификация с обща точност от над 95 %. Ограничението при проведеното изследване е, че по-времето на провеждането му (началото на месец юли) е трудно да се отделят по спектрални отражателни характеристики посевите на зимните култури.