The study aims to classify heterogeneous land use/cover and agricultural pattern belong to Catalca district, Istanbul using two supervised classification methods such as MLC and SVM Classification methods to compare performances of both classification methods on a new generation SPOT 7 image. The support vector machine classification approach slightly outperformed the maximum likelihood classification in both overall accuracy and Kappa statistics.

2. STUDY AREA & DATA

For this study Catalca district has been selected as the study area. Catalca is located in Kocaeli-Catalca part of Marmara region, Istanbul (Figure 1). It is an urban center where established on western boundary of Istanbul’s European side and the largest district of Istanbul with total area of 1115.50 square kilometers. The population of the region is 67.329 according to the population census of 2015 (Turkey Statistical Institute). In the region a majority of the surface area is covered with forests. The other land cover/use categories include artificial surfaces, water surfaces, wetlands, agricultural fields, pasture lands, forest lands and mining areas. The most important labor sector in Catalca district is agriculture. According to Istanbul provincial agriculture master plan, it has the most agricultural land distribution, agricultural production and productivity. It has
3. METHODOLOGY

3.1. Image Pre – Processing

The methodology of the study includes mainly three steps which are pre-processing of satellite image (radiometric, atmospheric and geometric correction), classification of satellite image with MLC and SVM, evaluating and comparing classification accuracies of both classification techniques.

In the image pre-processing stage radiometric and atmospheric correction were applied. Digital number values for the SPOT 7 satellite image were converted to radiance values and FLAASH Atmospheric Correction Model was used for atmospheric correction. Scene center location (lat/lon), sensor type, sensor altitude (km), ground elevation (km), pixel size (m), flight date, flight time GMT (HH:MM:SS), atmospheric model, aerosol model, initial visibility (km) were used as input parameters in the model. SPOT 7 image was geometrically corrected by image-to-map registration using 1:5000 scaled topographic maps and orthophotos. It took into consideration to be homogeneously distributed on the image of the selected control points. Connection between the image coordinate system and the ground coordinate system was accomplished by polynomial transformation. As a result of the conversion, it was obtained under the root mean square error of 0.5 pixels. Nearest neighbor method was used as resampling method.

3.2. Classification Methods

3.2.1 Maximum Likelihood Classification

Supervised classification is a technique that is based on the statistics of training areas representing different ground objects selected subjectively by users on the basis of their own knowledge or experience. In this study, maximum likelihood classification (MLC), which is the most common classification method in remote sensing, was used to derive land-use/cover categories of selected study area. In this method, the pixel is assigned to the class for which the probability of the pixel belonging is highest. MLC is based on Bayes’ Theorem (Jensen, 2005).

3.2.2 Support Vector Machine Classification

The support vector machine (SVM) method, which was originated in the late 1970s by Vapnik (1995), is one of the machine-learning algorithms. It is based on statistical learning theory and has recently been extensively used in
remote sensing for pattern recognition and classification. The method was originally designed for binary classification, and it allows the use of optimal algorithms to locate the best boundaries separating the binary classes in the feature’s space. The boundary is called the optimum separating hyperplane, which is aimed at maximizing the margin width.

The main advantage of SVM is the ability to make good generalizations of high-dimensional data with few training samples. SVM works with pixels in the boundaries of classes, which are called support vectors, and thus it is possible to get accurate classifications with small training sets.

Related parameters vary according to used kernel function such as Linear kernel, Polynomial kernel, Radial Basis Function kernel, Sigmoid kernel (Yang, 2011; Shi and Yang, 2012). In this study radial basis function kernel was used and optimal parameter values for the selected kernel function determined based on literature review. The parameters for the RBF kernel were set to 0.25 and 100 for γ (gamma= 1/number of the band) and C (error penalty), respectively, for the SVM classification. The pyramid parameter was set to a value of 0 to process the satellite data at full resolution (6 m).

4. RESULTS & CONCLUSIONS

In this study, two different supervised classification methods were employed using SPOT 7 data to analyze the potential of mapping land use/land cover in the Catalca, Istanbul. In the classification, at the first stage, training and test data was prepared to classify satellite image that is covering the study area. The same training data was used for all classifications problems in this study.

As a result of the classification methods, twelve land-use categories were distinguished in the selected region: urban-built up, industrial, road, quarry/mining area, sunflower, greenhouses, other agricultural fields, forest, grassland, bare land, water surfaces and cloud+shadow (Figure 3).

The output of the classified images without any error or bias is the accurate thematic map. There are a number of equations that can show the level of error statistically, such as producer accuracy, user accuracy, overall accuracy, and Kappa, which can be calculated using the error matrix (Foody, 2002). An accuracy assessment was performed using collected ground-truth data for all two types of classification using a standard error matrix. Kappa statistics and overall accuracy were used to compared to determine the performance of the selected methods for the selected heterogeneous region (Table 2).

<table>
<thead>
<tr>
<th>Classification Method</th>
<th>Overall Accuracy</th>
<th>Kappa Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Likelihood Classification Support Vector Machine Algorithm</td>
<td>83.0486%</td>
<td>0.7703</td>
</tr>
<tr>
<td></td>
<td>85.6109%</td>
<td>0.7981</td>
</tr>
</tbody>
</table>

Table 2. Classifier accuracies based on test data for MLC and SVM

The results showed that Spot 7 MS images could be used to produce land use/cover maps and statistics. Based on accuracy assessment results thematic maps were produced with overall accuracy and kappa statistic values of 83.05 % and 0.77 for MLC, 85.61% and 0.80 for SVM supervised classification methods.

The evaluation results demonstrated that the SVM algorithm with an overall accuracy and a kappa coefficient has a higher accuracy in comparison with the MLC algorithm in land use mapping.

This algorithm has been suggested as an optimal classifier for the extraction of land use/cover map because of its higher accuracy and better consistency with the study area.

ACKNOWLEDGEMENT

The authors would like to thank ITU Research and Application Center for Satellite Communications and Remote Sensing (CSCRS) for providing remotely sensed data.
7. REFERENCES


http://www.intelligence-airbusds.com Received date: 15 Sept. 2016.


Turkey Statistics Institute, 2015.
