PRODUCING INFORMATION FOR CORINE DATABASE BY USING CLASSIFICATION METHOD: A CASE STUDY OF SAZLIDERE BASIN, ISTANBUL

F. B. Sarıyılmaz[a*], N. Musaoğlu[a], N. Uluğtekin[a]

[a] ITU, Civil Engineering Faculty, Geomatics Engineering Department, 34469 Maslak Istanbul, Turkey - (geze, musaoglune, ulugtek)
@itu.edu.tr

KEY WORDS: Classification, CORINE Database, Sazlidere Basin, Istanbul, LANDSAT, Generalization

ABSTRACT:

The Sazlidere Basin is located on the European side of Istanbul within the borders of Arnavutkoy and Basaksehir districts. The total area of the basin, which is largely located within the province of Arnavutkoy, is approximately 177 km^2. The Sazlidere Basin is faced with intense urbanization pressures and land use / cover change due to the Northern Marmara Motorway, 3rd airport and Channel Istanbul Projects, which are planned to be realized in the Arnavutkoy region. Due to the mentioned projects, intense land use /cover changes occur in the basin. In this study, 2000 and 2012 dated LANDSAT images were supervised classified based on CORINE Land Cover first level to determine the land use / cover classes. As a result, four information classes were identified. These classes are water bodies, forest and semi-natural areas, agricultural areas and artificial surfaces. Accuracy analysis of the images were performed following the classification process. The supervised classified images that have the smallest mapping units 0.09 ha and 0.64 ha were generalized to be compatible with the CORINE Land Cover data. The image pixels have been rearranged by using the thematic pixel aggregation method as the smallest mapping unit is 25 ha. These results were compared with CORINE Land Cover 2000 and CORINE Land Cover 2012, which were obtained by digitizing land cover and land use classes on satellite images. It has been determined that the compared results are compatible with each other in terms of quality and quantity.

1. INTRODUCTION

It is very important to produce thematically accurate databases about land use / cover classes in terms of environmental assessments for decision makers. The "Land Cover" means to the bio - physical cover of land such as forests, crops or artificial surfaces, and the "Land Use" term means to the socio - economic use of land as agricultural land or residential area (URL - 1, 2016). It is very difficult and time consuming to extract land use / cover information in large areas with classical methods. With the possibilities provided by satellite images, it is an indispensable source to produce such information and to make time analyzes.

Land use / cover classes that are produced from satellite images has certain standards. There are four main approaches in the literature: ANDERSON, FAO, NLCD and CORINE (Coordination of Information on the Environment) (Anderson et al., 1976; Loveland et al., 1991; URL – 2, 2014; Di Gregorio et al., 2016). The CORINE program has been implemented, an information system on the state of the European environment has been created, nomenclatures and methodologies have been developed from 1985 to 1990 by the European Commission (URL – 3, 2017). The CORINE Land Cover database continued to be updated at regular intervals as the needs of environmental decision makers continued. The minimum mapping unit is 25 ha and the minimum width unit is 100 m. Areas smaller than 25 ha are allowed in the national land cover database as additional thematic layers, but should be aggregated or generalized in the European database (URL – 3, 2017). The mapping method is computer supported by photo interpretation for greater part of countries, however some countries perform semi-automatic methodologies by integrating in being land use data, digital classification of satellite images or cartographic generalization of high-resolution, national land use / cover data (URL – 4, 2017).

There are many up to date studies and researches in different study areas and subjects carried out using the CORINE Land Cover database in the literature (Diaz-Pacheco and Gutierrez, 2014; Perez-Hoyos, et al., 2014; Suau-Sanchez et al., 2014; Alavipanah et al., 2015; Alp et al., 2015; Baltzer, 2015; Keil et al., 2015; De Meij et al., 2015; Bartold, 2016; Calvache et al., 2016; Mantas et al., 2016; Giorgio et al., 2017; Rosina et al., 2017; Szumacher and Pabjanek, 2017, Sertel et al., 2017). Remote sensing and geographical information systems methods were used and land use / cover information from the CORINE database were utilized in these studies and researches. The researchers have contributed to the production, usage and development of the CORINE Land Cover database with the results and recommendations obtained. In this study, the appropriation of generalized pixel based classification results from LANDSAT data to produce data for CORINE Land Cover Database was investigated.

2. STUDY AREA

Istanbul is ranked 22nd in terms of the density of urban areas in the world and it is the most crowded city of Turkey (URL – 5, 2017). Istanbul, which connects Europe and Asia Continents, continues to receive migrations from other countries of our country due to the multiplicity and diversity of its economic and social facilities, and therefore significant changes are observed in urban areas within provincial boundaries. There are 7 drinking and usage water reserves within the province borders. These are Omerli, Elmali and Darlik dam lakes on the Anatolian side and Alibeykoy, Terkos, Sazlidere and Buyukcekmece dam lakes on the European side (Figure 1.)

* Corresponding Author

The Sazlidere Basin, which is the subject of the study, has an area of approximately 177 km^2 and is also located on the European side of Istanbul, within the borders of the Arnavutkoy and Basaksehir districts, and the great extent of the basin is within the borders of the Arnavutkoy district. In addition, the Arnavutkoy district center has developed into the basin boundary from the east of the basin.
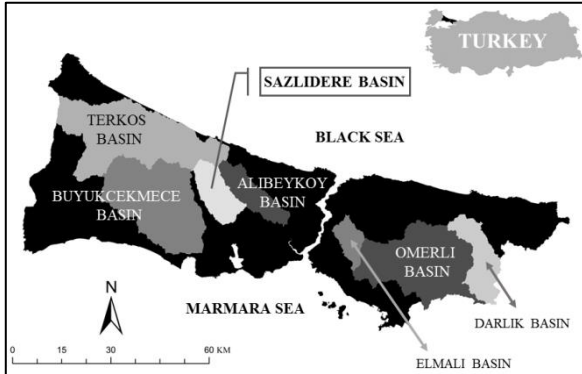

Figure 1. Basins that provide drinking water to Istanbul

3. MATERIAL USED

3.1 LANDSAT Data

In order to determine the change of land cover and usage in Sazlıdere Basin; LANDSAT MSS and ETM images of July 2000 and July 2012 were used. The satellite images used were provided free of charge from the LANDSAT USGS website and in the UTM WGS 84 coordinate system and datum. Table 1 shows LANDSAT MSS and ETM sensor properties (URL – 6, 2016).

	MSS Bands (µm)		TM / ETM Bands (µm)	
Band 1			30 m B	0.441-0.514
Band 2			30 m G	0.519-0.601
Band 3			30 m R	0.631-0.692
Band 4	80 m G	0.5-0.6	30 m NIR	0.772-0.898
Band 5	80 m R	0.6-0.7	30 m SWIR-1	1.547-1.749
Band 6	80 m NIR-1	0.7-0.8	120 / 60 m TIR	10.31-12.36
Band 7	80 m NIR-2	0.8-1.10	30 m SWIR-2	2.064-2.345
Band 8			15 m PAN (ETM)	0.515-0.896

Table 1. LANDSAT MSS and ETM sensor features

LANDSAT data of the study area were examined and their suitability was investigated. When the 2000, 2011 and 2012 dated LANDSAT images of the study area were examined, it was seen that the images of 2000 and 2012 were appropriate to study but the image of 2011 had cloud contains. Since there was no other image of 2011 dated taken in a different month, the 2012 dated LANDSAT image was used for comparison with CORINE Land Cover 2012, even though its spatial resolution is low.

3.2 CORINE Database and Classification System

In this study, the information classes were determined based on the CORINE classification system first level. CORINE Land Cover 2000 and CORINE Land Cover 2012 data were compared with the results of pixel based supervised classification results and generalized classification results of the 2000 and 2012 dated LANDSAT images.

At the present time, Turkey Government has fully completed the National Land Cover 1990, 2000, 2006 and 2012 and handed it over to the European Environment Agency (URL – 7, 2014). In CORINE 2000 Project 2000 dated LANDSAT image with 30 m spatial resolution and in CORINE 2012 project 2011 and 2012 dated RapidEye images with 5 m spatial resolution and RSD LISS III images with 25 m spatial resolution were used (URL – 7, 2014; URL – 8, 2016, URL – 9, 2017). Within the CORINE Land Cover database downloaded from the website, the boundaries of the study area were separated using vector data. Figure 2 shows CORINE Land Cover data in Sazlidere Basin area.
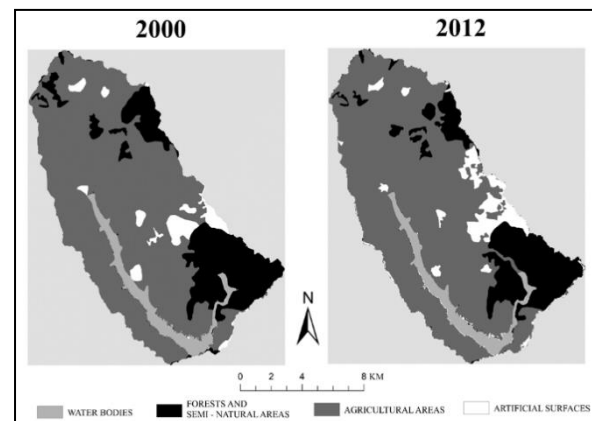

Figure 2. CORINE Land Cover 2000 and 2012 data

The CORINE classification system has a three-level hierarchical structure. At the first level; land cover and use types are divided into 5 main categories. At the second level; basic land cover and use units are divided into 15 sub-categories and 44 sub-categories in the third level. In this study, CORINE classification system first level classes were produced and the study area, Sazlidere Basin, covers 4 of them.

4. METHOD

4.1 Classification

There are different classification methods such as pixel based, object based or support vector machine to obtain land use / cover information from satellite data. In this study, pixel based classification was used to obtain CORINE first level classes. For this aim, unsupervised classification by using ISODATA algorithm and supervised classification by using maximum likelihood algorithm were processed together as a hybrid classification (Lillesand et al., 2004; Sunar and Özkan, 2011, Sarıyılmaz, 2012). The spectral classes resulting from the unsupervised classification process were modified by comparing with high spatial resolution image and used as input to the supervised classification process.

As a result of the classification, four information classes of the first level of the CORINE classification system were obtained for each image. These are; water bodies, forests and semi-natural areas, artificial surfaces and agricultural areas. Artificial surfaces (AS) include urban texture areas, industrial

or commercial units, transportation units, mine, foundry and construction sites, and non-agricultural plant sites. Agricultural areas (AA) include arable land, permanent crop fields, pasture lands and heterogeneous agricultural lands. Forests and semi-natural areas (FSNA) include forest areas, forests and grassy vegetation, and open areas with little vegetation or no vegetation cover. Water bodies (WB) are also defined as inland waters and seas (URL- 10, 2017).

In order to complete the classification process, the thematic accuracy of the results obtained should be determined (Richards, 1999). In this study, an accuracy analysis was carried out to determine the accuracy of the information classes obtained after completing the classification process. The most common method used to evaluate the accuracy of the thematic maps is to analyze the error matrices (Congalton and Green, 2008; Kumar, 2012). In order to express the thematic accuracy of the classification results, the overall accuracies of the images were calculated using the error matrix.

As a result of the accuracy assessment process, the overall accuracy of the 2000 dated classified image was calculated as 86% and the overall accuracy of the 2012 dated classified image was 87%. These values are in the acceptable range and it is appropriate to continue studying.

4.2 Generalization

It is essential to provide easy and fast understandable information on quantity and quality, depending on the purpose and scale of the results obtained after the acquisition and processing of spatial data. Making this information meaningful is done through maps produced in compliance with the standards accepted by the institutions or decision makers. This requires the generalization of database or information for cartographic model, depending on the subject and the scale. Cartographic generalizations, which are tightly connected to the purpose and scale factors of the produced map, are simplification and elimination (selection), geometric combination, exaggeration, displacement, emphasis and conceptual unification" (Ulugtekin, et al., 2013). In general terms, the generalization process is to decrease information redundancy in the result obtained and to include the most appropriate information in the content of the result (Bilgi, 2017).

In this study, images which were obtained as a result of conceptual generalization, for example, including all the settlement areas, industrial zones and roads into the thematic class of artificial surfaces, were still conceptually generalized by the thematic pixel aggregation process (Figure 3). At the same time, with this process, the pixel size of the image has been rearranged to be 25 ha. In this way, the images were scaled to the same level as CORINE Land Cover data and the salt and pepper effect in the images was removed. In order to ensure compatibility between classification results and CORINE Land Cover data, the conceptual generalization process was applied to the classified images considering the smallest mapping unit of CORINE Land Cover data, 25 ha. A pixel represents 0.09 ha for the 2000 dated image with a spatial resolution of 30 m. In this case 25 ha can be expressed with 277 pixels. For this image, the smallest mapping unit has been approached to 23 ha with 16x16 pixels. A pixel area represents 0.64 ha for the 2012 dated image with a spatial resolution of 80 m is likewise. In this case 25 ha can be expressed by 39 pixels. The smallest mapping unit for this image has been approached to 23 ha with 6x6 pixels.
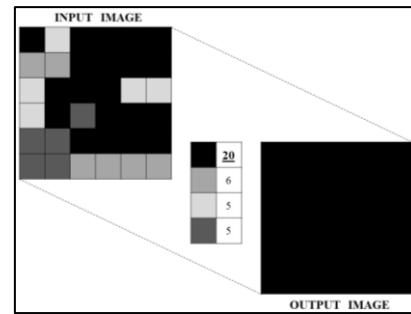

Figure 3. "Thematic Pixel Aggregation" process

5 RESULTS and CONCLUSIONS

Since the results of the accuracy analysis applied to the classified images are within the acceptable range, the spatial results were compared with the CORINE Land Cover database. 2000 and 2012 dated classified images and statistical results of the classification process are given in Figure 4 and Table 2.
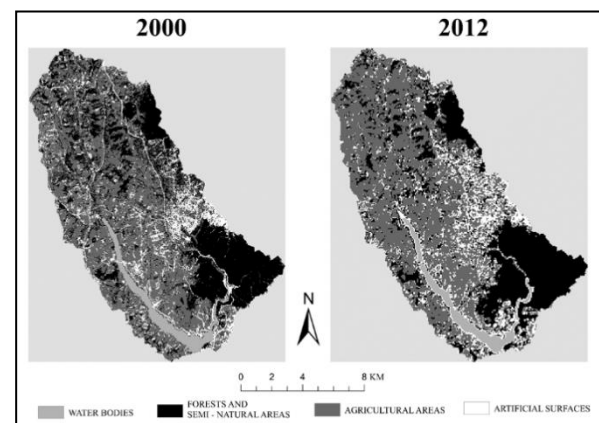

Figure 4. Classification results

Class	2000	2012
WB	4%	3%
FSNA	34%	29%
AA	48%	52%
AS	14%	16%
Total	100%	100%

Table 2. Distribution of land use / cover classes

At the next stage of the study, the CORINE Land Cover 2000 database (URL - 11, 2014) was compared with the 2000 dated classified image and the CORINE Land Cover 2012 database (URL - 12, 2017) was compared with the 2012 dated classified image for Sazlidere Basin area. Table 3 shows the areal comparison statistics.

Class	Classification (ha)		CORINE (ha)		Overlap Rate (%)	
	2000	2012	2000	2012	2000	2012
WB	670	588	918	918	73	64
FSNA	5745	5002	3379	3055	59	61
AS	2331	2680	767	1056	33	39
AA	8345	8821	12027	12062	69	73
Total	17091	17091	17091	17091		

Table 3. The comparison results for classification process

The mismatches between the classification results and the CLC data are due to differences in the data production method as well as the non-generalization of classification results. In order to make a quantitatively more significant comparison, classification results were generalized.

2000 and 2012 dated generalized images and statistical results of the classification process are given in Figure 5 and Table 4.
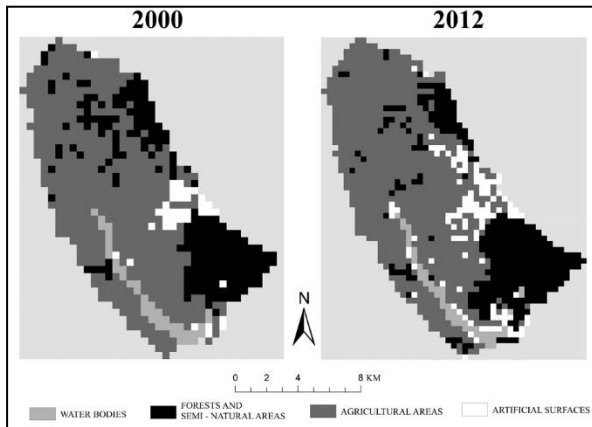

Figure 5. Generalized results

Class	2000	2012
WB	5%	4%
FSNA	26%	25%
AA	64%	62%
AS	5%	9%

Table 4. Distribution of land use / cover classes in generalized images

Table 5 shows the areal comparison statistics between classification results and CORINE Land Cover database first level classes after generalization process.

Class	Classification (ha)		CORINE (ha)		Overlap Rate (%)	
	2000	2012	2000	2012	2000	2012
WB	832	665	918	918	91	72
FSNA	4454	4263	3379	3055	76	72
AS	790	1499	767	1056	97	70
AA	11015	10664	12027	12062	92	88
Total	17091	17091	17091	17091		

Table 5. The comparison results for generalization process

When the Table 5 is examined, it is seen that the overlap rates between 2000 dated image results and CORINE Land Cover 2000 data is higher than the overlap rates between 2012 dated image results and CORINE Land Cover 2012 data. This is due to differences in spatial and spectral resolution of the data used in pixel based classification process researched in this study and CORINE Land Cover data production.

In this study, the consistency between the classes produced by the classification of LANDSAT images and CORINE Land Cover Database first level classes was investigated. For this aim, classification results were compared with same dated CORINE Land Cover data of study area, and low overlap rates were calculated as a result of comparison. This

is due to the fact that the classification results have not been generalized in addition to the method difference used in the production of the CORINE database. Therefore, in the following stage, classification results were conceptually generalized by the thematic pixel aggregation process. After the generalization, the results were compared again with CORINE Land Cover data and it was observed that the overlap rates of the data were increased.

The data used by European Commission in the CORINE Land Cover 2012 database production and the 2000 dated LANDSAT image used in this research have same spatial and spectral resolution characteristics. For this reason, these two different results have achieved higher rates of agreement with each other. Similar results can be obtained with different methods when using similar data. Consequently, it is appropriate that pixel based classification process results can be used to produce data for CORINE Land Cover Database after generalization process in areas containing mixed classes.

The CORINE Land Cover database is updated every 6 years since the 1990s. Using automated or semi – automated methods to update the database will allow producing results in a short time. For this reason, it is important to investigate the advantages of different classification methods on the CORINE Land Cover database in urban and rural areas with different land use / cover characteristics.

REFERENCES

Alavipanah, s., Wgmann, M., Qureshi, S., Weng, Q. and Koellner, T., 2015. The role of vegetation in mitigating urban land surface temperatures: a case study of Munich, Germany during the warm season. *Sustainability*, 7, pp. 4689-4706.

Alp, G., Yay Algan, I. and Sertel, E., 2015. Determination of agricultural land changes in Mugla, Turkey using remotely sensed data and CORINE methodology. 2015 Fourth International Conference on Agro-Geoinformatics.

Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R. E., 1976. A Land Use and Land Cover Classification System for Use with Remote Sensor Data, *Geological Survey Professional*, pp. 964.

Balzter, H., Cole, B., Thiel, C. and Schmullius, C., 2015. Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data using Random Forests. *Remote Sensing,* 7, pp. 14876-14898.

Bartold, M., 2016. Development of forest cover mask to monitor the health condition of forests in Poland using long-term satellite observations. *Forest Researced Papers*, 77 (2), pp. 141-150.

Bilgi, S., 2017. Evaluation of Spatial Knowledge of the Map by Quantitative Measurement Methods, *Afyon Kocatepe University Journal of Science and Engineering Sciences*, 17 (015502) pp. 154 – 164 (in Turkish).

Calvache, M. F., Prados, M. J. and Lourenço, J. M., 2016. Assessment of National Parks affected by naturbanization processes in Southern Europe. *Journal of Environmental Planning and Management,* 59 (9), pp.1629-1655.

Congalton, R. G. and Green, K., 2008. "Assessing the Accuracy of Remotely Sensed Data: Principles and Practices", Second Edition, CRC Pres.

De Meij, A., Bossioli, E., Penard, C., Vinuesa, J. F. and Price, I., 2014. The effect of SRTM and Corine Land Cover data on calculated gas and PM10 concentrations in WRF-Chem. *Atmospheric Environment*, 101, pp. 177-193.

Di Gregorio, A., Henry, M., Donegan, E., Finegold, Y., Latham, J.,Jonckheere, I. and Cumani, R., 2016. "Land cover Classification System Software version 3, Food and Agriculture Organization of the United Nations, Rome", http://www.fao.org/3/a-i5232e.pdf, (24.08.2016).

Diaz-Pacheco, J. and Gutiérrez, J., 2014. Exploring the limitations of CORINE Land Cover for monitoring urban land-use dynamics in metropolitan areas. *Journal of Land Use Science,* 9 (3), pp. 243-259.

Giorgio, G. A., Ragosta, M. and Telesca, V., 2017. Climate variability and industrial-suburban heat environment in a Mediterranean Area. *Sustainability,* 9 (775).

Keil, M., Esch T., Feigenspan, S., Marconcini, M., Metz, A., Ottinger and M., Zeidler, J., 2015. Creation of a high-resolution product CLC 2006_Backdating by a backward look from the Digital Land Cover Model DLM-De 2009 to 2006 - A Contribution to the German CORINE Land Cover 2012 Project within a bottom-up approach. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-7/W3, 36th International Symposium on Remote Sensing of Environment.

Kumar, M., 2012. Digital Image Processing, Photogrammetry and Remote Sensing Division Indian Institute of Remote Sensing, Dehra Dun, *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology* 81-102, http://www.wamis.org/agm/pubs/agm8/Paper-5.pdf, (05.12.2012).

Lillesand, T. M., Kiefer, R. W. and Chipman, J. W., 2004. "Remote Sensing and Image Interpretation". 5th Edition, Wiley, USA.

Loveland, T. R., Merchant, J. W., Ohlen, D. O. and Brown, J. F., 1991. Development of a Land-Cover Characteristics Database for the Conterminous U. S. *Photogrammetric Engineering and Remote Sensing,* 57 (11), pp. 1453-1463.

Mantas, V. M., Marques, J. C. and Pereira, A. J. S. C., 2016. A geospatial approach to monitoring impervious surfaces in watersheds using Landsat data (the Mondego Basin, Portugal as a case study). *Ecological Indicators,* 71, pp. 449-466.

Perez-Hoyos, A., Garcia-Haro, F. J. and Valcarcel, N., 2014. Incorporating sub-dominant classes in the accuracy assessment of large-area land cover products: application to GlobCover, MODISLC, GLC2000 and CORINE in Spain. IEEE Journal of Selected Topıcs in Applıed Earth Observatıons and Remote Sensıng, 7 (1), pp. 187-204.

Richards, J.A., 1999. "Remote Sensing Digital Image Analyst". Third Edition, Springer Verlag, Berlin NY.

Rosina, K., Hurbanek, P. and Cebecauer, M., 2017. OpenStreetMap to improve population grids in Europe. *Cartography and Geographic Information Science,* 44 (2), pp. 139-151.

Sarıyılmaz, F. B., 2012. Change Detection with Time Series: Istanbul, Sariyer, (Master thesis). Istanbul Technical University, Institute of Science and Technology, Istanbul, (in Turkish).

Sertel, E., Yay Algan, I., Alp, G., Musaoğlu, N. and Kaya, Ş., 2017. Production of 1:25.000 scale land cover/use maps by means of very hıgh resolutıon Spot 6/7 satellıte images**.** Turkish National Association of Remote Sensing and Photogrammetry, 9th Technical Symposium, 27-29 April, Afyon, Turkey.

Suau-Sanchez, P., Burghouwt, G. and Pallares-Barbera, M., 2014. An appraisal of the CORINE land cover database in airport catchment area analysis using a GIS approach. *Journal of Air Transport Management* 34, pp. 12-16. *Sustainability*, 9, 665.

Sunar, F. and Özkan, C. (2011). Unit 7: Classification. "Remote sensing". T. C. Anadolu University Publication No: 2320, Open Education Faculty Publication No: 1317, ISBN 978 – 975 – 06 – 0995 – 4, (in Turkish).

Szumacher, I. and Pabjanek, P., 2017. Temporal changes in ecosystem services in European cities in the continental biogeographical region in the period from 1990–2012.

Ulugtekin, N., Dogru, A. Ö. And Bildirici, İ. Ö., 2013. Design of GIS Maps, TMMOB Geographic Information Systems Congress, Ankara, Turkey: November 11 – 13, (in Turkish).

URL-1 European Statistics: http://ec.europa.eu/eurostat/documents/205002/208012/LUCAS2012_C1-InstructionsRevised_20130110b.pdf/10f750e5-5ea0-4084-a0e7-2ca36c8f400c (28.06.2016)

URL-2 http://landcover.usgs.gov/usgslandcover.php, (18.12.2014).

URL-3 http://land.copernicus.eu/user-corner/technical-library/techrep89.pdf, (04.08.2017).

URL-4 http://land.copernicus.eu/pan-european/corine-land-cover/clc-1990, (04.08.2017).

URL-5 http://www.turksat.gov.tr/, (01.02.2017).

URL-6 landsat.gsfc.nasa.gov, (08.11.2016).

URL-7 http://www.infomab.com/file/040913120311.pdf, (05.12.2014).

URL-8 http://land.copernicus.eu/pan-european/corine-land-cover, (24.08.2016).

URL-9 http://corine.ormansu.gov.tr/corine, (05.08.2017).

URL-10 http://corine.ormansu.gov.tr/corineportal/avrupacalismalar.html, (16.06.2017).

URL-11 http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-clc2000-seamless-vector-database, (20.03.2014).

URL-12 http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012/view, (15.04.2017).