IRANIAN LAND COVER MAPPING USING LANDSAT-8 IMAGERY AND RANDOM FOREST ALGORITHM

M. Amani [1], A. Ghorbanian [2], S. Mahdavi [1], A. Mohammadzadeh [2]

[1] Wood Environment & Infrastructure Solutions, St. John's, NL, Canada, A1B 1H3 - (meisam.amani, sahel.mahdavi)@woodplc.com
[2] Dept. of Photogrammetry and Remote Sensing, Geodesy and Geomatics Engineering Faculty, K. N. Toosi University of Technology, Tehran, Iran – a.ghorbanian@email.kntu.ac.ir, a_mohammadzadeh@kntu.ac.ir

KEY WORDS: Google Earth Engine, Remote Sensing, Land Cover, Iran, Landsat-8

ABSTRACT:

Land cover classification is important for various environmental assessments. The opportunity of imaging the Earth's surface makes remote sensing techniques efficient approaches for land cover classification. The only country-wide land cover map of Iran was produced by the Iranian Space Agency (ISA) using low spatial resolution Moderate Resolution Imaging Spectroradiometer (MODIS) imagery and a basic classification method. Thus, it is necessary to produce a more accurate map using advanced remote sensing and machine learning techniques. In this study, multi-temporal Landsat-8 data (1,321 images) were inserted into a Random Forest (RF) algorithm to classify the land cover of the entire country into 13 categories. To this end, all steps, including pre-processing, classification, and accuracy assessment were implemented in the Google Earth Engine (GEE) platform. The overall classification accuracy and Kappa Coefficient obtained from the Iran-wide map were 74% and 0.71, respectively, indicating the high potential of the proposed method for large-scale land cover mapping.

1. INTRODUCTION

The Earth's surface is characterized by the distribution of distinct land cover types, including water, soil, vegetation, and man-made infrastructure (Gomez et al., 2013). These influence the interaction of the Earth's surface with the atmosphere and can therefore affect the climate. Moreover, the characteristics of these land cover types determine the specific habitats of various types of flora and fauna. Since various land cover types have significant impacts on ecosystems, they have been recognized as critical climate variables (GCOS, 2003). Based on the above explanation, obtaining reliable information from the Earth's surface is essential for a variety of applications, including monitoring, managing and sustainably developing natural resources and environments (Sterling et al., 2013; Chen et al., 2015).

The possibility of imaging the Earth's surface using a variety of spatial and spectral resolutions makes remote sensing methods efficient approaches for land cover classification (Chen et al., 2015; Wulder et al., 2018; Mahdavi et al., 2018). For instance, optical and radar datasets have been extensively applied to classify various land cover types (Ghorbanian and Mohammadzadeh 2018; Woźniak et al. 2018). Remote sensing methods can efficiently delineate various types of land cover within large-scale areas through different classification processes and can help assist policy makers with further planning and monitoring of these areas.

Like many regions, remote sensing data have been employed to generate land cover maps of different areas in Iran (e.g., Saadat et al., 2011; Eisavi et al., 2015; Yousefi et al., 2015; Minaei and Kainz, 2016; Zakeri et al., 2017; Rahdari et al., 2017). For instance, Saadat et al. (2011) utilized Landsat ETM+ images to map land cover over the Golestan Dam watershed. Additionally, Eisavi et al. (2015) employed spectral and thermal datasets for land cover mapping using a Random Forest (RF) algorithm. Finally, Synthetic Aperture Radar (SAR) data were employed by Zakeri et al. (2017) to classify the land cover of the city of

Tehran. However, most of these studies have investigated relatively small areas and, thus, there is a need to implement advanced remote sensing techniques to efficiently classify the land cover of the entire country.

When the objective is to map a large area, such as a country, using multi-year satellite images, it is important to consider efficient methods to store, pre-process, classify, and analyze large datasets Xiong et al., 2017; Amani et al., 2019). Particularly, Google Earth Engine (GEE) has opened the door for planetary scale analyses and offers ease-of-use and a consolidated library of global satellite data to continually monitor the Earth's surface at high spatial and temporal resolutions (Gorelick et al., 2017; Amani et al., 2019). GEE is a web-based integrated development environment, which optimizes the parallel processing of geospatial data and provides rapid prototyping and visualization of complex spatial analyses (Gorelick et al., 2017; Kumar and Mutanga, 2018). The high potential of GEE to quantify various land surface locations and changes has been demonstrated in several large-scale applications, including mapping crops, rice paddies, forests, floods, fires, and malaria risk areas, and as well as mapping and monitoring surface water and classifying wetlands.

Although several studies have investigated land cover mapping in Iran, all were conducted over relatively small areas mainly using basic approaches. However, in this study, the entire country was classified into 13 different classes using advanced machine learning algorithms within GEE. To this end, Landsat-8 time series images along with a RF algorithm were utilized. The accuracy of the map was also evaluated using a large amount of field samples. Finally, it is worth noting that this study is the first to investigate the potential of GEE to produce an Iran-wide land cover map.

2. MATERIALS AND METHODS

2.1 Study Area

The study area is Iran (Figure 1), covering an area of more than 1.5 million km^2 between latitudes and longitudes of 24° and 40° N and 44° and 64° E, respectively. Iran has diverse climates varying from arid to subtropical, although the majority is dry (Haftlang and Lang, 2003). The geographical characteristics of Iran cause climatic variation, as Iran is a mountainous country covering a massive latitudinal range (Alizadeh-Choobari and Najafi, 2018). There are different types of land covers in the country, varying from rocky mountains to sandy deserts. Iran also contains the Zagros and Alborz mountain chains located in the west and north, respectively, which have created the key basins of this country. Major urban and agricultural settlements have been established in these basins and several lakes and marshlands have been created by the rivers that flow through them. Dense forests cover the northern part of Iran between the Alborz mountain chain and the shoreline of the Caspian Sea to the south. The eastern part of the country is mostly covered by deserts (e.g., Kavir and Lut deserts). Finally, long shorelines created by the Persian Gulf and Oman Sea form the southern border of Iran (Haftlang and Lang, 2003; Alizadeh-Choobari and Najafi, 2018).
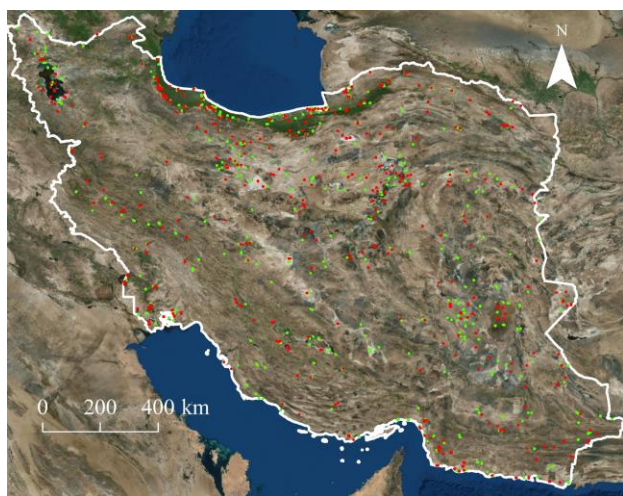

Figure 1. Study area and the distribution of training and test samples (red and light green colours, respectively).

2.2 Training and Test Datasets

In the classified map, 13 land cover classes were considered: Clay, Dense Forest, Farm Land, Kalut, Marshland, Outcrop, Range Land, Salty Land, Sand, Uncovered Plain, Urban, Water, and Wetland. These classes were similar to those considered in the country-wide map produced by the Iranian Space Agency (ISA) using Moderate Resolution Imaging Spectroradiometer (MODIS) imagery. The training and test datasets for the classification were visually identified and delineated using high resolution imagery available in ArcGIS and the ISA land cover map. The locations of the training and test samples were selected based on the existing map of the ISA and visual interpretation of the high-resolution images available in ArcGIS and Google Earth. Moreover, it was attempted to delineate samples (polygons) from different parts of Iran to properly represent various land cover types in the country. Figure 1 illustrates the distribution of training and test data and Table 1 provides the numbers and areas of samples considered for each

class. As is clear from this table, approximately 50 km^2 was considered when selecting the samples for each class.

NO.	Class	# polygons		Area (km^2)	
		Train	Test	Train	Test
1	Clay	19	25	22.40	27.49
2	Dense Forest	18	29	24.74	24.85
3	Farm Land	35	25	28.42	21.19
4	Kalut	30	29	26.76	23.49
5	Marshland	32	36	41.42	11.71
6	Outcrop	36	44	24.18	26.69
7	Range Land	55	41	25.83	28.49
8	Salty Land	27	29	25.69	24.08
9	Sand	18	20	22.78	26.61
10	Uncovered Plain	27	34	20.52	29.06
11	Urban	53	66	22.62	26.19
12	Water	47	30	23.50	26.70
13	Wetland	34	23	27.32	19.54

Table 1. The total number of training and test samples used in this study to classify Iran into 13 classes.

2.3 Satellite Data

The Landsat 8 satellite was launched in 2013 as a collaboration between the National Aeronautics and Space Administration (NASA) and United States Geological Survey (USGS). This satellite carries two imaging sensors: the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS). The OLI images the Earth surface in 9 spectral bands, including a panchromatic band with a 15-meter spatial resolution and 8 spectral bands with 30-meter spatial resolutions. TIRS collects the Earth surface radiance in two thermal bands with a 100-meter spatial resolution. The high signal-to-noise ratio of these sensors provides the opportunity to produce 12-bit images. The Landsat 8 imagery have been successfully used in a variety of applications, including forestry and agriculture, geology, land cover mapping, hydrology, and environmental monitoring.

In this study, the Landsat 8 images of Iran acquired between 2018-05-01 and 2018-11-30 (seven months) were used. Only the visible, Near Infrared (NIR) and Shortwave Infrared (SWIR) bands of Landsat 8 (bands 2-7) images were employed in the classification. Approximately, 94 Landsat-8 images cover the entirety of Iran. Moreover, ~188 Landsat-8 images are acquired over Iran each month and, thus, approximately 1,316 Landsat 8 images were processed within the GEE platform.

2.4 Methodology

Since an immense geographic region was analyzed in this research, it cannot be assumed that the atmospheric condition was consistent throughout the study area, precluding the use of the Digital Number (DN) values of pixels in lieu of the reflectance values (Amani et al., 2019). Therefore, multi-temporal Landsat 8 Surface Reflectance Tier 1 data were employed in this study. This product contains the atmospherically corrected, orthorectified surface reflectance for the visible, NIR and SWIR bands, and the orthorectified brightness temperature for the Thermal Infrared (TIR) bands. Moreover, cloud, shadow, water, snow, and saturated pixel masks are available for this product

(https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_SR). It is important to mask cloud and snow from optical satellite imagery before classification. This task was efficiently performed in this study because multi-date Landsat 8 images collected over seven months were employed. To this end, a median function was applied to downscale all datasets into a single mosaic image for which median temporal value was used to represent each pixel. A median function was selected to produce an image with no/least cloud and snow covers. Additionally, median function removed very dark or bright pixels which might be due to shadow and haze, respectively.

To date, various machine learning algorithms have been employed for land cover classification. In this regard, RF has demonstrated a high potential for classifying various land covers (Amani et al., 2017; Mahdavi et al., 2017; Amani et al., 2018). RF is a commonly used classifier that consists of several decision trees containing multiple nodes. In each of these nodes, pixels are divided into mutually exclusive groups. The bottom node of each tree represents one of the final classes (Breiman, 2001). RF has several tuning parameters, of which the number of trees and the number of variables in each node are the most important for classification (Amani et al., 2017). Based on trial and error and considering the limitations of GEE, 80 and the square root of the number of variables were selected as the most optimal values of the two mentioned parameters, respectively. It is worth noting that 50% of ground truth data were randomly selected to train the classification algorithm, and the remaining 50% were used for accuracy assessment (see Table 1).

Accuracy assessment of the produced maps was conducted both visually and statistically. For the visual assessment, the generated map was compared to the existing land cover map provided by the ISA, as well as high-resolution images available in Google Earth™ and Arcmap™. For the statistical accuracy assessment, 50% of the ground truth data were applied to calculate the confusion matrix. The three parameters of overall, producer, and user accuracies were then reported for the final classified map.

3. RESULT AND DISCUSSION

Figure 2 illustrates the final Iran-wide land cover map produced using multi-temporal Landsat 8 images at a 30 m spatial resolution. As is clear, 13 land cover classes were considered in the final map. Based on the visual analysis and comparison with high resolution imagery, it was observed that the classes were reasonably accurately identified. For instance, the northern forested part of the country was correctly classified as Dense Forest. Water bodies were also correctly identified, and their surrounding areas were classified as either Wetland or Marshland.
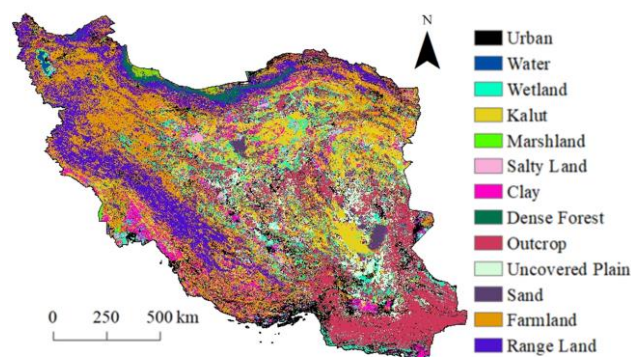

Figure 2. Iran-wide land cover map produced using Google Earth Engine.

The overall classification accuracy and Kappa Coefficient for the produced Iran-wide map were 74% and 0.71, respectively. This level of accuracy was reasonable considering the large area that was classified into 13 various land cover types in this study. The producer and user accuracies for each class are also provided in Table 2. As is clear, most of the classifications had relatively high accuracies. In this regard, the classification of Dense Forest, Water, and Sand had the highest accuracies, respectively. However, the lowest classification accuracies were obtained for Marshland and Clay, each with producer and used accuracies lower than 50%. One of the main reasons for inconsistent producer and user accuracies for different classes rooted in the fact that the training and test polygons were generated using visual interpretations and there might be uncertainty for some of the classes. Therefore, the samples collected in the field, with a high level of accuracy, should be included in the future studies to improve the level of accuracies and produced a more reliable Iran-wide landcover map. Additionally, some of the classes are spectrally similar and there is considerable confusion between these classes. This fact increased the misclassified samples for some of the classes and reduced the producer and used accuracies.

Class	Producer Accuracy (%)	User Accuracy (%)
Urban	73	81
Water	80	91
Wetland	49	56
Kalut	65	56
Marshland	33	50
Salty Land	75	67
Clay	43	52
Dense Forest	90	95
Outcrop	73	70
Uncovered Plain	56	69
Sand	75	86
Farmland	71	63
Range Land	50	78
Average	64	70

Table 2. The producer and user accuracies of the classifications obtained from Iran-wide land cover map.

The current study is the first to classify the land cover types throughout Iran using GEE. Several approaches to increase the accuracy of the produced map were identified during this research and will be addressed in future studies. For instance, researchers have extensively argued that a combination of

optical and SAR data is optimal for producing accurate land cover maps. Thus, future studies should combine Landsat 8, Sentinel 2, and Sentinel 1 datasets to produce a more accurate Iran-wide map. Additionally, including more features (e.g., spectral indices, ratio features, and textural features) usually increases the classification accuracy and should be further investigated in future studies of Iranian land cover types. Furthermore, in this study, the training and test samples were identified through visual analysis, which decreased the reliability of the ground truth data. It is therefore recommended that samples used to construct classifications in the future be collected through field surveys. Finally, it should be noted that several of the 13 classes considered in this study are spectrally similar and should be merged in future classifications to increase accuracy.

4. CONCLUSION

Remote sensing data have proved significantly useful for mapping and monitoring various land cover types. In this study, the first Iran-wide land cover map was produced using high level machine learning algorithms within GEE. It was concluded that this platform offers a high potential for large-scale land cover mapping and monitoring that requires the processing of hundreds to thousands of satellite images. In the future, studies should aim to improve classification accuracies by incorporating other sources of satellite data, particularly SAR and high-resolution Digital Elevation Model (DEM) images, optimizing classifier inputs, and developing more advanced machine learning algorithms. Once the accuracy of the map presented herein has been improved, the proposed method can also be used as a basis to begin operational monitoring of changes in Iranian land cover types.

ACKNOWLEDGEMENTS

REFERENCES

Akbari, M., Mamanpoush, A.R., Gieske, A., Miranzadeh, M., Torabi, M., Salemi, H.R., 2006. Crop and land cover classification in Iran using Landsat 7 imagery. International Journal of Remote Sensing, 27(19), 4117-4135.

Alizadeh-Choobari, O., Najafi, M.S., 2018. Extreme weather events in Iran under a changing climate. Climate dynamics, 50(1-2), 249-260.

Amani, M., Mahdavi, S., Afshar, M., Brisco, B., Huang, W., Mohammad Javad Mirzadeh, S., White, L., Banks, S., Montgomery, J., Hopkinson, C., 2019. Canadian Wetland Inventory using Google Earth Engine: The First Map and Preliminary Results. Remote Sens. 11, 842.

Amani, M., Salehi, B., Mahdavi, S., Brisco, B., Shehata, M., 2018. A Multiple Classifier System to improve mapping complex land covers: a case study of wetland classification using SAR data in Newfoundland, Canada. International journal of remote sensing, 39(21), 7370-7383.

Amani, M., Salehi, B., Mahdavi, S., Granger, J. E., Brisco, B., Hanson, A., 2017. Wetland classification using multi-source and multi-temporal optical remote sensing data in Newfoundland and Labrador, Canada. Canadian Journal of Remote Sensing, 43(4), 360-373.

Breiman, L. 2001. Random forests. Machine learning, 45, 5-32.
Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., Zhang, W., 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 103, 7-27.

Eisavi, V., Homayouni, S., Yazdi, A.M., Alimohammadi, A., 2015. Land cover mapping based on random forest classification of multitemporal spectral and thermal images. Environmental monitoring and assessment, 187(5), 291.

Ghorbanian, A., Mohammadzadeh, A., 2018. An unsupervised feature extraction method based on band correlation clustering for hyperspectral image classification using limited training samples. Remote sensing letters, 9(10), pp.982-991.

Gomez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 116, 55-72.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27.

Haftlang, K.K., Lang, K.K.H., 2003. The book of Iran: a survey of the geography of Iran. Alhoda UK.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A.A., Tyukavina, A., Kommareddy, A., 2013. High-resolution global maps of 21st-century forest cover change. science, 342(6160), 850-853.

Kumar, L., Mutanga, O., 2018. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. Remote Sensing, 10(10), 1509.

Mahdavi, S., Salehi, B., Amani, M., Granger, J. E., Brisco, B., Huang, W., Hanson, A., 2017. Object-based classification of wetlands in Newfoundland and Labrador using multi-temporal PolSAR data. Canadian Journal of Remote Sensing, 43(5), 432-450.

Mahdavi, S., Salehi, B., Granger, J., Amani, M., Brisco, B., Huang, W., 2018. Remote sensing for wetland classification: A comprehensive review. GIScience & remote sensing, 55(5), 623-658.

Mason, P.J., Manton, M., Harrison, D.E., Belward, A., Thomas, A.R., Dawson, D.K., 2003. The second report on the adequacy of the global observing systems for climate in support of the UNFCCC. GCOS Rep, 82.

Minaei, M., Kainz, W., 2016. Watershed land cover/land use mapping using remote sensing and data mining in Gorganrood, Iran. ISPRS International Journal of Geo-Information, 5(5), 57.

Rahdari, V., Soffianian, A., Pourmanafi, S., Mosadeghi, R., Mohammadi, G.H., 2018. A hierarchical approach of hybrid image classification for land use and land cover mapping. Geographica Pannonica, 22(1), 30-39.

Saadat, H., Adamowski, J., Bonnell, R., Sharifi, F., Namdar, M., Ale-Ebrahim, S., 2011. Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 66(5), 608-619.

Sterling, S.M., Ducharne, A., Polcher, J., 2013. The impact of global land-cover change on the terrestrial water cycle. Nature Climate Change, 3(4), 385.

Woźniak, E., Kofman, W., Lewinski, S., Wajer, P., Rybicki, M., Aleksandrowicz, S., Włodarkiewicz, A., 2018. Multi-temporal polarimetry in land-cover classification. International journal of remote sensing, 39(22), pp.8182-8199.

Wulder, M.A., Coops, N.C., Roy, D.P., White, J.C., Hermosilla, T., 2018. Land cover 2.0. International Journal of Remote Sensing, 39(12), 4254-4284.

Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Thau, D., 2017. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS Journal of Photogrammetry and Remote Sensing, 126, 225-244.

Yousefi, S., Khatami, R., Mountrakis, G., Mirzaee, S., Pourghasemi, H.R., Tazeh, M., 2015. Accuracy assessment of land cover/land use classifiers in dry and humid areas of Iran. Environmental monitoring and assessment, 187(10), 641.

Zakeri, H., Yamazaki, F., Liu, W., 2017. Texture analysis and land cover classification of Tehran using polarimetric synthetic aperture radar imagery. Applied Sciences, 7(5), 452.