Assessing the Accuracy of the Esa Worldcover 2021 for the Local Region of Lalapasa/Edirne, Turkey and Recommending Possible Accuracy Improvement Strategies

Paria Ettehadi Osgouei
History Department
Koc¸ University
Istanbul, Turkey
pettehadi@ku.edu.tr

Elif Sertel
Geomatics Engineering Department
Istanbul Technical University
Istanbul, Turkey
sertele@itu.edu.tr

M. Erdem Kabadayi
History Department
Koc¸ University
Istanbul, Turkey
mkabadayi@ku.edu.tr

***Abstract*—Global Land Cover (LC) datasets are important geo-information sources for environmental, climate, agriculture, and landscape applications. The ESA WorldCover project (2020 and 2021) provides a global scale land cover map with the pre- defined 11 generic classes for almost the current state. This study aims to evaluate the accuracy of the ESA LC data for a local region in Lalapasa/Edirne to provide insights into this data for possible local-level applications. Our study revealed that while the grassland, shrubland, and bare classes have inaccuracies that need to be further addressed, tree cover, water bodies, and cropland LC classified were correctly mapped for the studied region. We proposed strategies to improve the accuracy of some classes in the ESA LC map with integrated usage of open geospatial datasets and object-based classification. We encourage merging segments with their best-fitting surrounding segments if they are smaller than a minimum mapping unit of 1 ha. By doing this, we aim to improve the representation of the integrity and compactness of built-up regions and agricultural lands.**

***Index Terms*—Land cover map, ESA WorldCover, Accuracy Assessment, Accuracy Improvement, Object Oriented**

I. Introduction

Land use and land cover (LULC) is a significant environmental variable that impacts cross-cutting topics and disciplines such as land management, climate change, agriculture, hydrology, ecology, disaster response, biodiversity, and sustainable development [1], [2], [3], [4]. Accurate representation of LULC is extremely important to understand the impacts of weather, climate, hydrology, and carbon cycle, to model biospheric emission, to characterize the landscape, and to reliably quantify future land use projections [5], [6], [7]. To address the need for accurate, timely, and high-resolution geospatial information on land cover at the global scale, the European Space Agency (ESA) developed the WorldCover (WC) project. The ESA WC dataset, a global land cover map at 10m resolution, was produced and validated using freely available Sentinel-1 (S1) radar data and Sentinel 2 (S2) optical data. ESA WC datasets improve land cover maps derived solely from optical data by using S1 radar data. Two versions of the ESA WC dataset were created for 2020 and 2021, using different algorithm versions (v100 and v200, respectively) (Buchhorn et al., 2020). The new ESA WC (2021) had an improved classification accuracy overall, increasing from 74.4% to 76.7%. While ESA WC products are widely used, there may be inconsistencies at the local and regional levels across these databases. It is important to assess their precision and applicability for tracking changes in regional land cover. This research aims to assess the accuracy of the ESA 2021 LC map of the Lalapasa district and further investigate the errors and misclassifications in this map to propose strategies for improving the ESA LC map of the study region. We recommend revisions for the following:

- highly mixed LC classes, namely shrubland-grassland, and built-up-bareland,
- underestimated roads as part of the built-up class,
- overestimated LC classes, specifically inside the spatially extensive LC classes.

II. study area

Lalapasa is a district of Edirne province in Turkish Thrace. It is 27 kilometers north of Edirne's downtown and located in the city's north (Fig. 1), neighboring Bulgaria to the north and west. The district's population decreased from 8406 (in 2007) to 6358 (in 2021). Edirne's economy is mainly based on agriculture, and most of the land cover of the Lalapasa district is covered with agricultural lands. Within the current demographic and socio-economic context of Southeast Europe, and especially in Thrace, rural depopulation and agricultural land abandonment are major threats to sustainable agricultural production. In our chosen region, a local accuracy assessment and possible improvements of the ESA LC data would have concrete additional benefits. Accurate and up-to-date information on land use and land cover is essential for sustainable land use planning and management. Improving the accuracy of the ESA LC data for the region could help identify areas of agricultural land that are at risk of abandonment, as well as areas that are suitable for sustainable land use practices such

as agroforestry, conservation agriculture, or precision farming [8]. This information could be used to support policies and programs aimed at promoting sustainable agriculture and rural development, such as land consolidation, rural infrastructure development, or support for small-scale farming.

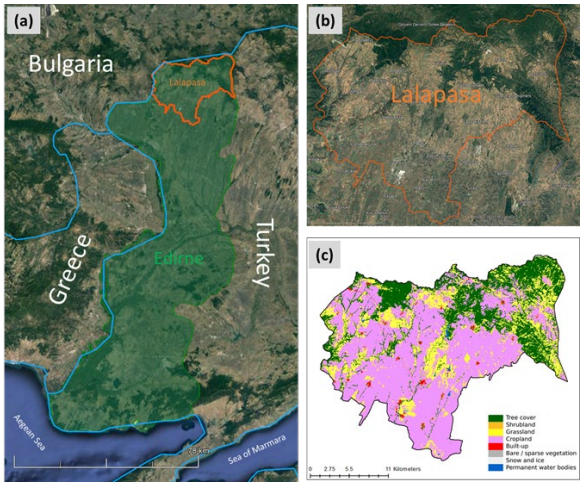


Fig. 1. Study area (a) Edirne province (highlighted by green polygon), (b) Lalapasa district boundaries, and (c) ESA LC 2021 map of Lalapasa district.

III. METHODOLOGY

We evaluated the ESA 2021 LC map of the Lalapasa district of Edirne/Turkey by identifying 290 reference points representing different LC classes. The reference class values of these points were obtained from up-to-date, very high-resolution Google Earth images. The ESA LC map includes 11 generic classes: ”bare/sparse vegetation,” ”built-up,” ”cropland,” ”grassland,” ”herbaceous wetland,” ”mangrove,” ”moss and lichen,” ”permanent water bodies,” ”shrubland,” ”snow and ice,” and ”tree cover.” We randomly generated 50 sample points per class for tree cover, cropland, grassland, and built-up and 30 points per class for permanent water bodies, bare/sparse vegetation, and shrubland. According to the ESA LC 2021 map of the Lalapasa region, eight LC classes are available: bare/sparse vegetation, built-up, cropland, tree cover, grassland, shrubland, permanent wetland (excluding herbaceous wetland, mangrove, moss and lichen, and snow and ice). We formed a confusion matrix using 290 reference points and calculated the overall user’s and producer’s accuracy values. To propose quality improvements for the ESA LC map of the Lalapasa district, we consider benefiting from freely available geospatial data and analyzing potential datasets. We suggest the integration of the OpenStreetMap (OSM) road dataset to improve the road and highway network details as the part of built-up class [9]. The ESA LC map was generated through a pixel-based approach and has a spatial resolution of 10 m, in some cases, which resulted in salt and pepper effects. We propose to use an object-scale Minimum Mapping Unit (MMU) to improve the classification accuracy of some classes by following the majority approach. We reclassified the spatially-small LC pixels by merging small tree covers or grasslands within larger cropland or built-up areas. We also propose dividing tree class into the forest and permanent crop to differentiate between forest and agricultural trees, considering spectral and textural aspects, in studying changes in agricultural dynamics [10]. The differentiation of permanent crops, such as olive trees and grapevines, within tree cover class is crucial for better utilization of the ESA LC data for agricultural mapping, agricultural land abandonment analysis, forest monitoring, and regional and urban planning needs. In the wider Mediterranean climate zone and in some areas along the coast of the Black Sea, particularly in Georgia and Crimea, vineyards and olive trees are important components of agricultural production. Understanding the extent and location of vineyards and olive growing areas, as well as other permanent crops, is essential for sustainable land use planning, management, and monitoring in these regions [11]. Moreover, accurate mapping of agricultural land cover and land use could support the development of sustainable agri-environmental policies and programs, such as payment for ecosystem services, conservation of biodiversity, or promotion of sustainable land use practices. By identifying and mapping permanent crops, such as vineyards and olive groves, we can better understand the distribution and contribution of these crops to the local and regional economies, as well as to ecosystem services such as soil conservation and carbon sequestration.

IV. RESULT

A. Accuracy Assessments of the ESA LC 2021 Map in Lalapasa

Our initial visual examination of the ESA LC 2021 map illustrated that the LC map for the Lalapasa district was satisfactory on a broad scale. The built-up regions, agricultural lands, forest areas, and water bodies were adequately detected and mapped. We observed significant misclassifications of the shrubland as grassland/tree cover while there were slight confusions between the built-up region and bare/sparse vegetation LC class, we have already pointed to these mixed class problems in our previous work [12]. The ESA LC 2021 classification accuracy of the Lalapasa district derived from the error matrix is presented in Table I. The overall accuracy is 70.2%, and the Kappa value is 0.57. User’s and producer’s accuracy values of individual classes vary from 6.7% to 100%. With the exception of the LC classes for shrubland, grassland, and bare/sparse vegetation, the classification map’s user and producer accuracies are generally good.

B. ESA LC classification improvements

The initial visual examinations of the ESA LC map revealed that the shrubland and bare land are significantly misclassified. A post-editing phase is recommended for these LC classes, which were mainly mixed and misclassified as grassland or built-up, respectively. The shrubland and grassland LC classes can be revised using multi-temporal Sentinel 2A and Google Earth VHR images. We identified misclassifications of small

TABLE I
SUMMARY OF THE CLASSIFICATION ACCURACIES FOR ESA LC 2021 OF LALAPASA TOWN

Land cover class	User's Acc. (%)	Producer's Acc. (%)
tree cover	84.48	90.74
shrubland	28.57	6.67
grassland	42.35	67.92
cropland	66.67	75.00
built-up	76.60	70.59
bare/sparse vegetation	83.33	33.33
permanent water bodies	83.33	100

objects within the segments, which are assigned to built-up and cropland classes (Fig. 2 and Fig. 3). By specifying the MMU of 1 ha, the small group of pixels surrounded by larger meaningful segments (built-up objects in Fig. 2 examples and agricultural land objects in Fig. 3) with relative border to neighboring segment greater than 0.45 can be merged with pixels in the surrounding LC classes.

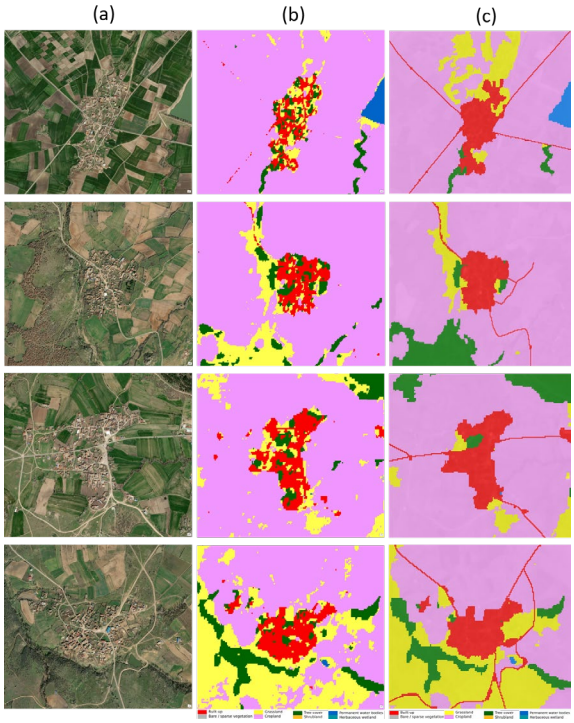


Fig. 2. Sample locations of misclassified inner pixels (grassland, cropland, tree cover) of a superior land cover (built-up area) (b) panels against the improved ESA classification (c) and Google Earth imagery (a).

Roads and highways were mapped in the ESA LC map as parts of the built-up class. Except for some major highways and access roads, the road networks were mixed with sur- rounding land covers. Fig. 4 shows examples of areas where roads are misclassified. This error could be due to either the unrepresentative reference samples or the insufficient spatial resolution of the input data in which the roads are not spatially or spectrally distinct from adjacent land cover classes. We propose the use of the OSM road dataset to improve the

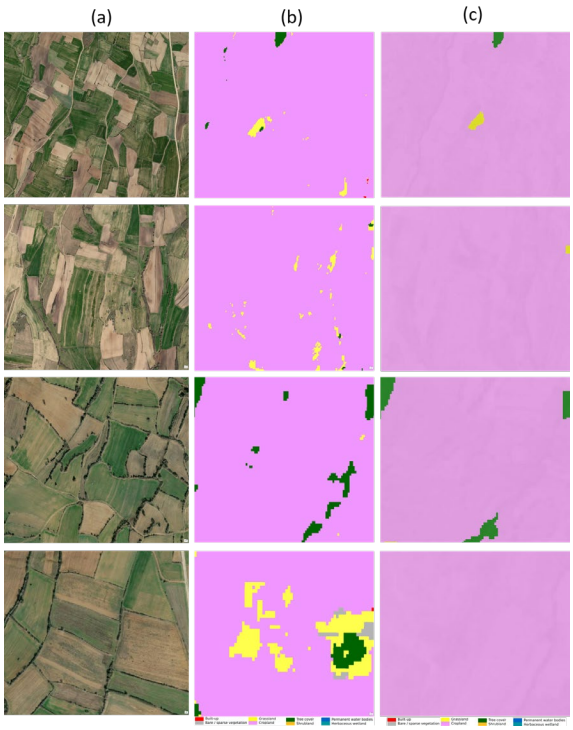


Fig. 3. Sample locations of misclassified inner pixels (grassland, cropland, tree cover) of a superior land cover (cropland) (b) and (e) panels against the improved ESA classification (c, f) and Google Earth imagery (a, d).

classification accuracy of built-up LC and better represent transportation networks. The use of OSM road data leads to significant improvement of the built-up class of the ESA land cover classification (Fig. 4).

C. ESA LC classification improvements

Improvement yielded by the new approach lead to a better classification accuracy of the ESA 2021 LC map of the Lalapasa district. The enhanced ESA LC 2021 classification accuracy of the Lalapasa is presented in Table II. The overall accuracy improved to 86.01%, and the Kappa value improved to 0.83. User's and producer's accuracy values of individual classes vary from 76% to 100%. Accuracy for the shrubland LC class was clearly improved. This is also true for the bare/sparse vegetation and built-up LCs, with 10–177% improvement in user's and producer's accuracies.

V. CONCLUSION

Environmental, climatic, agricultural, and landscape applications rely extensively on the global land cover (LC) dataset. The ESA WC project offers a map of land cover on a worldwide scale with predefined classes, but the accuracy at the local level can be increased by integrating free geospa- tial datasets and object-based classification. In the instance of the Lalapasa district in Edirne, the first ESA LC map was adequate on a large scale, but there were errors in the shrubland, grassland, and bare/sparse vegetation classes. The

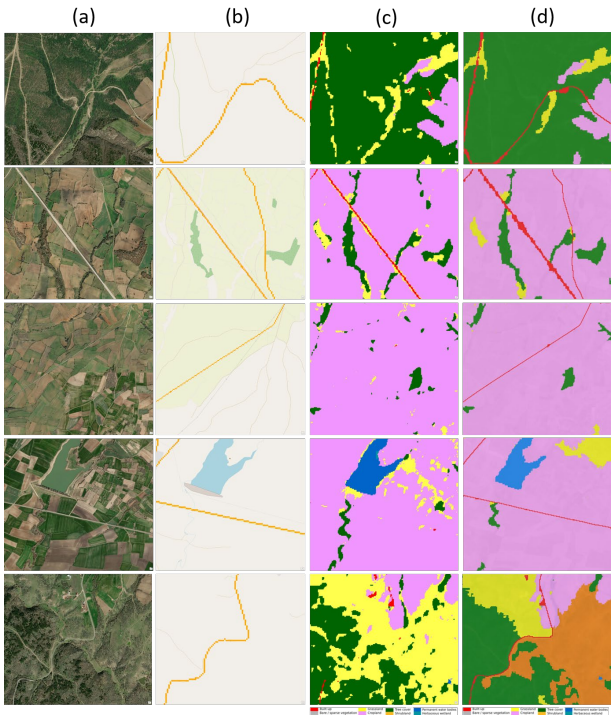


Fig. 4. Sample locations of improved ESA classification (d) against misclassified road LC class ESA classification (c) panel, Google Earth imagery (a) panel, and OSM roads highlighted with orange lines on OpenStreetMap (b) panel.

TABLE II
SUMMARY OF THE CLASSIFICATION ACCURACIES FOR AN IMPROVED VERSION OF THE ESA LC 2021 OF LALAPASA TOWN.

Land cover class	User's Acc. (%)	Producer's Acc. (%)
tree cover	96.15	89.29
shrubland	68.75	84.62
grassland	84.48	75.38
cropland	88.68	83.93
built-up	84.31	91.49
bare/sparse vegetation	83.33	92.59
permanent water bodies	90.00	100

classifications of these LCs should be revised using the multi-temporal Sentinel 2A and Google Earth VHR images in a post-editing phase to address the misclassifications. The LC classes for agriculture, built-up water, and tree cover, on the other hand, were accurately mapped. However, to improve the depiction of the integrity and compactness of built-up areas and croplands we suggest a technique to combine segments smaller than MMU with their best-fitting neighbors. Besides, we report the significant improvements of the built- up class by the incorporation of OSM road dataset. These improvements resulted in an increase in the upgraded ESA LC 2021 classification accuracy for the Lalapasa district to 86.01% and an increase in the Kappa value to 0.83. This made the classification map more dependable and appropriate for local-level applications. For accurate classification and mapping of land cover types to track changes in agricultural dynamics, we also think that dividing the tree class into the two subclasses of forest and permanent crop is essential.

VI. ACKNOWLEDGMENTS

This work was supported by the European Research Council (ERC) project: "A GeoAI-based Land Use Land Cover Segmentation Process to Analyse and Predict Rural Depopulation, Agricultural Land Abandonment, and Deforestation in Bulgaria and Turkey, 1940-2040" under the European Union's Horizon 2020 research and innovation program Grant Agreement No. 101100837, acronym GeoAI_LULC_Seg.

REFERENCES

[1] S. S. Hasan, L. Zhen, M. G. Miah, T. Ahamed, and A. Samie, "Impact of land use change on ecosystem services: A review," *Environmental Development*, vol. 34, p. 100527, 2020.
[2] W. Song and X. Deng, "Land-use/land-cover change and ecosystem service provision in china," *Science of the Total Environment*, vol. 576, pp. 705–719, 2017.
[3] X.-P. Song, M. C. Hansen, S. V. Stehman, P. V. Potapov, A. Tyukavina, E. F. Vermote, and J. R. Townshend, "Global land change from 1982 to 2016," *Nature*, vol. 560, no. 7720, pp. 639–643, 2018.
[4] N. Clerici, F. Cote-Navarro, F. J. Escobedo, K. Rubiano, and J. C. Villegas, "Spatio-temporal and cumulative effects of land use-land cover and climate change on two ecosystem services in the colombian andes," *Science of the Total Environment*, vol. 685, pp. 1181–1192, 2019.
[5] P. Ettehadi Osgouei, E. Sertel, and M. E. Kabadayı, "Integrated usage of historical geospatial data and modern satellite images reveal long-term land use/cover changes in bursa/turkey, 1858–2020," *Scientific Reports*, vol. 12, no. 1, p. 9077, 2022.
[6] M. Buchhorn, M. Lesiv, N.-E. Tsendbazar, M. Herold, L. Bertels, and B. Smets, "Copernicus global land cover layers—collection 2," *Remote Sensing*, vol. 12, no. 6, p. 1044, 2020.
[7] D. Sulla-Menashe, J. M. Gray, S. P. Abercrombie, and M. A. Friedl, "Hierarchical mapping of annual global land cover 2001 to present: The modis collection 6 land cover product," *Remote Sensing of Environment*, vol. 222, pp. 183–194, 2019.
[8] S. Li and X. Li, "Global understanding of farmland abandonment: A review and prospects," *Journal of Geographical Sciences*, vol. 27, pp. 1123–1150, 2017.
[9] "OpenStreetMap contributors, 2021," planet dump retrieved from https://planet.osm.org. [Online]. Available: https://www.openstreetmap.org.
[10] M. E. Kabadayı, P. Ettehadi Osgouei, and E. Sertel, "Agricultural land abandonment in bulgaria: A long-term remote sensing perspective, 1950–1980," *Land*, vol. 11, no. 10, p. 1855, 2022.
[11] U. Alganci, G. N. Kuru, I. Yay Algan, and E. Sertel, "Vineyard site suitability analysis by use of multicriteria approach applied on geo-spatial data," *Geocarto International*, vol. 34, no. 12, pp. 1286–1299, 2019.
[12] P. Ettehadi Osgouei, S. Kaya, E. Sertel, and U. Alganci, "Separating built-up areas from bare land in mediterranean cities using sentinel-2a imagery," *Remote Sensing*, vol. 11, no. 3, p. 345, 2019.