5th Intercontinental Geoinformation Days

igd.mersin.edu.tr

A quantitative and qualitative assessment from official statistics to spatial statistics: Agricultural greenhouses detection over time integrating of remote sensing and transfer learning-based machine learning approach

Fuat Kaya*[1], Gordana Kaplan [2], Levent Başayiğit [1]

[1]Isparta University of Applied Sciences, Faculty of Agriculture, Department of Soil Science and Plant Nutrition, Isparta, Türkiye
[2]Eskisehir Technical University, Institute of Earth and Space Sciences, Eskişehir, Türkiye

Keywords
Remote sensing
Greenhouse mapping
Spatial statistics
PlanetScope
Sentinel 2A
Spatiotemporal dynamics

Abstract

The availability of medium-resolution satellite data such as the open-access Sentinel-2 as well as high-resolution commercial satellite imagery from PlanetScope presents significant opportunities for the agricultural sector and allows us to gain insight into the land surface, land use, and their management. Agricultural Greenhouses (AGs) are the fastest-growing food or commercial ornamental production approach around cities, driven by different factors. The current study is aimed to conduct temporal greenhouse maps in the covered upland greenhouse region (Isparta-Deregümü region, Southwestern Türkiye) from 2016 to 2021 using open-access Sentinel 2 and PlanetScope imagery and machine learning algorithm in a high-performance computing environment (R Core Environment). As a result of the qualitative evaluation of satellite imagery with two different spatial and spectral resolutions, PlanetScope, which has a higher spatial resolution, was determined to be useful in the detection of AGs. Temporal greenhouse maps were generated using random forest algorithm at two-time periods, and the overall accuracy of the predictions was around 90%. While the total greenhouse area in the current area increased by 47% from 2016 to 2021 in official statistics, the methodology allowing to obtain spatial statistics detected this increase by 76%. The current study significantly improves the link between spatial statistics and official statistics.

1. Introduction

Sustainable agricultural production constantly needs up-to-date spatial data that is key to various land management decisions (İbrahim ve Gobin 2021). Agricultural Greenhouses (AGs), one of the pioneers of controlled agriculture, have been expanding uncontrollably in different geographies in recent years (Li et al., 2020). The significant increase in greenhouses in certain urban areas not only changes the shape of the seasonal food supply but also can completely change the landscape of farmland in expanding regions (Ou et al., 2019).

In this context, to relieve the pressure of agricultural production on other sectors and to maintain the balance between food production and the environment, scientists in different disciplines use up-to-date technologies to make definitive decisions about the extent of greenhouse expansion. Thus, it may be necessary to better understand the temporal-spatial dynamics of a city in a short time on a regional scale. Spatial effective monitoring of greenhouses can only be done most accurately by integrating Geographic Information System (GIS) and Remote Sensing (RS) data with advanced statistical learning methods (Kurucu and Erden 2015).

In monitoring greenhouses from satellite images (Aguilar et al., 2020), current studies tend to offer many different methods that integrate machine learning to improve classification accuracy and determine their dynamics in the long term (Koc-San 2013). Open-access Sentinel and Landsat series satellite images have been the subject of studies with medium spatial resolutions (Aguilar et al., 2020; İbrahim and Gobin 2021).

This study aims to i) present a qualitative assessment using PlanetScope and Sentinel series satellite images for temporal greenhouse mapping in an area where typical highland greenhouse cultivation is active (Buyurgan et al., 2019); ii) identify greenhouse areas and present a quantitative assessment with machine learning and transfer of the learned model, iii) determine the distribution of greenhouses between 2016

* **Corresponding Author**

*(fuatkaya@isparta.edu.tr) ORCID ID 0000-0003-0011-9020
(kaplangorde@gmail.com) ORCID ID 0000-0001-7522-9924
(leventbasayigit@isparta.edu.tr) ORCID ID 0000-0003-2431-5763

Cite this study

Kaya F, Gordana K & Başayiğit L (2022). A quantitative and qualitative assessment from official statistics to spatial statistics: Agricultural greenhouses detection over time integrating of remote sensing and machine learning algorithm. 5th Intercontinental Geoinformation Days (IGD), 64-68, Netra, India

and 2021 by integrating spatial statistics information with official statistics.

2. Method

2.1. Study area

The current study area is located in the central part of Isparta Province, Türkiye (279504 - 285538 E, 4183775 - 4188045 N - WGS 1984 UTM North Zone 35) (Figure 1).

The study area has a typical Mediterranean climate, characterized by a rainy spring, a hot summer, and a cold and wet winter. Since it is close to the province of Antalya, which has a high vegetable-fruit export potential, production in the plateau starts in this region when the production on the Mediterranean coast ends in the greenhouses. The exporter can continue this export throughout the year (Buyurgan et al., 2019).

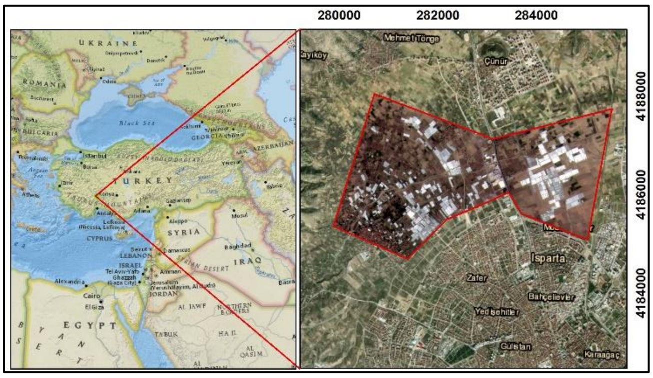

Figure 1. Location of the study area and distribution of AGs in PlanetScope satellite image (September 2021) with natural colours band combination

2.2. Data Acquisition for official statistics and remote sensing imagery

According to TURKSTAT (2022), while the greenhouse area in the center district of Isparta was 1715 decares in 2016, in 2021, it was recorded as 2520 decares (Figure 2). Considering the most recent data, the majority of the greenhouse areas (about 80%) are located in the central district of Deregümü, Çünür and Yakaören villages of Isparta district (Anonymous, 2022).

An increase of approximately 34% in the field of greenhouse agriculture was recorded in 2017 compared to 2016.

September 2016, 2017, and 2021 images of the open-access Sentinel 2A, which contains the short-wave infrared band sensitive to anthropogenic structures, and the PlanetScope satellite, which includes the visible and near-infrared band with higher spatial resolution, were provided (ESA 2022; Planet Team 2022).

2.3. Qualitative evaluation

An evaluation was presented considering both satellite images' spectral and spatial differences. A composite image was produced using the B12-8A-B4 (SWIR, NIR narrow, Red) band combination of the Sentinel 2 satellite. This process was performed on metafile satellite images taken at 20 meters resolution. Planet Scope satellite images were obtained at 3 m spatial resolution. From the Planet Scope satellite imagery, natural color band combination was created.

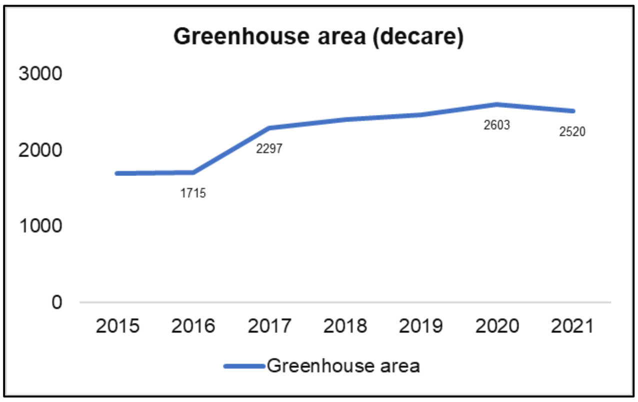

Figure 2. Agricultural greenhouses area amounts in Isparta Central District.

2.4. Quantitative evaluation

Supervised machine learning algorithms use data samples or experience to train classification results to optimize. In this study, random forest classifier was used for modeling. The "Greenhouse absence-presence" model, learned from the observations in 2021, was transferred to the 2016 images.

2.4.1. Dataset for training samples and modelling process

Control data were collected by photo-interpretation of "historical images" in Google Earth® in September 2021. 122 points were labelled as “non-greenhouse”, and 128 points were labelled as “Greenhouse”. To solve this binary classification problem, we used the "randomForest" package R (Breiman 2001; R Core Team 2022). The model was established by randomly choosing 70% of the observations as training points, while the rest of 30% were randomly selected as validation points to evaluate the accuracy of the classification results in 2021. Different criteria were used to evaluate the produced greenhouse maps for 2021 year in this study, including overall accuracy, kappa coefficient, producer’s accuracy, and user's accuracy (Congalton, 1991).

2.4.2. Area Changes Analysis

Pixels identified as "Greenhouse" in 2016 and 2021 were vectorized using the "ArcGIS - Arctoolbox-Conversion Toolbox-Raster to Polygon" tool and area calculations were converted in "decares". The obtained results were compared with TURKSTAT (2022) statistics. The methodological flow chart was presented in Fig. 3.

3. Results

3.1. Temporal greenhouse maps and their qualitative evaluation

After the generated multi-time band combination maps, evaluations are presented for 2 different satellites to explore the scene composition change of the greenhouses.

The results showed that there is a greenhouse increase in the area north of the study area, as it can be seen at Figure 4. This increase can be clearly distinguished qualitatively by the SWIR band combination of the Sentinel 2A satellite. Also in 2021, greenhouse areas increased significantly in the eastern region of the study area.

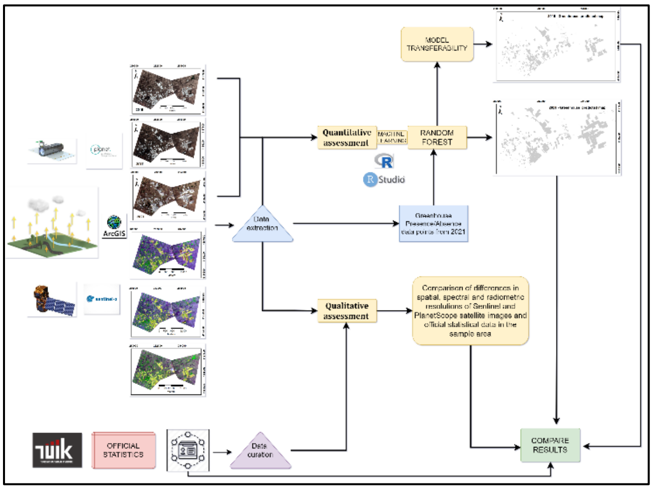

Figure 3. Flow chart of the methodological conducted for AGs detection and comparison with official statistics

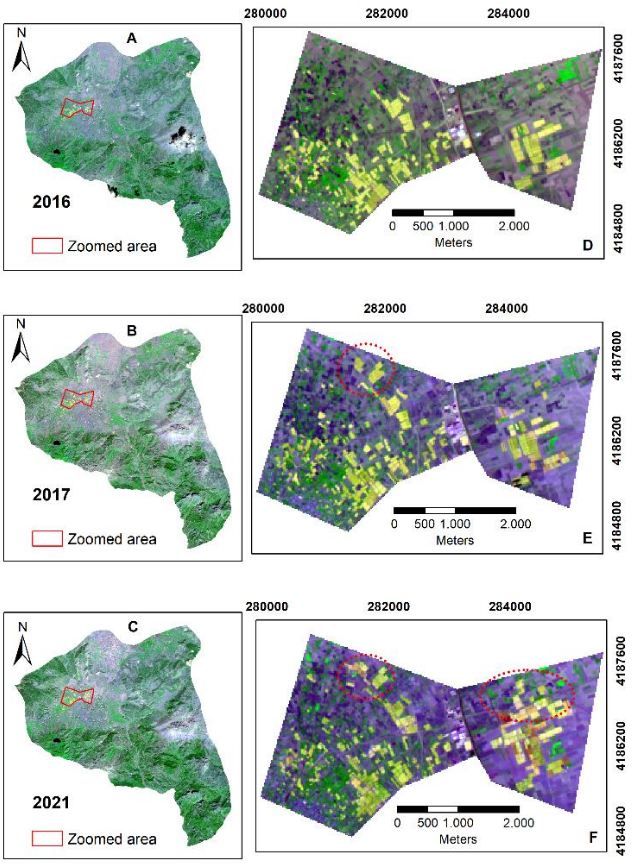

Figure 4. Qualitative evaluation of AGs over the years using the SWIR band combination of the Sentinel 2A-MSI satellite.

While Sentinel-2's band combination map containing the shortwave infrared bands was useful for the foto-interpretation, it was noteworthy that the Planet Scope satellite could better reflect the boundaries of greenhouses (Figure 5). This is expected as PlanetScope has significantly better spatial resolution than Sentinel-2.

Especially the time series (Figure 4) of Sentinel 2A satellite with SWIR band was not preferred for quantitative evaluation. Because it is seen that these images are very affected by the building material used in the greenhouses. Greenhouses with fine plastic materials can be mixed with bare soil in these images (Figure 4-E, F).

3.2. Greenhouse Maps and Their Quantitative evaluation and Area Change of Greenhouses during 2016–2021

Temporal greenhouse maps were produced using RF classifier for 2016 and 2021 (Figure 6). The overall accuracy of the RF model, which was also transferred to 2016, in 2021 is more than 90% and the user's accuracy of the greenhouse class is over 95%, representing map accuracy (Table 1). The performance metrics of classification process shown in Table 1 showed that the mapping results of the greenhouses could meet the needs for measuring their spatio-temporal dynamics in this study. It can be stated that classification errors occur in plate roofs of buildings and greenhouses where soil is seen a lot under the cover.

Looking at the official statistics (Table 1), 805 da increase in greenhouse area from 2016 to 2021 was determined while 895 da because of the Planet scope satellite being classified with the RF model (Table 1). This finding is an indication that the increase in greenhouse areas in the area of interest (AOI) we focus on is higher than in the district in general. Here, relatively, the presence of sheet roof structures may trigger an excessive increase. Because in our AOI via quantitative approach, a greenhouse area estimates of more than 80% of the official statistical value (Table 1) was made.

Table 1. Accuracy assessment of greenhouse maps (2021) and the total area of the greenhouses by comparing dates (O: Overall, A: Accuracy, P: Producer's, U: User's)

Accuracy assessment greenhouse map					
Model	Class	PA	UA	OA	Kappa
RF	Non-Greenhouse	95.54%	87.70%	90.18 %	0.80
	Greenhouse	89.13%	96.09%		
The total area changes of greenhouses					
	Official statistics		Spatial statistics		
Years	2016	2021	2016	2021	
AGs (da)	1715	2520	1170	2065	

3.3. Importance levels of satellite bands

The Blue band of the Planet Scope satellite was the most important variable in distinguishing plastic greenhouses planned for planting different plants, which had significantly higher "digital number" values (Figure 7-c) compared to non-greenhouse areas.

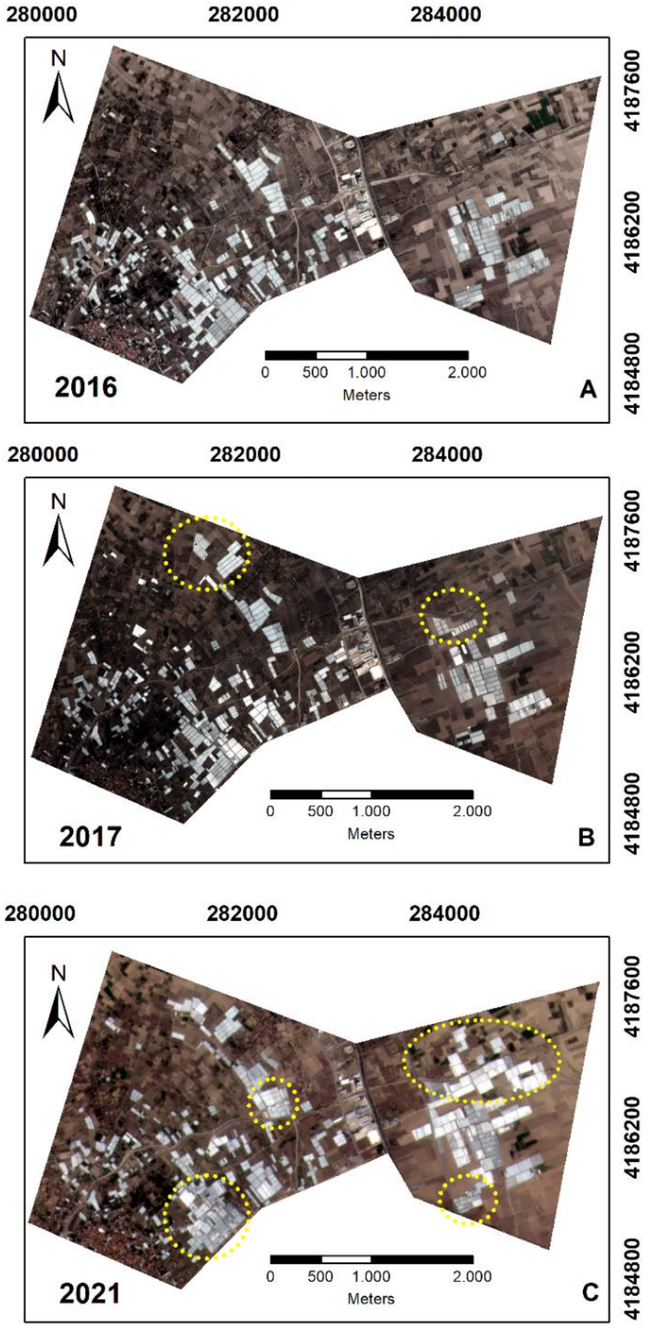

Figure 5. Qualitative evaluation of AGs over the years using the natural band combination of the PlanetScope satellite.

4. Discussion

Greenhouse mapping as time series and analysis of related dynamics are important for agronomists and different disciplines to understand and evaluate the sustainable development of protected agriculture.

The proposed multi-temporal greenhouse mapping and comparison with official statistics method used open access medium and high-resolution satellite images and RF supervised classification algorithm.

Greenhouse areas in 2017, where there was a sharp increase (about 34%) in official statistics, were qualitatively more noticeable with the Planet Scope satellite.

Quantitative calculations based on remote sensing in our study area, which includes the borders of 3 villages where production is intensified in the central district of Isparta, is expected to be below the official statistics (Table 1). The next goal here is to focus on more quantitative methods for the entire district center.

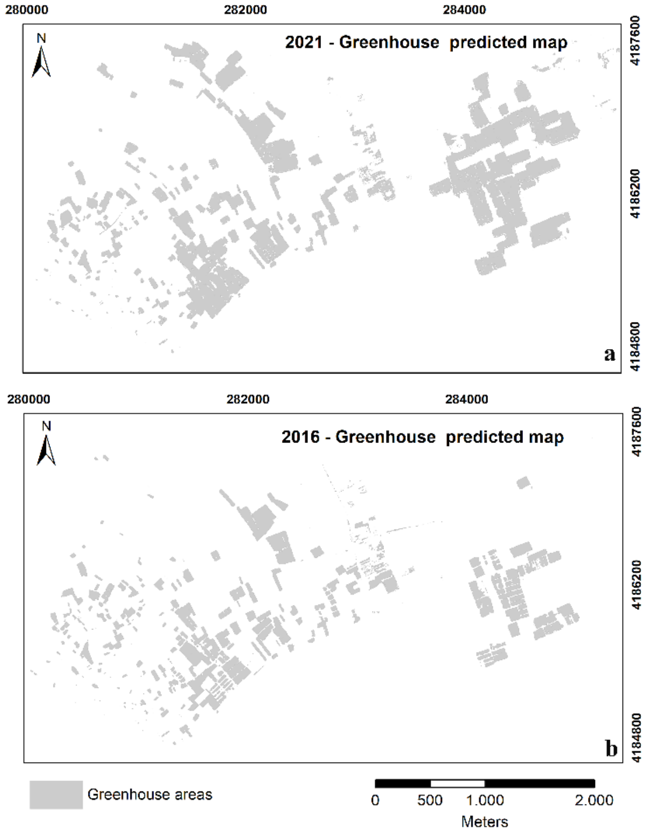

Figure 6. Temporal density of agricultural greenhouses expansion (a:2021, b:2016).

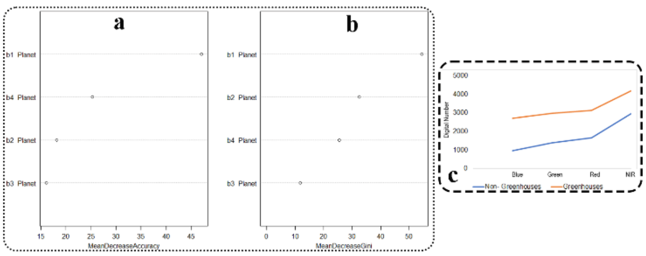

Figure 7. Importance levels of the PlanetScope bands used to predict the agricultural greenhouses using the RF algorithms (a-b). Digital number profile of agricultural greenhouse status (c).

5. Conclusion

In this study, we first made maps of the development of highland greenhouse cultivation in Isparta from 2016 to 2021. Compared to greenhouse detection studies that traditionally involve accuracy assessment, our study significantly advances the link between spatial distribution and official statistics. Furthermore, compared to studies based on official statistical data such as agricultural extension, our study includes spatial statistics of temporal remote sensing data and may contribute a more intuitive approach for future studies.

Acknowledgement

We acknowledge the usage of modified Copernicus Sentinel 2A Level 2A data and PlanetScope Ortho Scene Level 3B Product data.

References

Aguilar, M. Á., Jiménez-Lao, R., Nemmaoui, A., Aguilar, F. J., Koc-San, D., Tarantino, E., & Chourak, M. (2020). Evaluation of the Consistency of Simultaneously Acquired Sentinel-2 and Landsat 8 Imagery on Plastic Covered Greenhouses. Remote Sensing, 12, 2015. https://doi.org/10.3390/rs12122015

Anonymous. (2022). Yayla Şartlarında Örtüaltı Yetiştiriciliğinde Hasat Başladı. Accesssed on https://isparta.tarimorman.gov.tr/Haber/478/Yayla-Sartlarinda-Ortualti-Yetistiriciliginde-Hasat-Basladi

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32.

Buyurgan, K., Altunbaş, S. & Gözükara, G. (2019). Farklı fizyografya ünitelerinde gelişen yayla seracılığının zamansal ve mekansal değişiminin uzaktan algılama ve CBS teknikleri ile belirlenmesi: Elmalı/Antalya örneği. Derim, 36, 217-227. https://doi.org/10.16882/derim.2019.614303

Congalton R, (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37, 35–4

ESA (2022). European Space Agency. Copernicus Open Access Hub.

Ibrahim, E., & Gobin, A. (2021). Sentinel-2 Recognition of Uncovered and Plastic Covered Agricultural Soil. Remote Sensing, 13, 4195. https://doi.org/10.3390/rs13214195

Koc-San, D. (2013). Evaluation of different classification techniques for the detection of glass and plastic greenhouses from WorldView-2 satellite imagery. Journal of Applied Remote Sensing, 7, 073553. https://doi.org/10.1117/1.JRS.7.073553

Kurucu, Y., & Erden, H. (2015). Control of farmer statements integrated in national farmer register system by remote sensing data. 2015 Fourth International Conference on Agro-Geoinformatics (Agro-geoinformatics), 97–100. Istanbul, Türkiye. https://doi.org/10.1109/AgroGeoinformatics.2015.7248144

Ou, C., Yang, J., Du, Z., Liu, Y., Feng, Q., & Zhu, D. (2020). Long-Term Mapping of a Greenhouse in a Typical Protected Agricultural Region Using Landsat Imagery and the Google Earth Engine. Remote Sensing, 12, 55. https://doi.org/10.3390/rs12010055

Planet Team (2022) Planet Explorer.

R Core Team. (2022) R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria.

Turkstat. (2022). Turkish Statistical Institute, Agriculture database. Date of access: 17.10.2022, website: https://biruni.tuik.gov.tr/medas/?locale=tr