



Article

Estimating Small Farming Plots' Key Crop Production at a Regional Level Utilizing Sentinel Imagery in Southern Europe

Theodore A. Tsiligridis ^{1,*} , Sergio Godinho ² , Katerina Ainali ¹ and Rui Machado ²

¹ Informatics Laboratory, Department of Agricultural Economics and Rural Development, Agricultural University of Athens, 75 Iera Odos, 11855 Athens, Greece; kainali@aua.gr

² MED—Mediterranean Institute for Agriculture, Environment and Development & CHANGE—Global Change and Sustainability Institute, Institute for Advanced Studies and Research, Universidade de Évora, Largo dos Colegiais 2, 7004-516 Évora, Portugal; godinho.sergio@gmail.com (S.G.); rdpm@uevora.pt (R.M.)

* Correspondence: tsili@aua.gr

Abstract

This study utilizes Sentinel imaging to monitor small farming plots (<5 ha) and objectively estimate their significance in agricultural output. It develops, tests, and evaluates methodologies within and across ten southern European NUTS-3 regions—commonly referred to as prefectures, provinces, or departments—located in Greece, Italy, France, Portugal, and Spain. The study provides stakeholders with the necessary tools to estimate their spatial distribution, crop diversity, crop area extent, and agricultural yields. The approach is expanded as needed to similar landscapes, serving as a model for other European NUTS-3 regions and providing a comprehensive view of the remote sensing solution. It employs random forest crop identification and classification analysis to build crop-type maps and record the countless small agricultural plots. The Sentinel data utilized are from 20 NUTS-3 pilot regions selected across 11 EU countries, focusing on 29 estimated key crop areas. The research demonstrates remarkable classification accuracy for a previously identified wide range of key crop data across each of the selected NUTS-3 regions. Crop area estimates derived from field data were adjusted and stratified. The RS-based methodology reduces propagation errors in estimating the area and production of crop categories with low accuracy levels, enhancing the overall understanding of the crucial role of small farming plots. For highly accurate key crop products (per crop-type categorization, with $F_{Score} > 75\%$), production estimates are calculated by multiplying predicted self-reported crop yields by unbiased key crop small farming plot area estimates. However, for the referenced period of cultivation, only 16 out of the above 29 estimated key crop areas are recorded for SFs by the regional official statistics. In addition, of the 16 reported key crop areas listed above, eleven are from nine NUTS-3 regions spread across four southern EU countries, with the remaining five coming from four NUTS-3 regions of two non-southern EU countries. The analysis shows a strong correlation between estimated key crop areas from SFs reported in regional official statistics and Sentinel-based key crop area estimates obtained from small farming plots, with $R^2 = 0.96$ for 16 estimated key crop area datasets across all NUTS-3 regions in the EU and $R^2 = 0.98$ for 11 estimated key crop area datasets across all southern NUTS-3 regions. As a result, in NUTS-3 regions lacking regional official statistics data, RS can serve as a reliable alternative source for estimating the extent of key crop areas.



Academic Editors: Kelyn Schenatto, Marlon Rodrigues and Glauco Vieira Miranda

Received: 23 July 2026

Revised: 5 September 2026

Accepted: 10 September 2026

Published: 15 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

Keywords: remote sensing algorithms; earth observations; small farms; crop maps; yield estimation

1. Introduction

Most farms in the world are small. Since they are at the center of sustainable agricultural development, their role is vital. According to recent estimates, over 608 million farms exist worldwide, with a higher proportion of SFs in developing countries. Of those, 70% occupy only 7% of all land used for agriculture, while the largest 1% of all farms hold 70% of all agricultural land [1]. Current knowledge derived from an evidence review and meta-analysis shows that smaller farms often have higher yields and maintain more crop and non-crop biodiversity at the farm and landscape scales than do larger farms (greater than 5 ha). Therefore, the role of SFs in food and food security, although important, remains insufficiently quantified at high spatial resolution.

Although there is no worldwide acceptable definition of a small farm (SF) as the scales vary significantly, in the EU, and according to the European classification system and Eurostat methodologies, SFs or small holdings, are defined as having less than 5 ha or 8 economic size units (ESUs). A distinction between very small farms (VSFs), which are those less than 2 ha; medium farms (MFs), defined as those between 5 and 50 ha; and large farms (LFs), or those greater than 50 ha, may also be useful in certain circumstances, irrespective of the criteria set on their size. Based on the most recent reports from Eurostat (data retrieved in April 2026), the number of agricultural holdings in the European Union (EU) dropped from 9.07 million in 2020 to 8.8 million in 2023, a figure substantially lower than that corresponding to the rest of the developed world. In 2020 (2023), about 67.0% (62.8%) of them were SFs. Large farms (more than 50 hectares) accounted for 7.5% (7.6%) of the total but managed 68.2% (68.3%) of the agricultural land. The majority of them (70.1%) exist in Romania, Italy, Poland, and Spain. They are dispersed across the EU according to different spatial patterns that correspond to the agricultural and territorial characteristics of each region. Nonetheless, the SFs are often linked to peripheral regions of the EU, like all of the Mediterranean countries in southern Europe (S-EU), Scotland, Ireland, and Northern Scandinavia, as well as areas with low agricultural production, such as mountain ranges.

Open issues remain regarding the exploration of indicators and the production of maps, as a step forward in clarifying their importance, their distribution in European regions, and how they are positioned in relation to the agricultural context and the overall territorial characteristics. It is also important to note that the regional official statistics (ROS) of the EU rarely include estimates of SFs' crop yields. A recent study presents a new classification of SFs at NUTS-3. Note that the NUTS (Nomenclature of Territorial Units for Statistics) classification is a hierarchical system subdividing the territory of the EU into regions at three levels (NUTS 1, 2, and 3) and that for several socioeconomic analyses, the smaller territorial units at the NUTS-3 level are known as prefectures, provinces, departments, or regional units. By providing accurate, timely information for the cultivation of small farming plots (SFPlots), remote sensing (RS) can assist food systems in a variety of ways, particularly in the classification of food production. For this, classified satellite images and land cover (LC) maps produced by photo-interpretation are useful resources. In areas where SFPlots predominate, they might distinguish crop varieties for monitoring and measure food output. Even if the resolution is inadequate, it can be enhanced by fusing images to identify ground objects at finer scales. However, since the classification algorithms are not able to map all the present crop varieties into separate classes, the task of cropland mapping is difficult to extend at a global scale [3]. It should be highlighted that SFs' property-level knowledge, which is often obtained from a county's cadastral database, is required to understand how the SFPlots are linked to the SFs and, generally, how agricultural plots are related to farms of any size (Farm-to-Field relationship) [4]. Nevertheless, there are instances where the cadastral data may be insufficient, as the cadastral units do not always reflect functional management units. Additionally, considerations about property fragmentation can be

skewed due to a lack of information about the owners [5]. As a result, RS can provide information on the estimated crop production of SFPlots that is either complementary or at least comparable to the corresponding estimated crop production of SFs presented in rOS at the NUTS-3 level.

This study explores a classification approach for crop mapping over ten of the twenty-five NUTS-3 regional units (RUs) in S-EU, which are situated in southern France, Italy, Greece, southern Portugal, and southern Spain and characterized by a variety of microclimates, distinctive territorial features, and a combination of landscapes [2]. Consequently, the variety of crops cultivated in SFPlots of the aforementioned RUs provides knowledge and useful guidance to other RUs with similar traits. The research was carried out from 2017 to 2020 and examined how RS technology uncovers the role of SFs in regional food systems. It assessed the SFPlots' location, crop diversification, area size, and production yield—metrics that are frequently underreported because rOS rarely keeps track of them. Nowadays, the method allows mapping and monitoring SFPlots accurately, ensuring that the data is comprehensive, consistent, and in line with rOS data. However, very few previous works can be identified in the subject, particularly before the reference period. The study was completely new when it was performed, but now there is a body of knowledge on similar approaches and their added value [6,7]. During the last two decades, significant efforts have been made to map LULC changes at global, regional, and local scales, utilizing cropland information derived from RS data [8–13].

The implemented RS-based approach is assessed using the Sentinel imagery, along with field-collected data and additional auxiliary information [14–16]. The procedure first uses image segmentation to determine and select small plots, which ultimately encompass most of the required SFPlots. The LC crop maps of the RUs under consideration are then created and tested using the random forest (RF) classifier, which also serves as a test bed for other RUs in the EU. We aim to relate crop yields to the spatial distribution of crop types produced in each RU. However, narrow coverage areas and dispersed crop types might enhance misclassification and create noise. In order to mitigate these effects, we limited our attention to the so-called key crop products that were chosen based on factors associated with crops' spatial representativeness in the area.

It should be noted that the key crop production estimates combine the key crop areas with the field-level yields obtained from key informant surveys. Only the highly accurate key crop products are used for the unbiased crop area computation and the crop production estimates, indicating that the *FScore* per-crop-type classification should be more than 75%. Although the above choice can be relaxed in certain cases, it ensures that the key crop production will only be estimated using the best classified crops within each RU. Thus, the study uses estimates gathered by regional experts to calculate the percentage of production within each RU that might be associated with small-scale farming systems. This is because systematic data on the total output of each key crop at the NUTS-3 level (or RU) were missing. Details on the surveys that support the scope of this study can be seen in [17,18].

The primary results of this study are the characterization of the RUs with respect to the distribution of SFPlots, crop types, and production potential. Using Sentinel data, the study also evaluates the effectiveness of utilizing RS-based methodological approaches to monitor SFPlots in other NUTS-3 regions of Europe. Notably, when estimating the area and production of the crop types with low or very low classification accuracy, the RS-based approach minimizes the propagation error. These levels, however, might weaken the strength of the overall conclusions drawn about SFPlots' primary contributions. Thus, evaluating the present and future output of small-scale farming systems at a regional level remains

the major objective and one of the most significant factors to consider in accomplishing the goal of understanding the production patterns of SFPlots and subsequently of SFs.

In the following, the Materials and Methods section (Section 2) is divided into three subsections: In the Study Area section (Section 2.1), we look at the characteristics of the ten RUs which are located in southern Europe (close to the Mediterranean basin) and are referred to as S-RUs for convenience; in the RS-based Overall Approach section (Section 2.2), we present the applied RS-based approach; and lastly, in the Preprocessing section (Section 2.3), we study the details of the image acquisition, the mosaic creation, and the computation of certain secondary indices that were utilized during the analyses and, finally, the building of the agricultural/non-agricultural mask for each of the examined RUs. Section 3 comprises the results on designing the field sample, collecting reference data, determining quality control, segmenting photos, and finally creating the crop map for each S-RU. Finally, the Discussion in Section 4 concludes the analysis and explores the possibilities for further research on aspects of this study and for including the present approach in agrostistics, agricultural policy, etc.

2. Materials and Methods

2.1. Study Area

In the context of food and nutrition security [19,20], a total of twenty-one (21) RUs across eleven (11) European countries and one (1) African country participated in the analysis. These included Bulgaria (Montana), Czech Republic (Jihočeský kraj), France (Vaucluse), Greece (Imathia, Larisa, Ilia), Italy (Lucca, Pisa), Latvia (Latgale, Pieriga), Lithuania (Vilniaus Apskritis), Poland (Rzeszowski, Nowosadecki, Nowotarski), Portugal (Alentejo Central, Oeste), Romania (Bistrița-Năsăud, Giurgiu), Spain (Castellón, Córdoba), and Tunisia (Haouari). As mentioned in Section 1, the RUs were selected using the methodology described in [2]. This research study introduced a novel system for classifying SFs at the NUTS-3 level within the EU, taking into consideration both the agricultural and territorial significance of each RU, as well as a variety of farm sizes.

To provide the analysis and the results obtained by applying the RS-based approach, we revised the RUs presented in [16] and applied it to six RUs selected from Bulgaria (RU01), Italy (RU11 and RU12), and Greece (RU08, RU09, and RU10). We excluded RU01: Montana-BL from the current analysis, as the whole country belongs to Central Europe, and added five additional RUs, including one from southern France and two from each of southern Spain and Portugal. Therefore, the current study area consists of ten (10) administrative (NUTS-3) RUs of the S-EU countries. These are the RU06: Vaucluse-FR of France (FR: 632,702 km²); the RU08: Imathia-GR, RU09: Larisa-GR, and RU10: Ilia-GR of Greece (GR: 131,694 km²); the RU11: Lucca-IT and RU12: Pisa-IT of Italy (IT: 302,073 km²); the RU22: Alentejo Central-PT and RU23: Oeste of Portugal (PT: 92,230 km²); and finally the RU26: Castellon-SP and RU27: Cordoba-SP of Spain (SP: 506,030 km²). The Appendix A contains information regarding the geographical and economic characteristics of the ten S-RUs. As this study aims to provide detailed information on the RS-based approach applied to typical S-RUs, it is essential to thoroughly review and compare the most significant data that are directly relevant. The following part reviews the overall key findings for the aforementioned 20 European RUs.

2.2. RS-Based Overall Approach

In small-scale farming applications such as crop type mapping and crop area estimation, image classification is a powerful RS tool that demonstrates the effectiveness of Sentinel image data. Notably, the accuracy of the produced crop type maps and RF classification performance varied according to the geographic features of each RU. The

classification accuracies across the RUs can be explained in a variety of ways, including the number of satellite images used per region, the different dates of their reception, the spatial and spectral heterogeneity of each agricultural landscape, and the availability and representation of the field dataset (crop-type polygons to train the classification models).

During the field survey campaign conducted in the summer season (June and August 2017), a total of 12,230 crop parcels (points) from all 21 European RUs were considered. Of those points, 10,694 were used for crop identification, with the remainder being extra interesting points obtained during the SFPlots' visits. Studies provide details, focusing on the adopted RS-based approach [15,21].

- The *overall accuracy* (OA) values ranged across 21 RUs from 59.6% to 91.4%, with a *mean OA* ($\overline{OA}$) of 81.6%.
- The *Kappa* index values ranged across RUs from 0.50 to 0.89 with a *mean Kappa* ($\overline{Kappa}$) of 0.74.
- The *mean FScore* ($\overline{FScore}$) was 70.2% obtained for several crop types with *FScore* index values varying significantly among the classes, within the same class, and across the various RUs.
- The analysis of all 21 RUs shows that 14 of those present higher *OA* values than the mean value of 81.6%.
- The regions with the highest *OA* were RU08: Imathia-GR, with *OA*: 91.4%; RU24: Bistrita Nasaud-RO, with *OA*: 91.4%; and RU19: Rzeszowski-PL, with *OA*: 90.0%.
- The lowest *OA* values were observed in RU26: Castellon-SP, with *OA*: 59.6%; RU12: Pisa-IT, with *OA*: 65.7%; and RU14: Latgale-LV, with *OA*: 72.6%.

Note that [17] offers a detailed explanation of the selection criteria for essential agricultural products for each RU. The process involves selecting key products for each RU according to how important they are to SFs in terms of output, consumption, revenue, or cultural value. Table 1 lists the 40 key product cases, which correspond to 34 distinct ones chosen from each of the 10 RUs. Furthermore, [17] has the complete list of the 91 chosen and analyzed key products from the 25 RUs spread across 15 European countries.

Table 1. The ten SE-RUs along with their key products, as well as the number of interviews, surveys, and participants in focus group discussions.

Country	RU Region	Key Informant Interviews	SFs Surveys	Focus Group Size	Products Analyzed
France	RU06 Vaucluse	50	10	-	Wine, Cherry, Olive oil
Greece	RU08 Imathia	11	39	19	Peaches, Cherries, Wine grapes, Beef
	RU09 Larisa	12	38	21	Almond, Sheep & goat milk, Apples, Pulses
	RU10 Ilia	13	42	32	Currants, Oranges, Olive oil, Pickled Vegetables
Italy	RU11 Lucca	6	32	47	Vegetables, Wine, Olive oil, Fruit
	RU12 Pisa	6	24	61	Wine, Vegetables
					Wheat, Bovine meat
Portugal	RU22 Alentejo Central	11	38	24	Vegetables, Wine, Olive oil, Fruit
	RU23 Oeste	5	36	20	Wine, Pears, Eggs, Potatoes
Spain	RU26 Castellon	22	27	17	Citrus, Almond, Pork, Olive oil
	RU27 Cordoba	20	40	18	Olive oil, Wheat, Wine, Milk

Before the key crop area findings were discussed and presented, the coherence of the key crop areas produced from Sentinel-derived data and the key crop areas resulting from regional OS (rOS) was assessed. Using linear regression, estimates of crop areas obtained from rOS and unbiased SFPlot area estimates from Sentinel-based data showed a high correlation ($R^2 = 0.96, p < 0.001$). Of the twenty-one RUs considered, seven, including one from S-EU (RU01: Montana-BL, RU04: Jihoceky Kraj-CZ, RU06: Vaucluse-FR, RU14: Latgale-LV, RU15: Pieriga-LV, RU24: Bistrita-RO, and RU25: Giurgiu-RO), had limited rOS data cultivated by SFs. Additionally, in RU20: Nowosadecki-PL, there were no rOS pertaining to the area covered by one key crop product (apples). Thus, the acquired Sentinel imagery is a useful alternative source of fairly accurate estimates of crop area extent for locations where there was no information (at least until the time of reporting), offering more opportunities to track changes. However, since it is essential to represent the spectral footprint of each key crop type, their effectiveness and utility in creating highly accurate crop maps are rather limited. This is an issue when SFs dominate the agricultural landscape and crop diversification is an important aspect of the spatial pattern.

Regarding the importance of SFs, this research offers an integrated approach to assess how SFPlots contribute to regional food production in ten S-RUs that are typical of the Mediterranean diversity. The methodology combines information on agricultural crop types and field sizes estimated from RS imagery with self-reported yield data from a field survey in a novel way. Using the spatial data from the Sentinel imagery and image segmentation procedures, it successfully creates an accurate key crop type map for each RU. One way to achieve this is to find and utilize a possible field-to-farm size relationship. In this situation, it could be possible to select most agricultural plots and utilize them in place of farms. Note that by evaluating field sizes, we can improve our understanding of geographical patterns of farm sizes based on empirically investigating the field-size-to-farm size relationship. However, according to recent studies, neither the field size nor the farm size can be accurately evaluated [4]. This is because traditional sources of farm size data are agricultural censuses and household surveys, although they are normally only accessible at aggregate spatial scales and at rate time intervals. Furthermore, spatially explicit and comprehensive information about specific agricultural fields can be obtained from various sources, including cadastral information systems, agricultural subsidy applications in certain regions, and earth observation data [13,21]. The case is more complicated in the situation of SFs, and therefore, the characterization of the field-to-farm size relationship is not currently possible due to our inadequate empirical evidence.

2.3. Preprocessing

The data set for the ten S-RUs includes 58 cloud-free composite images from Sentinel-2A featuring 13 spectral bands and spatial resolutions between 10 m and 60 m. The number of utilized scenes for each RU is as follows:

- RU06 Vaucluse-FR: 1 spring and 3 summer images.
- RU08 Imathia-GR: 4 spring and 4 summer images.
- RU09 Larisa-GR: 4 spring and 4 summer images.
- RU10 Ilia-GR: 2 spring and 2 summer images.
- RU11 Lucca-IT: 4 summer images.
- RU12 Pisa-IT: 4 summer images.
- RU22 Alentejo Central-PT: 1 spring and 4 summer.
- RU23 Oeste-PT: 2 spring and 8 summer.
- RU26 Castellon-SP: 2 spring and 3 summer.
- RU27 Cordoba-SP: 2 spring and 4 summer.

The composites were created using individual Sentinel-2A scenes collected between April and September 2017 to establish a multi-temporal classification scheme. In all instances, cloud coverage was below 10%; however, in some cases, such as those involving the Greek RUs, cloud coverage was less than 5%.

The images were downloaded from ESA's Sentinel SciHub. The analysis utilized ten bands with the following spatial resolution:

- 10 m resolution: B2 (Blue: 490 nm), B3 (Green: 560 nm), B4 (Red: 665 nm), and B8 (NIR-1: 842 nm),
- 20 m spatial resolution: B5 (Red edge-1: 705 nm), B6 (Red edge-2: 740 nm), B7 (Red edge-3: 783 nm), B8a (NIR-2: 865 nm), B11 (SWIR-1: 1610 nm), and B12 (SWIR-2: 2190 nm).

Before the analysis, the Sentinel-2 images underwent atmospheric correction using the Dark Object Subtraction (DOS)-1 method. Subsequently, mosaic images were created by merging and clipping the images to fit within the bounds of each RU. It is noteworthy that each mosaic image corresponds to a specific date in each region.

Four vegetation indices were calculated in the next preprocessing step and utilized as auxiliary variables in the classification. Their selection was motivated by their complementary sensitivity to vegetation greenness, canopy density, senescence, and moisture-related spectral properties. Together, these indices provide a more comprehensive spectral characterization of small agricultural parcels, improving their discrimination from surrounding vegetation, bare soil, and other land-cover types. These indices are as follows:

- The Normalized Difference Vegetation Index (NDVI).
- The Enhanced Vegetation Index (EVI).
- The Plant Senescence Reflectance Index (PSRI).
- The Short-wave Infrared Reflectance 3/2 Ratio (SWIR32) vegetation index.

The NDVIs were computed and stacked to obtain one NDVI image per date and RU. The aim is to increase the spectral separability of the different classes of land cover. This is predominantly in agricultural landscapes where spatial and spectral heterogeneity is prevalent and where there is a high degree of crop diversity. Besides, the mean, variance, texture mean, and homogeneity were calculated using a 3×3 window size gray-level co-occurrence matrix (GLCM) feature. These indices can be used as auxiliary variables in the classification process. Ultimately, a raster layer was created by stacking all the clipped bands, vegetation indices, and the GLCM features.

By removing the non-agricultural areas from the estimated SFPlots distribution process, the final step of the preprocessing phase creates a mask for every RU. Therefore, in addition to the spatial data, the following geo-processing steps are used to build the mask:

- The collection of the CORINE LC maps for each RU.
- The creation of around 1000 randomly stratified sample points, equally split between agriculture and non-agriculture fields over each RU. This procedure ensures that the 1000 points will cover the diversity of the main LC types existing in each RU.
- The codification of each point as *agri* for agriculture and *n_agri* for non-agriculture by visual interpretation of high-resolution Google Earth imagery and Sentinel-2A true color composition.
- The split of the main dataset into training (80%) and test (20%) datasets using the *CreateDataPartition* function from the *caret* R package [22].
- The implementation of an R script for the image classification process.
- The accuracy assessment, and prediction, using random forest (RF) as a classifier.
- The classification based on the RF ML algorithm.
- The accuracy assessment based on the test dataset (OA, *Kappa* indices), and finally,

- The prediction to create the agriculture/non-agriculture mask using the best RF model.

Figure 1 shows the OA of all 20 RUs, considering that 7 of those were derived using Sentinel-1 images and the other 14 were derived using Sentinel-2A images. In the above, RU20 and RU21 of Poland use the same Sentinel images, whereas RU28 does not belong to the EU. Also, Figures 2 and 3 show the steps we have to follow in order to create an RU mask, taking as an example RU12: Pisa-IT. Further, Figures 4–7 show the agriculture/non-agriculture masks of the RU08: Imathia-GR, RU09: Larisa-GR, RU10: Ilia-GR, and RU11: Lucca-IT, respectively.

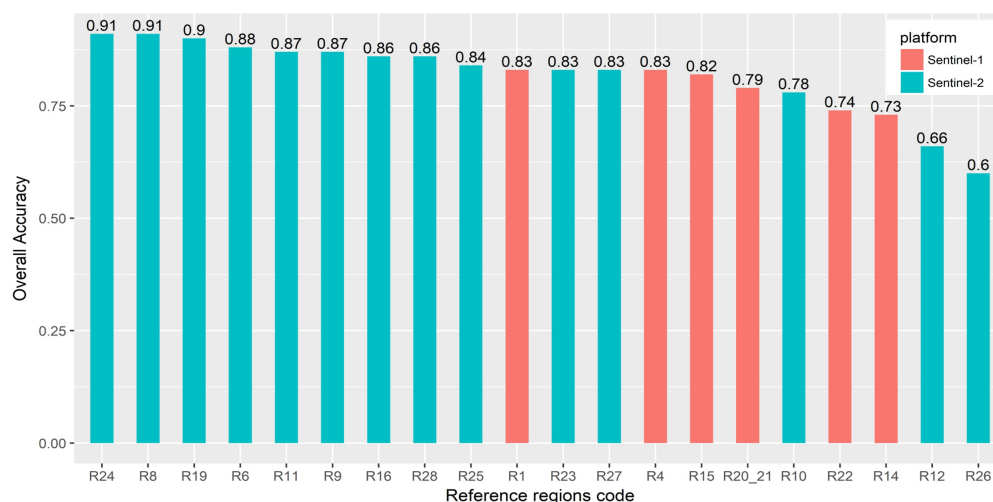


Figure 1. OA values deduced from the S1 and S2 imagery of the 20 RUs participating in the analysis.

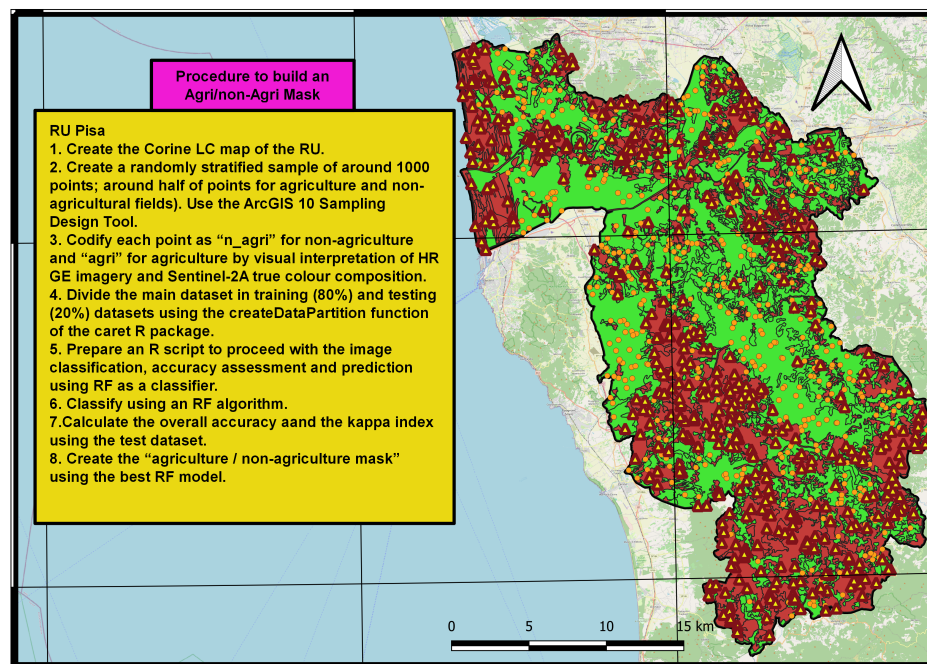


Figure 2. Steps to create an agri/non-agri mask for the RU11: Pisa(IT). Agri-areas are indicated with green color, non-agri areas with brown and sample points with yellow.

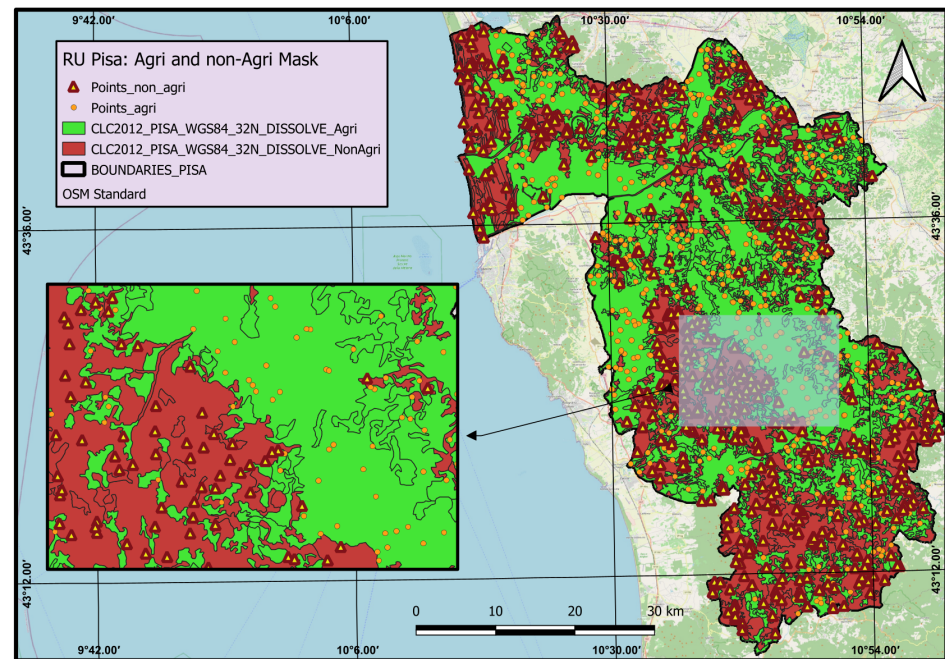


Figure 3. Agri/non-agri mask for the RU11: Pisa(IT). Agri-areas are indicated with green color, non-agri areas with brown and sample points with yellow.

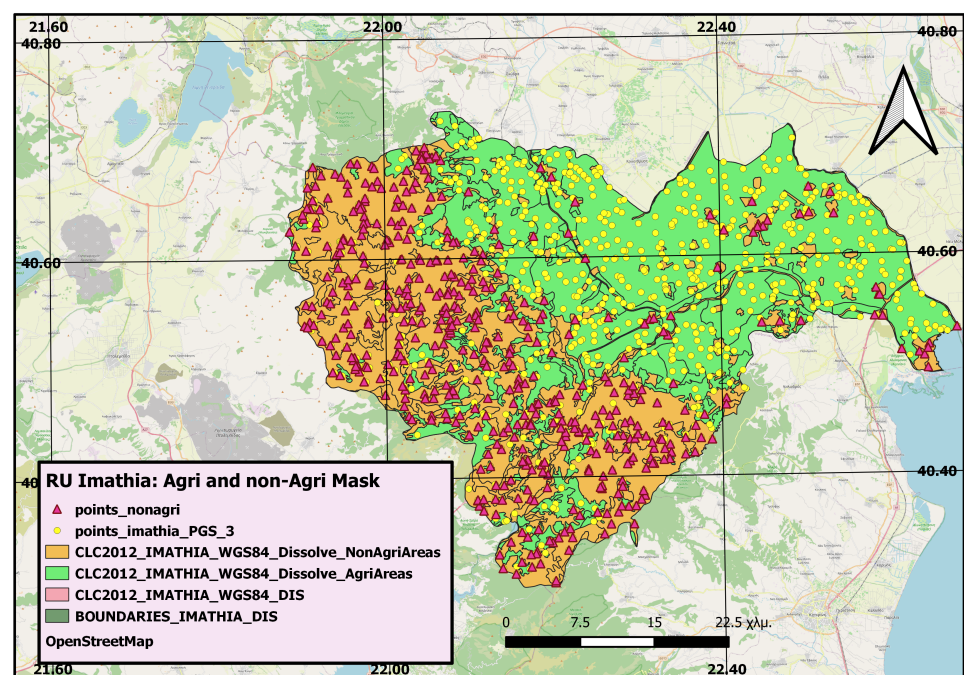


Figure 4. Agri/non-agri mask for the RU08: Imathia(GR). Agri-areas are indicated with green color, non-agri areas with light brown and sample points with yellow.

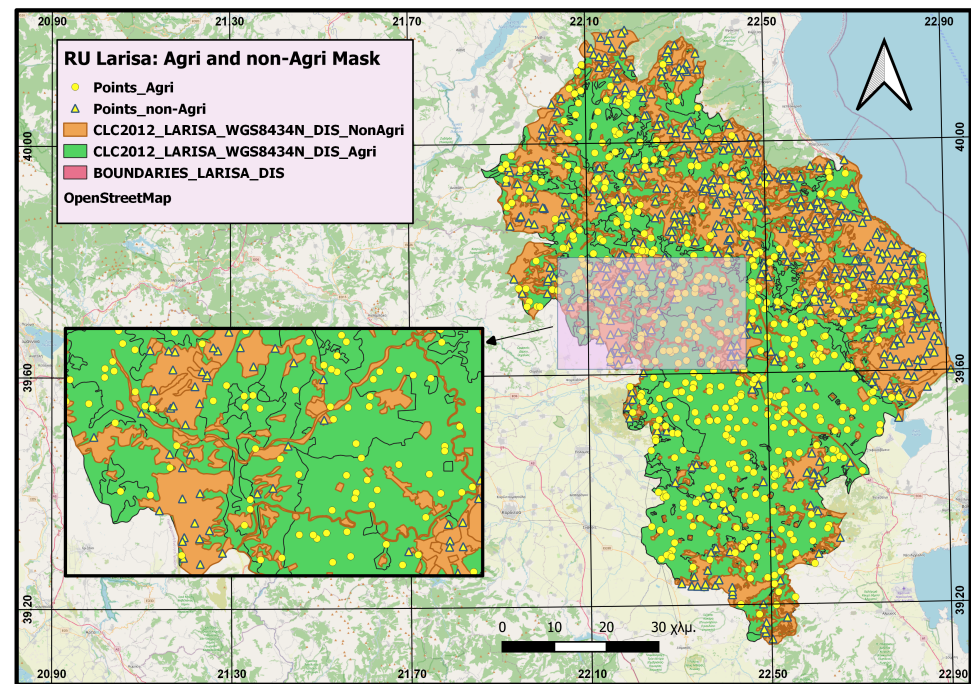


Figure 5. Agri/non-agri mask for the RU09: Larisa(GR). Agri-areas are indicated with green color, non-agri areas with light brown and sample points with yellow.

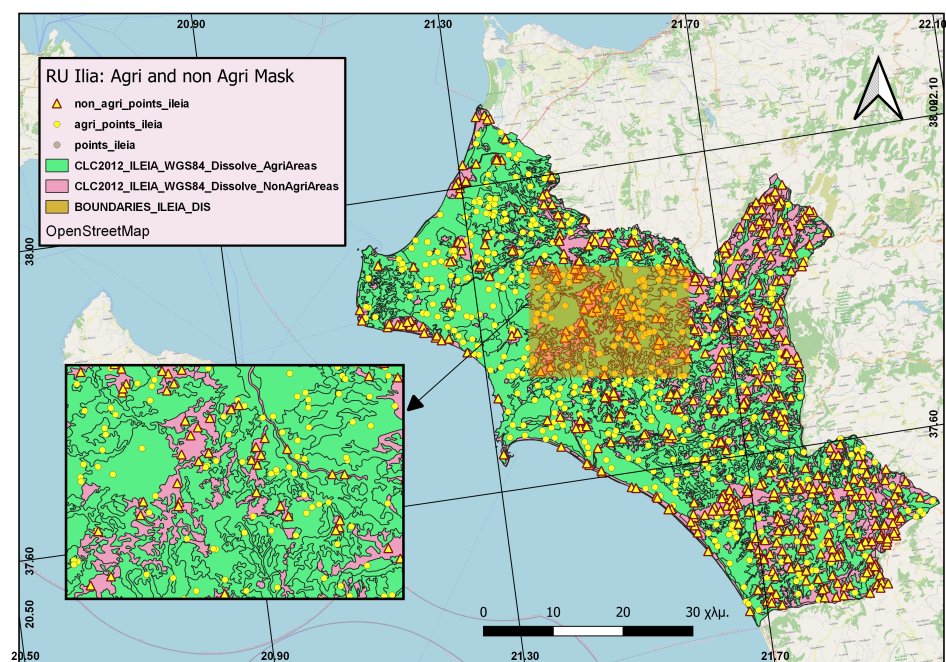


Figure 6. Agri/non-agri mask for the RU10: Ilia(GR). Agri-areas are indicated with green color, non-agri areas with light brown and sample points with yellow.

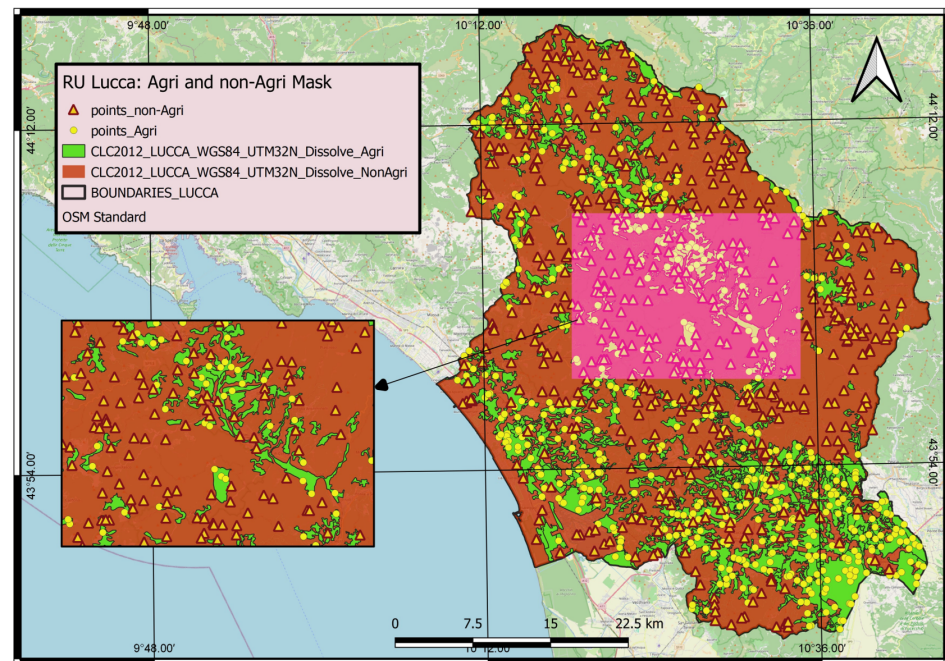


Figure 7. Agri/non-agri mask for the RU11: Lucca(IT). Agri-areas are indicated with green color, non-agri areas with brown and sample points with yellow.

Crop Types Maps Production

The production of a crop type map per RU requires further analysis, which is performed through the following steps:

- *Field sample design and collection of data:* In order to perform dataset calibration and validation in the image classification process, fieldwork was carried out to identify crop types and enable the connection of satellite image data with the actual features and materials. Also, to reduce the high cost of field operations, a specialized sampling methodology was applied that combined the diversity of the agricultural terrain with a few generally accepted accessibility criteria.

In particular, in the first step, twenty-five (25) square blocks of land that were defined in a $2 \times 2 \text{ km}^2$ (400 ha) grid were utilized to collect data for each RU. Then, about 20 sample points for each square block, or over 500 in total per RU, were selected for examination on a stratified basis. Their distribution in each selected square block was based on the percentage of the area that each cluster covered. Note that the selection criteria included the densest road network; the minimum distance of 3.0 km between the blocks; the highest threshold of the categorization; and the elimination of those that intersected the corresponding region borders. Since the Shannon Equitability (or Evenness) Index:

$$J = H/H_{max} = H/\ln(n) = - \sum_{i=1}^n p_i \ln(p_i) / \ln(n)$$

measures how evenly cluster types are distributed; it is used to evaluate the overall diversity of agricultural land cover along the grid in each RU, with values of J ranging from zero (uneven/dominated) to one (perfectly even). In the above, H represents the Shannon Diversity Index; H_{max} is the maximum possible diversity value, which equals the natural log of the total number of classes n , indicating also the class richness, namely, the number of classes determined by the unsupervised classifications of spectral data from Sentinel 2A images, and p_i is the proportional abundance of the i^{th} cluster type. Note that the higher (lower) the value of H , the higher (lower) the crop

diversity along the grid in each RU, and in fact, J scales H relative to the maximum possible diversity if all classes were perfectly equal.

From the aforementioned analysis and to facilitate the fieldwork, the square blocks with the highest agricultural diversity were chosen, namely, those with values in the upper quartile of the index J , while having a distance to roads that was less than the mean distance to roads. In addition, the square blocks were separated by 3 km, indicating that adjacent square blocks were not selected. Notably, the Euclidean distance from the relevant linear infrastructures is calculated using OpenStreetMap's road network database, allowing easy access to the destination square blocks.

During the fieldwork, each field point's crop was reported and captured on camera. Moreover, each point's thematic and geographic accuracy was confirmed by superimposing the points over Sentinel-2A and high-resolution Google Earth images, in addition to visually inspecting the field digital photos that the field teams had taken. This process was vital to the excellent quality of the data: certain points were removed because the crop identification was ambiguous or because the plot was in an unreachable place, as was the case with the GPS's inaccurate geographical coordinates. Nevertheless, the final dataset excluded all the in situ observations grouped from fieldwork as tillage lands, ploughed areas, and unidentified crop types.

Finally, an unsupervised classification was performed to ascertain the optimal number of classes that the model produced. Figures 8–11 display the maps of the Greek and Italian RUs considered in this work, including the details of the chosen square blocks and the spatial distribution of field points used to identify crop types. The optimal number of classes that the model produced was then ascertained by doing an unsupervised classification.

- *Quality checking:* To produce a consistent output across the RUs under consideration, the quality control validation procedure utilized for all in situ data collected by the relevant surveyors' teams has to be adhered to. As a result, the quality control procedure was used as a validation method for all the field data gathered for each RU, correcting any thematic and spatial propagation errors across the entire methodological cycle used to build each RU crop type map. In 16 of the 21 RUs, an average of 8.5% (std = 2.75%) of the sample points (min = 5.1% and max = 15.8%) were examined for quality control. It should be noted that RU16: Vilnius Apskritis-LT, RU22: Alentejo Central-PT, RU23: Oeste-PT, RU27: Cordoba-SP, and RU28: Haouaria-TN were not included in the analysis due to incomplete data. Separate surveyor teams from Portugal and Greece carried out the quality control work and the analysis of the remaining 16 of the 21 RUs. A surveyor team visited each sample point to identify and gather information about the type of crop being cultivated. To ensure consistency, the same procedure was used to gather the necessary crop data at the selected locations. The quality control surveyors utilized a random sample of an average of 10.0% (50 out of the previous 500 points chosen per RU) of the data to re-examine the in situ data in all RUs. However, the crop type of the rechecked data was verified and registered, and the crop was photographed, by different surveyors from those who carried out the initial in situ data collection. In addition to the spatial distribution of the square blocks and the field points, Figures 8–11 show the ground truth points (highlighted with a different color than that of the field points) in each square block that the Greek team visited. This implies that each square block would receive no more than two or three revisits.

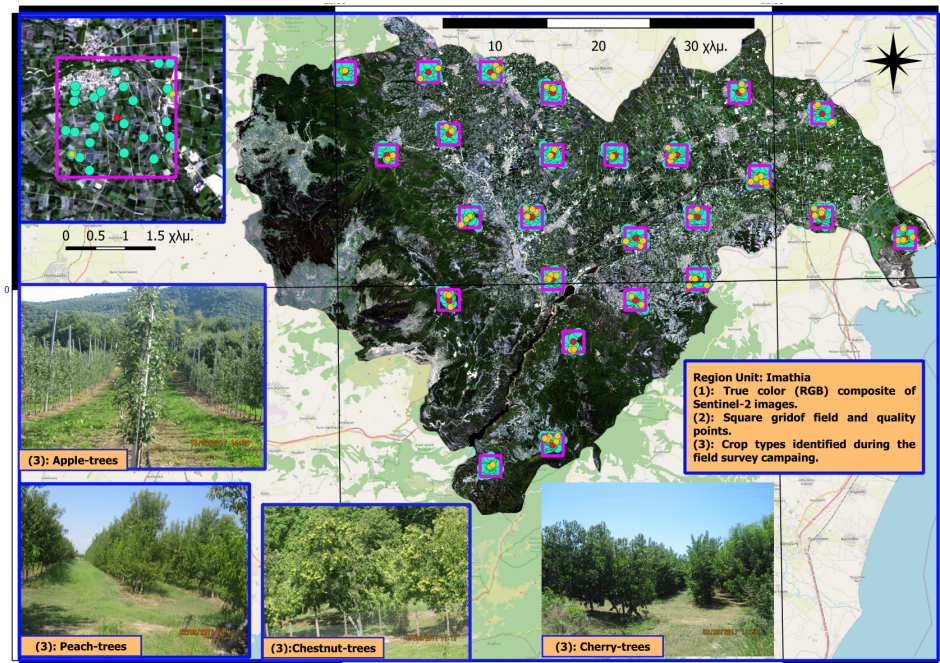


Figure 8. The map of RU08: Imathia(GR) showing the spatial distribution of field points (highlighted in light blue) used for crop type identification. Ground truth points (highlighted in yellow) are also indicated.

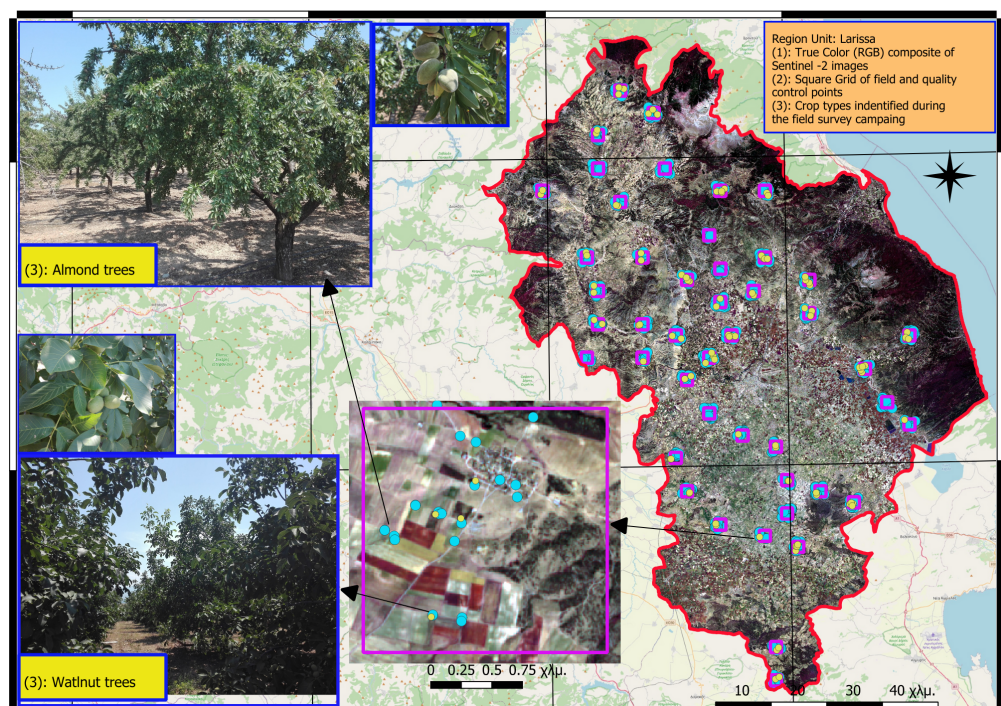


Figure 9. The map of RU09: Larisa(GR) showing the spatial distribution of field points (highlighted in light blue) used for crop type identification. Ground truth points (highlighted in yellow) are also indicated.

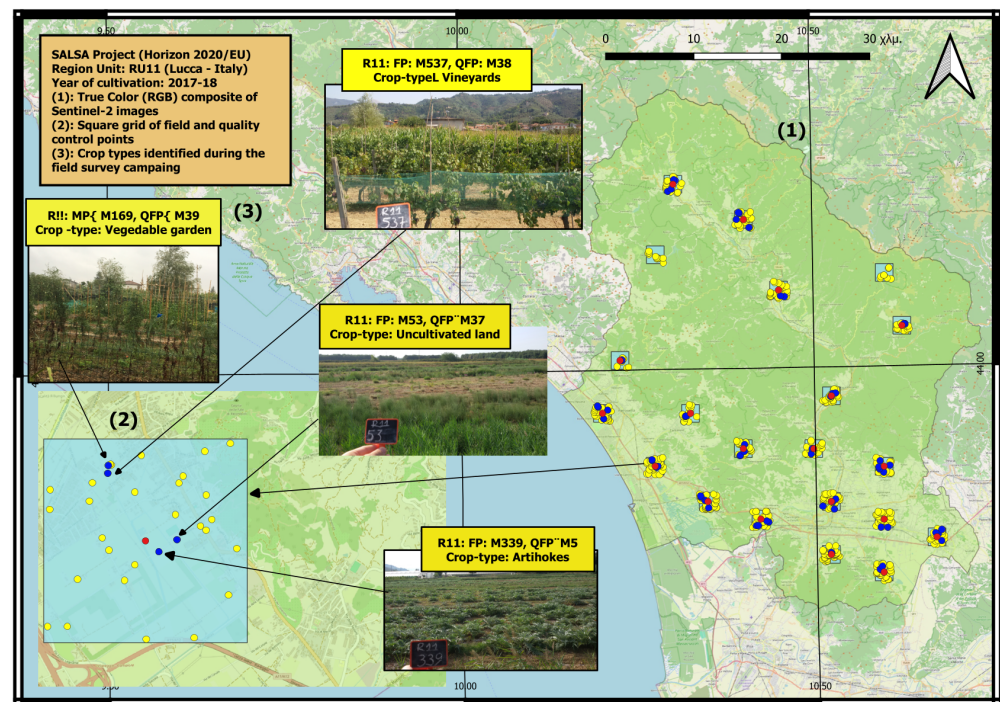


Figure 10. The map of RU11: Lucca (IT) showing the spatial distribution of field points (highlighted in yellow) used for crop type identification. Ground truth points (highlighted in blue) are also indicated.

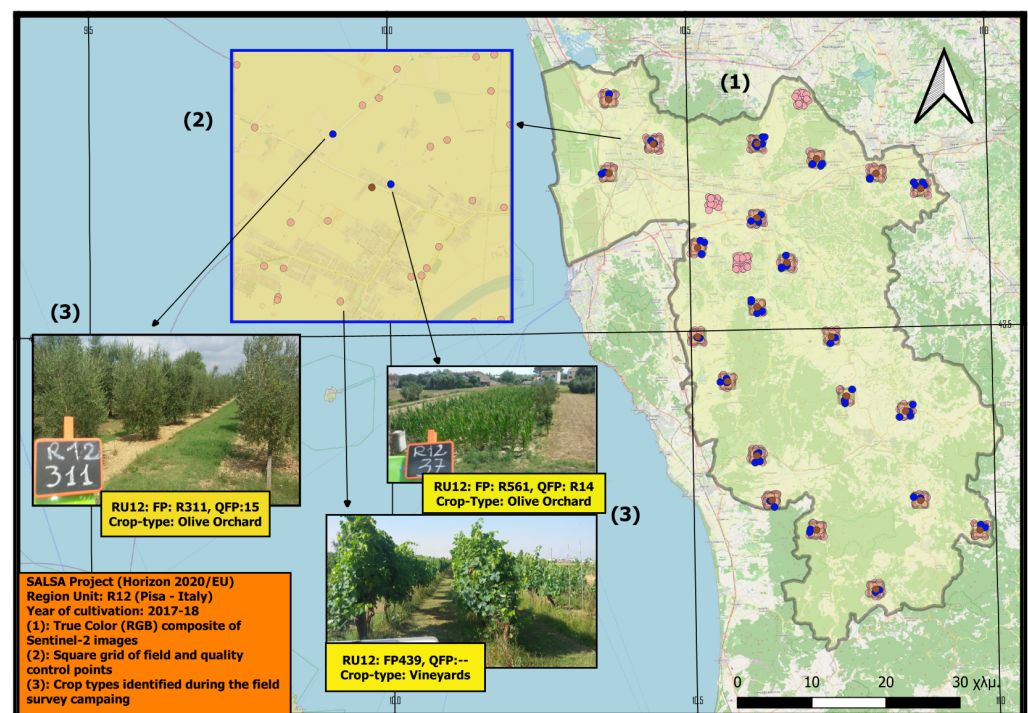


Figure 11. The map of RU12: Pisa (IT) showing the spatial distribution of field points (highlighted in yellow) used for crop type identification. Ground truth points (highlighted in blue) are also indicated.

- *Image segmentation:* Using the Normalized Difference Vegetation Index (NDVI) computed for each accessible Sentinel-2 date, we constructed a layer stack to handle image segmentation. All 21 RUs were required to employ the segmentation process. However, the segmentation was carried out using the Sentinel-2 images that were available in 2016 for the ten RUs that did not have cloud-free images during the period of analysis. To handle satellite images and geographical and spatial informa-

tion simultaneously, the edge length was calculated using the Canny edge detector (CED) algorithm on the GEE cloud computing platform. Moreover, significant proxy information on the borders of FPlots can be obtained by segmenting an image into polygons. The polygon (segment) size will be sufficiently small if the images are over-segmented, which will result in a very large number of false SFPlots. On the contrary, if the images are under-segmented, the size of the polygon will be very large, mixing several SFPlots into one large polygon. The selection of SFPlots from the segmentation output was made simpler by this insight. Two instances from RU-Ilia, Greece, are shown in Figure 12. Large SFPlots predominate in the first (upper left square, 2×2 km), while SFPlots predominate in the second (upper right square). As can be seen, there is a significant variation in the overall edge length between the SFPlots and LFPlots contexts due to the significantly higher configuration variability.



Figure 12. Paradigms illustrating how the CED algorithm calculates the variations in total edge lengths between small and large farm plot sizes.

- Crop dataset creation** The borders of the segments that had been created by the image segmentation procedure were joined together with all the dispersed observed points so as to move forward with the image classification. This allowed for the identification of segments that contain the observed locations (see Figure 13). These segments were depicted as polygons determined by the boundaries of the parcels containing the crop plots identified during the fieldwork. This implies that each point could be linked to a spatial unit that corresponds to a crop plot. Moreover, each reference region's output, income, consumption, and cultural significance were taken into consideration while choosing a set of regional key crop products. This was part of the methodological design to understand and assess the regional food systems and the contribution of SFs to food and nutrition security (FNS). In Table 1, we showed the list of the 40 key products selected in each of the 10 RUs taken into consideration. The approved criteria might have made it simpler to collect sufficient field data because they were connected to the geographical representation of the crops in the area. These criteria reduce classification errors by excluding crop types that cover a residual percentage of the covered area from the analysis.

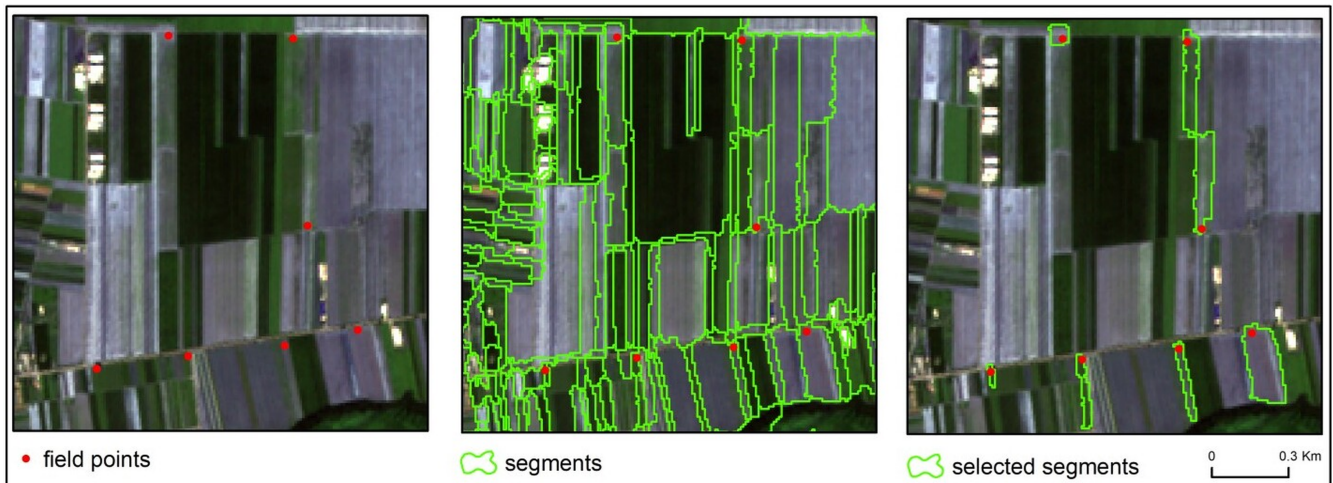


Figure 13. Intersection of field points with segments (agricultural plots) generated during the image segmentation stage

- *Image classification:* The RF algorithm is a well-known, effective RS tool introduced in [23] and used in farming systems for predicting models of various crop types (classes). According to [7], it functions as a pixel-based ensemble supervised ML classifier, which was employed to produce the crop-type maps. A unified interface for training, fine-tuning, and assessing ML models, such as RF classifiers and regressors, is provided by the Classification And REgression Training package in R (*caret R*) [22]. By utilizing bagging, or the bootstrap aggregation method, to select a subset of the training data, it builds several decision trees for prediction. This approach enhances predictive performance, showing comparable accuracies with other ML algorithms demonstrated in various crop mapping studies [24–26].

To proceed with the RF classification process, the crop segments dataset of the ten RUs under consideration was split into training (75%) and test (25%) subsets (see Table 2), using the *CreateDataPartition* function from the *caret R* package [22]. This maintained the overall class distribution across the dataset, produced balanced splits in the data, and guaranteed random sampling within each class. The R package *RandomForest* was utilized to carry out the RF classification by adjusting its two primary parameters. The first, *mtry*, controlled the number of predictor variables randomly sampled to determine each split [7], and it was set to $\sqrt{p}$, with p the number of predictor variables. The second parameter, *ntree*, determined the total number of independent trees to grow. To train the RF models, several training subsets were tested using a randomly selected number of trees, with the best one selected from 1000 trees [27].

- *Accuracy assessment* The test subset was utilized to measure accuracy by calculating the confusion matrix, namely, a cross-tabulation of the crop-mapped data against the corresponding reference crop data. The pixel-based supervised RF classification performed on the selected Sentinel data indicated that the accuracy of crop type maps varied across 19 rather than 20 RUs of the EU, with a mean OA value of 81.41% and an average *Kappa* value of 0.736. This is due to Poland utilizing the same Sentinel images for both of its RUs, and the RU28-Haouari of Tunisia being a non-EU country and not participating in this analysis. Additionally, the mean values of OA and *Kappa* obtained from all the selected Sentinel data of the ten S-RUs were 79.70% and 0.726, respectively. The results demonstrate that the corresponding mean values of OA and *Kappa* obtained from all the selected Sentinel data of the remaining nine(ten) non-southern RUs (non-S-RUs) of the EU varied as 83.31% and 0.747, respectively. Note that the OA of the nine(ten) non-S-RUs performs better than that of the ten

S-RUs; however, RU08: Imathia-GR and RU24: Bistrița-Năsăud-RO outperform all the considered RUs of the EU. Figure 14 demonstrated the various possibilities.

Apart from the *OA*, the *User accuracy (UA)* and the *Producer accuracy (PA)* were also computed from the confusion matrix. The *FScore* metric, which ranges from 0 to 1, expresses the balance between the *Commission Error (CE)*, i.e., $UA = 100\% - CE$, and the *Omission Error (OE)*, which assesses the test accuracy obtained by their harmonic mean; the higher the *FScore*, the more accurate the classification.

Table 2. The number of segments (SFPlots) of the ten S-RUs that underwent the quality control check and were used in the classification model, along with the *OA* and *Kappa* values of the crop types obtained by utilizing the Sentinel imagery.

Reference RU	Segments for the Classification			OA (%)	Kappa
	Total	Training	Validation		
RU06: Vaucluse-FR	561	419	142	88.3	0.850
RU08: Imathia-GR	500	371	129	91.4	0.830
RU09: Larisa-GR	514	373	141	86.8	0.770
RU10: Ilia-GR	486	368	118	78.3	0.730
RU11: Lucca-IT	372	272	100	87.4	0.830
RU12: Pisa-IT	510	356	154	65.7	0.550
RU22: Alentejo Central -PT	612	460	152	73.8	0.605
RU23: Oeste-PT	651	489	162	82.5	0.770
RU26: Castellon-SP	625	462	163	59.6	0.500
RU27: Cordoba	398	291	107	83.2	0.780
Mean Values				79.7	0.726

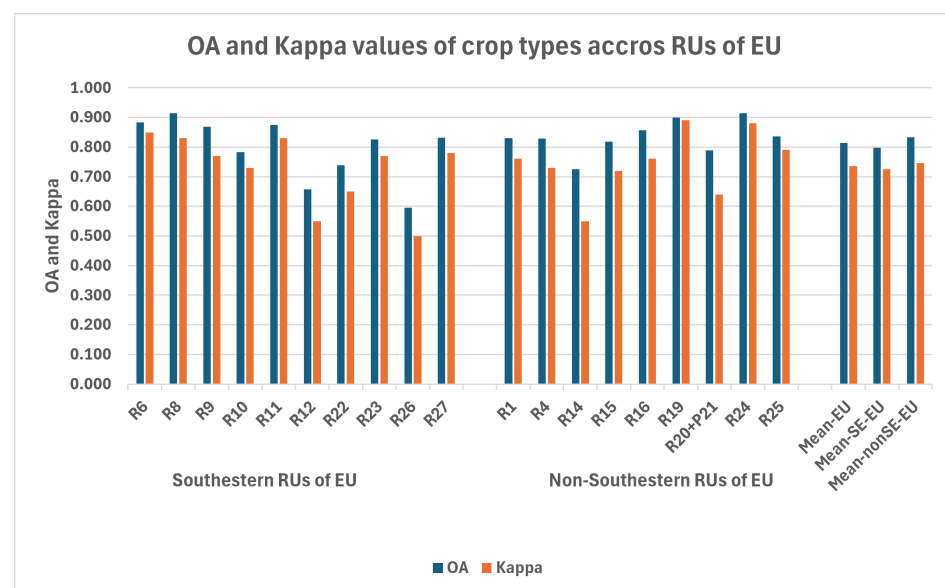


Figure 14. The *OA* and *Kappa* values of crop types across the considered S-RUs and non-S-RUs in the EU.

3. Results

The main purpose of this study was to establish a framework designed for an RS-based approach to obtain accurate area estimates of key crops in SFPlots in the examined RUs (NUTS-3 level) across a variety of European countries. We focused on ten southern RUs located in five Mediterranean EU countries based on the SF typology provided in [2]. The assessment was carried out by utilizing Sentinel image-derived RS products, including crop type maps with ground-truth validation. The approach presented the corresponding SFPlots by mapping the chosen key crop products [17] and applying image segmentation.

The study only looked at crops that matched specific basic requirements for spatial representation in each RU, which helped to reduce noise and misclassification produced by dispersed small crop patches.

Table 3 shows the levels of *PA* and *UA*, the *OE* and *CE*, as well as the *FScore* per crop type per S-RU derived from the corresponding confusion matrices of classification. Also, it presents the number of crop types per region, along with their mean *FScore* values and their corresponding standard deviations. Although the number of crop types may vary, the *FScore* assessment metric, which is the harmonic mean of *PA* and *UA*, can cover both performance characteristics in a simple calculation. To ensure that key crop production estimates are derived only for the best-classified crops per RU, the unbiased key crop area computation and the key crop production estimations are performed only for the best-classified crops, which surpass *FScore* values of 75%. This threshold criterion is aligned with the third quantile value of 74.80% for the *FScore* values of key crop types in all 21 RUs selected, as well as with the corresponding value of 73.04% of the ten S-RUs in the southern EU countries. However, the preceding constraint can be relaxed, allowing for the estimation of at least one key crop product per RU to be included in the analysis. That occurred in RU-Larisa-GR, for which the *FScore* for “vegetables (legumes)” was 73.2%, slightly lower than 75% but higher than the third quantile value indicated above. Table 3 indicates that the *FScore* value of 73.02% for “vegetables (legumes)” exceeds the *FScore* values of the remaining key crop products “other orchards (apples)” and “nuts (almonds)”.

Table 3. The levels of *Producer* and *User Accuracy* (*PA*, *UA*), as well as the *Omission*, the *Commission Errors* (*OE*, *CE*), and the *FScore* are presented for each crop type, within each S-RU derived from the respective confusion matrices. Additionally, the Table includes the number of crop types per S-RU, along with their mean and the standard deviation values for the *FScore*, the *PA*, and the *UA* measures. Note that: (1) “*”: Indicates that the classified crop type utilizes a more generalized crop cover. (2) We use italics to indicate the key crop types in the respective RU.

RU Code	Crop Type	Crop Area	PA (%)	OE (%)	UA (%)	CE (%)	FScore (%)
RU06: Vaucluse	Aromatic plants		91.7	8.3	96.5	3.5	94.0
	Cereals		94.2	5.8	91.4	8.6	92.8
	Maize		80.7	19.3	57.4	42.6	67.1
	Meadows, pastures and forage crops		85.8	14.2	88.3	11.7	87.1
	<i>Olive groves</i>		0.7	99.3	1.2	98.8	0.8
	Orchard		79.3	20.7	96.1	3.9	86.9
	Sunflower		79.0	21.0	97.7	2.3	87.3
	Vegetables		46.5	53.5	68.3	31.7	55.3
	<i>Vineyards</i>	10,578.4	94.8	5.2	89.5	10.5	92.1
No. of crop types: 9							
Mean values:			72.52		76.27		73.71
S.D. values:			30.62		31.36		30.31
RU08: Imathia	Cereals		82.8	17.2	72.7	27.3	77.4
	<i>Cherries *</i>		7.6	92.4	21.4	78.6	11.2
	Cotton plants		96.1	3.9	89.0	11.0	92.4
	Maize		93.8	6.2	95.8	4.2	94.8
	Meadows, pastures and forage crops		85.3	14.7	91.1	8.9	88.1
	Orchards		99.5	0.5	95.5	4.5	97.4
	Other crops		24.7	75.3	73.7	26.3	37.0
	<i>Peaches *</i>	8782.06	84.1	15.9	77.0	23.0	80.4
	Rice		93.5	6.5	94.3	5.7	93.9
	Vegetables		57.4	42.6	79.5	20.5	66.7
	<i>Vineyards/Grapes</i>		62.2	37.8	88.1	11.9	72.9

Table 3. Cont.

RU Code	Crop Type	Crop Area	PA (%)	OE (%)	UA (%)	CE (%)	FScore (%)
	total no: 11						
	Mean values:		71.55		79.83		73.84
	S.D. values:		30.67		21.21		27.08
RU09: Larisa	Cereals		99.3	0.7	90.0	10.0	94.4
	Cotton plants		78.2	21.8	92.0	8.0	84.5
	Maize		80.0	20.0	94.2	5.8	86.5
	Meadows, pastures and forage crops		47.6	52.4	65.0	35.0	55.0
	Olive groves		18.3	81.7	32.8	67.2	23.5
	Other orchards		75.7	24.3	64.3	35.7	69.5
	Vegetables/Pulses	814.552	58.2	41.8	98.7	1.3	73.2
	Vineyards		75.8	24.2	51.4	48.6	61.3
	No. of crop types: 9						
	Mean values:		64.04		71.56		66.28
	S.D. values:		24.49		23.09		22.04
RU10: Ilia	Cereals		70.8	29.2	89.2	10.8	78.9
	Meadows, pastures and forage crops		82.7	17.3	78.1	21.9	80.4
	Olive groves	20,618.20	88.5	11.5	82.8	17.2	85.5
	Oranges		54.7	45.3	35.4	64.6	43.0
	Other orchards		28.6	71.4	52.6	47.4	37.0
	Vegetables		79.9	20.1	74.1	25.9	76.9
	No. of crop types: 8						
	Mean values:		70.06		74.06		71.04
	S.D. values:		19.59		20.66		18.68
RU11: Lucca	Maize		94.7	5.3	98.7	1.3	96.7
	Meadows, pastures and forage crops		81.2	18.8	95.0	5.0	87.6
	Olive groves	2180.83	93.3	6.7	81.9	18.1	87.2
	Other crops		68.3	31.7	91.5	8.5	78.2
	Vineyards	789.903	89.8	10.2	75.3	24.7	81.9
	Vegetables		79.9	20.1	74.1	25.9	76.9
	No. of crop types: 5						
	Mean values:		70.06		74.06		71.04
	S.D. values:		19.59		20.66		18.68
RU12: Pisa	Cereals	1304.62	69.6	30.4	93.3	6.7	79.7
	Meadows, pastures and forage crops		67.4	32.6	65.6	34.4	66.5
	Olive groves		64.3	35.7	58.3	41.7	61.1
	Other crops		68.4	31.6	71.3	28.7	69.8
	Vineyards		63.0	37.0	64.6	35.4	63.8
	Watermelons	2289.35	74.6	25.4	80.3	19.7	77.3
	No. of crop types: 5						
	Mean values:		85.46		88.48		86.32
	S.D. values:		10.93		9.66		6.99
RU22: Alentejo Central	Cereals		77.4	22.6	67.2	32.8	71.9
	Meadows, pastures and forage crops		46.3	53.7	51.3	48.7	48.7
	Olive groves		72.8	27.2	65.6	34.4	69.0
	Orchard		60.0	40.0	69.3	30.7	64.3
	Vegetables		57.7	42.3	63.4	36.6	60.4
	Vineyards	1867.13	82.5	17.5	93.0	7.0	87.5
	Vegetables/Grapes		74.6	25.4	80.3	19.7	77.3

Table 3. Cont.

RU Code	Crop Type	Crop Area	PA (%)	OE (%)	UA (%)	CE (%)	FScore (%)
No. of crop types: 6							
Mean values:			66.12		68.30		66.97
S.D. values:			13.72		13.66		12.92
RU23: Oeste	Cereals		93.8	6.2	88.3	11.7	90.9
	Maize		9.8	90.2	20.7	9.3	13.3
	Meadows, pastures and forage crops		61.5	38.5	72.5	27.5	66.6
	Orchard		89.9	10.1	87.3	12.7	88.6
	Potatoes		88.5	11.5	18.7	81.3	30.9
	Pumpkins		97.7	2.3	63.6	36.4	77.0
	Vegetables		62.9	37.1	92.6	7.4	74.9
	Vineyards	3627.34	76.1	23.9	92.2	7.8	83.4
	Pears *	2407.64	88.0	12.0	94.0	6.0	90.9
No. of crop types: 9							
Mean values:			74.20		69.99		68.50
S.D. values:			27.43		30.26		27.85
RU26: Castellon	Almonds		71.3	28.7	53.8	46.2	61.3
	Bare soil		34.6	65.4	47.1	52.9	39.9
	Cereals		55.4	44.6	72.7	27.3	62.9
	Citrus1	17,016.44	85.7	14.3	90.6	9.4	88.1
	Citrus2		46.4	53.6	70.5	29.5	56.0
	Meadows, pastures and forage crops		59.9	40.1	59.8	40.2	59.8
	Olive groves		51.2	48.8	37.7	62.3	43.4
	Other crops		25.6	74.4	51.7	48.3	34.2
No. of crop types: 8							
Mean values:			54.90		60.49		55.70
S.D. values:			19.23		16.81		16.96
RU27: Cordoba	Cereals	2742.21	91.9	8.1	84.2	15.8	87.9
	Meadows, pastures and forage crops		70.6	29.4	77.7	22.3	74.0
	Olive groves		87.1	12.9	87.4	12.6	87.2
	Orchard		63.1	36.9	89.2	10.8	73.9
	Vegetables		9.6	90.4	29.7	70.3	14.5
	Vineyards	1857.41	92.4	7.6	79.5	20.5	85.5
	Wheat *		70.6	29.4	53.7	46.3	61.0
No. of crop types: 7							
Mean values:			71.63		69.33		69.14
S.D. values:			21.96		28.77		25.96

The simplest way to estimate crop area from a classified image is to count all the pixels assigned to each crop in the classification result and multiply by the corresponding pixel size. However, crop area estimates can be biased due to errors in crop type classification [28,29], so a direct calibration estimator of the area was used following the procedure provided in [30,31]. The approach uses ground truth information to calculate the conditional probability:

$$A_j = \sum_{i=1}^n (A_i * P(j|i))$$

where A_j represents the unbiased area of crop type j , A_i is the total area of pixels classified as crop type i , and $P(j|i)$ is the conditional probability of being crop type j (ground truth) when classified as crop type i . Table 3 is accompanied by Figures 15–23, which show some representative examples of the spatial distribution of key crop segments (“olive groves”,

“vineyards”, “cereals”, “citrus”, “vegetables”, “pears” and “peaches”) for the ten S-RUs in the southern EU, using the RF classifier.

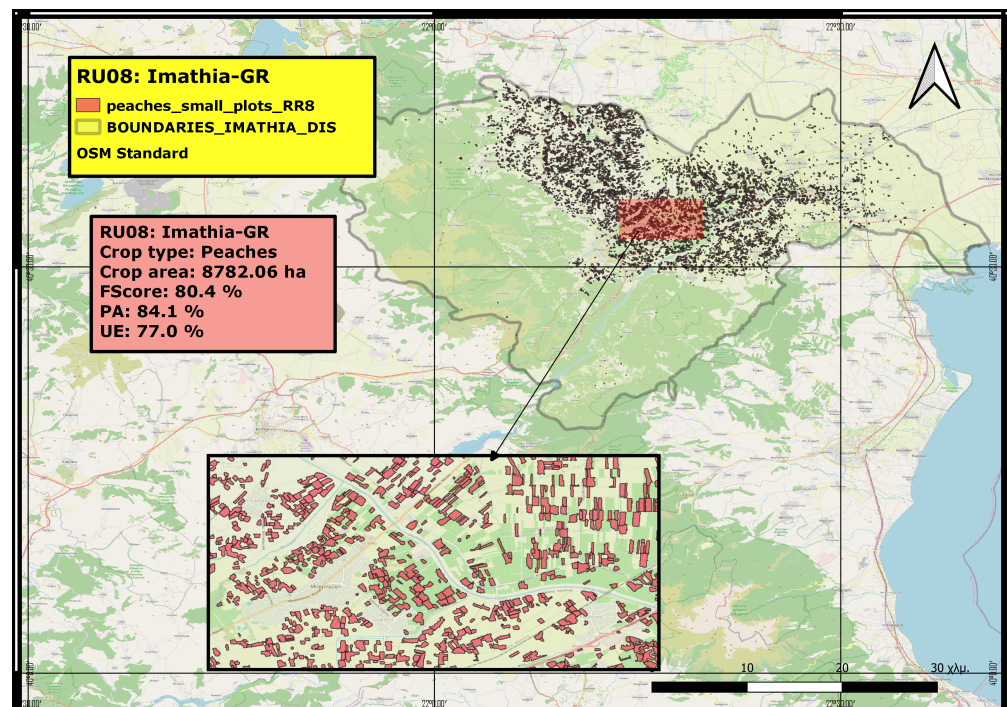


Figure 15. RU08: Imathia(GR); Key crops: “Peaches ” segments derived from the RS approach; year of cultivation 2017–2018.

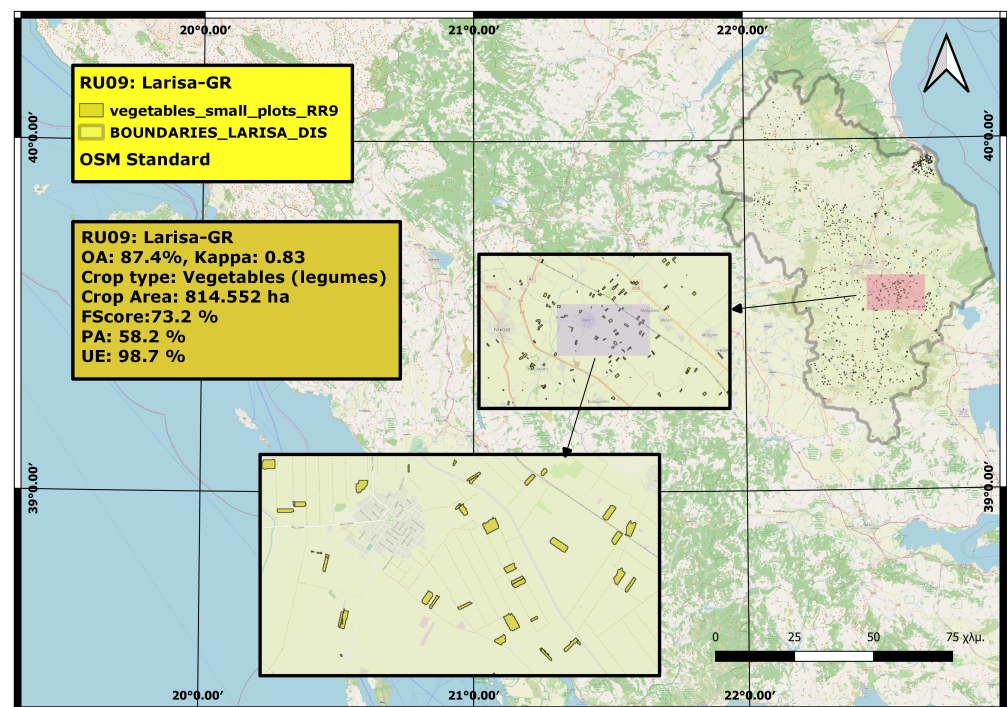


Figure 16. RU09: Larisa(GR); Key crops: “Vegetables (mainly legumes)” segments derived from the RS approach; year of cultivation 2017–2018.

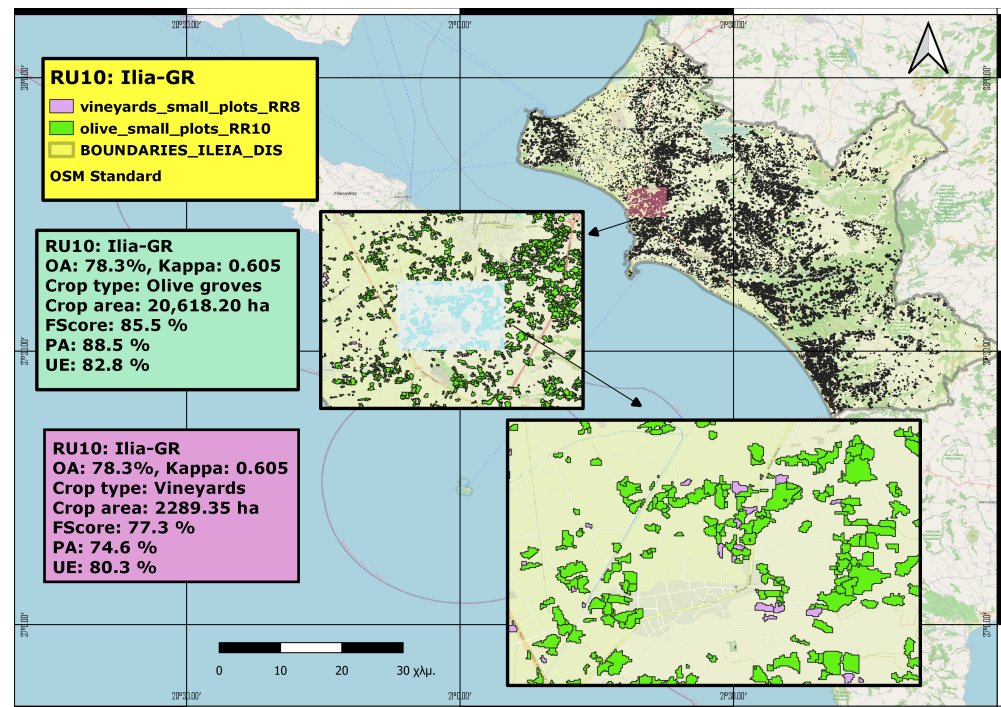


Figure 17. RU10: Ilia(GR); Key crops: “Olive groves” and “Vineyards” segments derived from the RS approach; cultivation year 2017–2018.

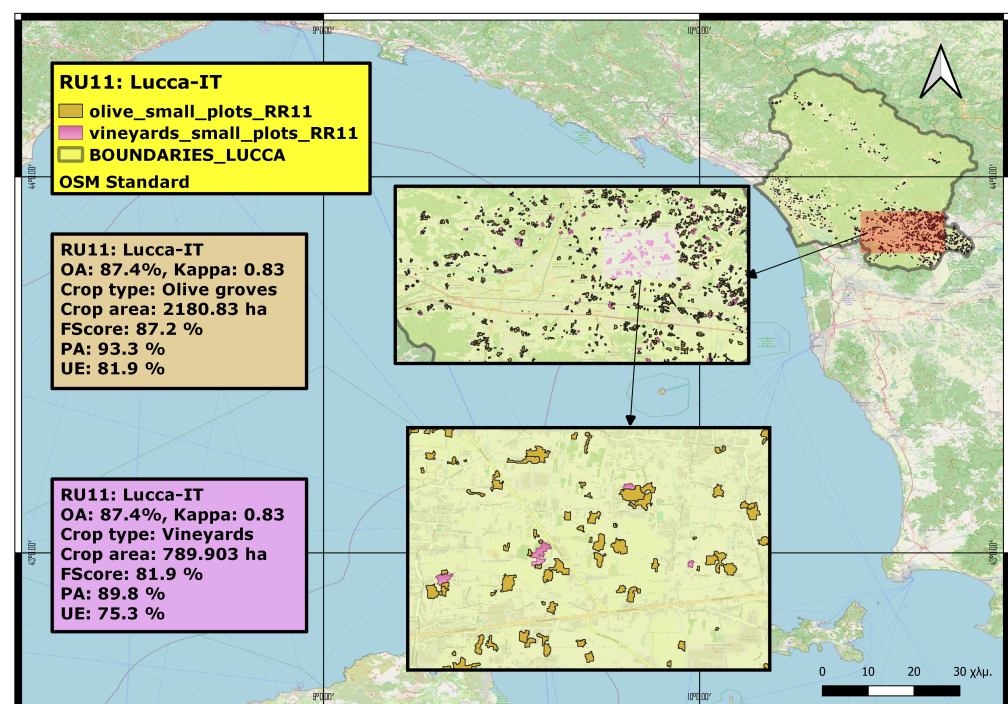


Figure 18. RU11: Lucca(IT); Key crops: “Olive groves” and “Vineyards” segments derived from the RS approach; cultivation year 2017–2018.

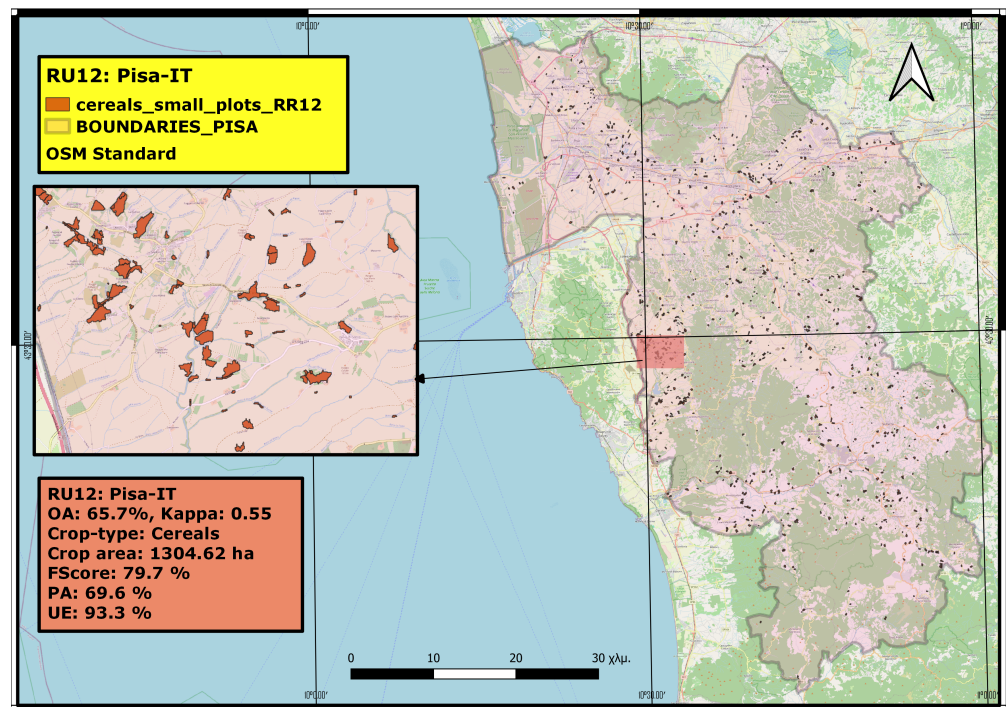


Figure 19. RU12: Pisa(IT); Key crops: “Cereals” segments derived from the RS approach; year of cultivation 2017–2018.

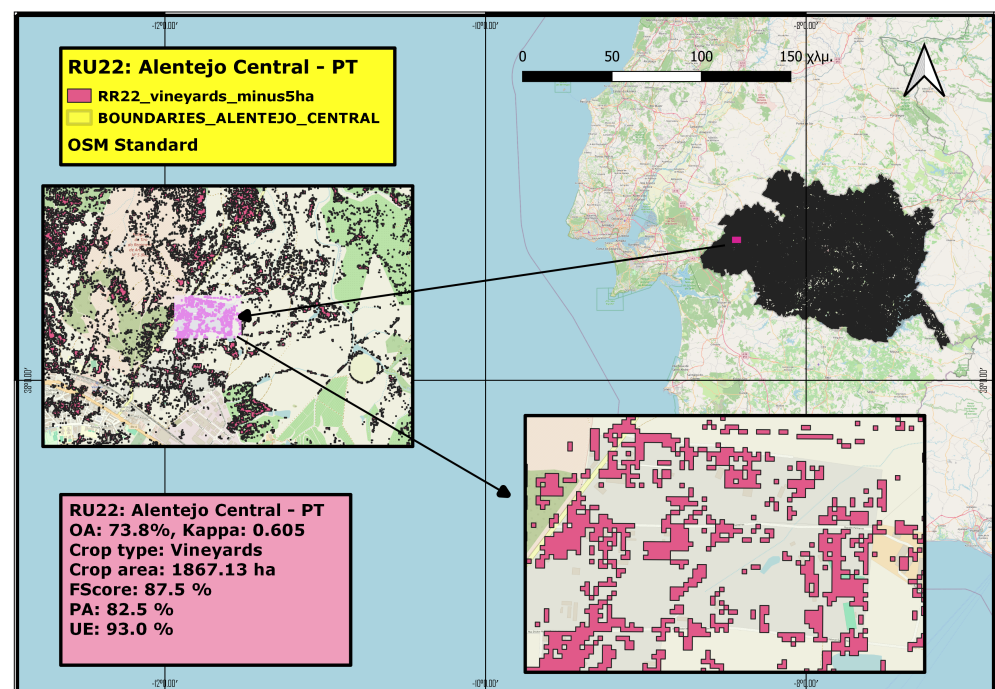


Figure 20. RU22: Alentejo Central (PT); Key crops: “Vineyards” segments derived from the RS approach; cultivation years 2017–2018.

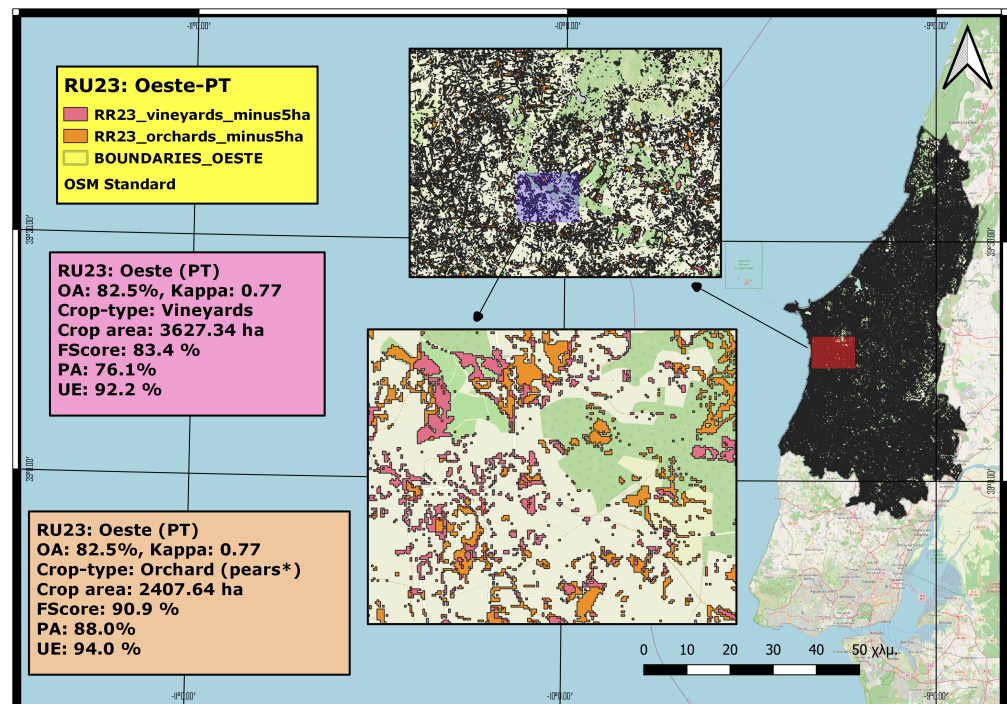


Figure 21. RU23: Oeste (PT); Key crops: “Vineyards” and “Orchards (mainly pears)” segments derived from the RS approach; cultivation year 2017–2018.

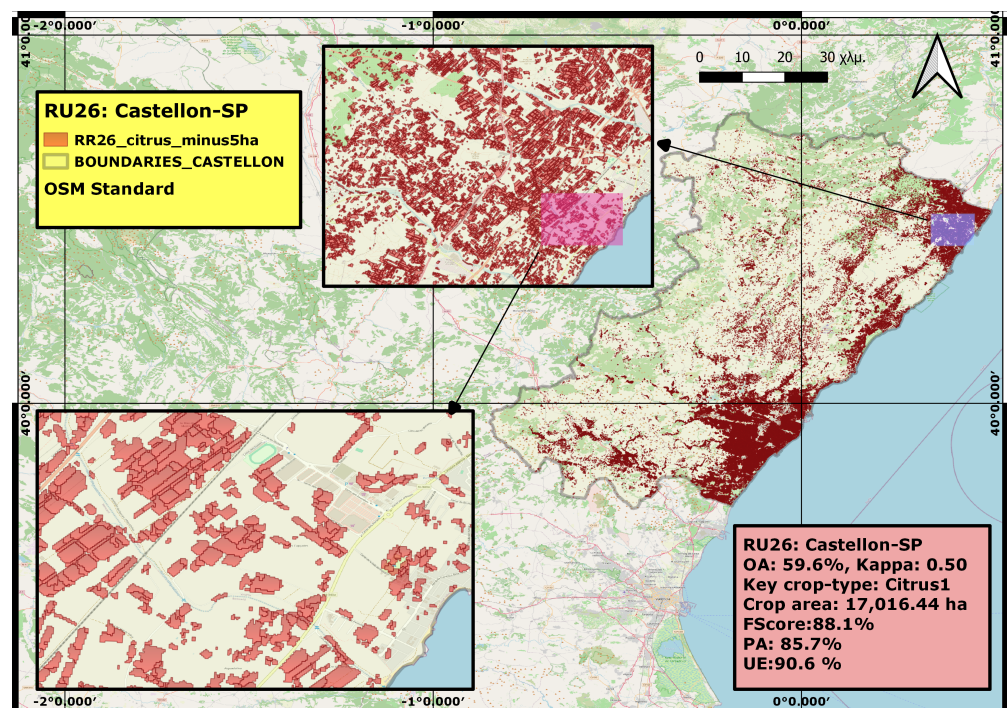


Figure 22. RU26: Castellon(SP); Key crops: “Citrus” segments derived from the RS approach; year of cultivation 2017–2018.

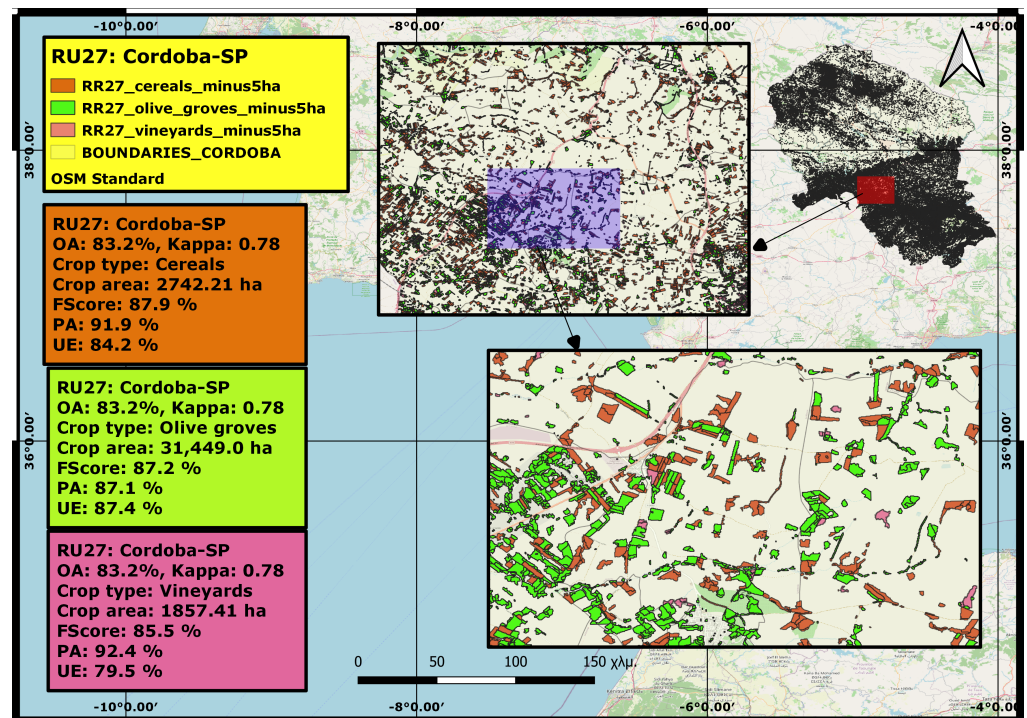


Figure 23. RU27: Cordoba(SP); Key crops: “Cereals”, “Olive groves”, and “Vineyards” segments derived from the RS approach; cultivation year 2017–2018.

Table 4 presents the key crop type for each RU case, along with the *FScore* values, the estimated SFPlot areas derived from the RS-based approach, and the SF areas obtained from rOS (Small plot areas belonging to MFs and LFs are not counted, as in the case of SFPlot areas derived from the RS-based method.). It also shows the number of farmers who participated in the interviews, as well as the estimated self-reported key crop yields and production estimates for each key crop product using the estimated SFPlot areas in each RU. The analysis only included RUs within the EU. There are 29 key crop cases in total, distributed across 20 RUs in 11 EU countries. The rOS surveys on SF agricultural areas reveal a lack of information about the relevant key crops in 13 of 29 cases. These are marked with “na” in the corresponding column of Table 4 and are listed below. Montana-BL (“cereals”); Jihocecky Kraj-CZ (“wheat”); Vaucluse-FR (“vineyards”); Lucca-IT (“olive groves”, “vineyards”); Pisa-IT (“cereals”); Latgale-LV (“cereals”); Pieriga-LV (“cereals”); Nowosadecki-PL (“cereals”); Bistrita-RO (“vegetables”, “orchards”); and Giurgiu-RO (“cereals”, “sunflowers”). Note that in the two cases of RU11: Lucca-IT (cases 11 and 12), the value of 4465.35 hectares belongs to a crop group that comprises the crops of “fruits”, “olives”, “cereals”, “vines”, “vegetables”, and “potatoes”; therefore, it cannot be included in the regression analysis.

The relationship between the 16 key crop cases for which data for area estimates are available from both sources—the RS-based method and the rOS surveys—is depicted in Figure 24. With $R^2 = 0.9621$, the regression analysis shows a fairly good coefficient of determination. Because the $F_{stat} = 355.5347$ statistic is higher than the critical value $F_{1,14}^{crit} = 6.5149$, the linear relationship between the variables is statistically significant at the 1% level ($p\text{-value} = 2.3865 \times 10^{-11} < 0.01$).

Table 4. List of the 29 key crop cases in RUs-EU along with their FScore values, the unbiased RS-estimated key crop areas of SFPlots, the rOS-survey SFs values of key crop areas, the number of individuals interviewed, the self-reported yields (in tons/ha), and finally, the RS-estimated crop production (in tons/year). ‘**’ In the absence of field-level wheat information, the mean national wheat yield was utilized. ‘‡’: For Ilia, the information about vineyard yields was obtained in two different ways: 5 out of the 17 farmers surveyed provided data in tons of grapes per hectare, while 12 reported yields as dried grapes per hectare. To ensure comparability with other vineyard regions, the yields in tons of grapes per hectare were utilized. ‘◊’: The UAA of main crops in SFs includes relevant crop areas totaling 4465.35 ha consisting of “fruits”, “olives”, “cereals”, vines, vegetables, and potatoes. However, these values are not included in the regression analysis. ‘_’: FScore values that are less than 75% are underlined. Estimated values based in RS approach are indicated in bold provided there is an estimated OS survey value to compare with. Finally, for all self-reported yields for which the group interview size is greater than 4 (those less than five are indicated with italics) the corresponding RS-production values are calculated and indicated in bold.

Cases	Reference RU	Crop Types	FScore (%)	RS-Estimated SFPlot Key Crop Areas (ha)	OS Surveys on SFs Key Crop Areas (ha)	Group Interview Size	Self-Reported Yields (tons/ha)	RS-Estimated Production (Tons/Year)
1	RU01: Montana-BL	Cereals	83.4	3523.98	na	1	4.00	14,095.92
2	RU04: Jihocecky Kraj-CZ	Wheat	85.5	2140.48	na	0	6.10 *	13,096.94
3	RU06: Vaucluse-FR	Vineyards	92.1	10,578.40	na	2	4.15	43,900.40
4	RU08: Imathia-GR	Peaches	80.4	8782.06	8125.00	38	29.99	263,373.98
5	RU09: Larisa-GR	Vegetables	<u>73.2</u>	814.55	690.43	6	1.48	1205.54
6	RU10: Ilia-GR	Olive grove	85.5	20,618.20	24,488.06	28	1.31	27,009.84
7	RU10: Ilia-GR	Vineyards	77.3	2289.35	1783.21	5‡	14.68	33,607.66
8	RU11: Lucca-IT	Olive grove	87.2	2180.83	na◊	18	3.98	8679.70
9	RU11: Lucca-IT	Vineyards	81.9	789.90	na◊	13	6.88	5434.53
10	RU12: Pisa-IT	Cereals	75.4	1304.62	na	5	4.06	5296.76
11	RU14: Latgale-LV	Cereals	<u>68.1</u>	6596.26	na	18	2.65	17,480.09
12	RU15: Pieriga-LV	Cereals	88.8	3389.91	na	7	2.01	6813.72
13	RU16: Vilniaus Aps-LT	Vegetables	76.5	1592.14	539.00	11	2.91	4633.13
14	RU19: Rzeszowski-PL	Cereals	91.7	15,603.10	22,896.00	34	4.11	64,128.74
15	RU19: Rzeszowski-PL	Potatoes	86.9	8714.34	6420.00	29	18.41	160,431.00
16	RU20: Nowosadecki-PL	Cereals	<u>70.8</u>	10,779.93	11,069.00	38	3.22	34,711.37
17	RU20: Nowosadecki-PL	Apples	81.4	1705.50	na	9	31.24	53,279.82
18	RU21: Nowotarski-PL	Cereals	<u>70.8</u>	3020.26	2,319.48	26	2.75	8305.72
19	RU22: Alentejo Cen-PT	Vineyards	87.5	1867.13	1299.00	16	5.63	10,511.94
20	RU23: Oeste-PT	Pears	90.9	2407.64	3379.00	16	9.23	22,222.52
21	RU23: Oeste-PT	Vineyards	83.4	3627.34	3931.00	21	6.66	24,158.08
22	RU24: Bistrita-Nas-RO	Vegetables	75.2	532.13	na	0	—	—
23	RU24: Bistrita-Nas-RO	Orchards	98.2	4799.07	na	21	10.81	7630.52
24	RU25: Giurgiu-RO	Cereals	98.2	17,416.00	na	14	3.99	69,489.84
25	RU25: Giurgiu-RO	Sunflower	75.9	6152.74	na	4	—	—
26	RU26: Castellon-SP	Citrus	88.1	17,016.44	14,583.70	10	35.46	603,402.96
27	RU27: Cordoba-SP	Cereals	87.8	2742.21	2245.00	10	2.95*	8089.52
28	RU27: Cordoba-SP	Olive grove	87.2	31,449.00	37,599.00	12	0.76	23,901.24
29	RU27: Cordoba-SP	Vineyards	85.5	1857.41	1252.00	10	8.68	16,122.32
Totals				194,290.93	140,296.67			1,579,098.86

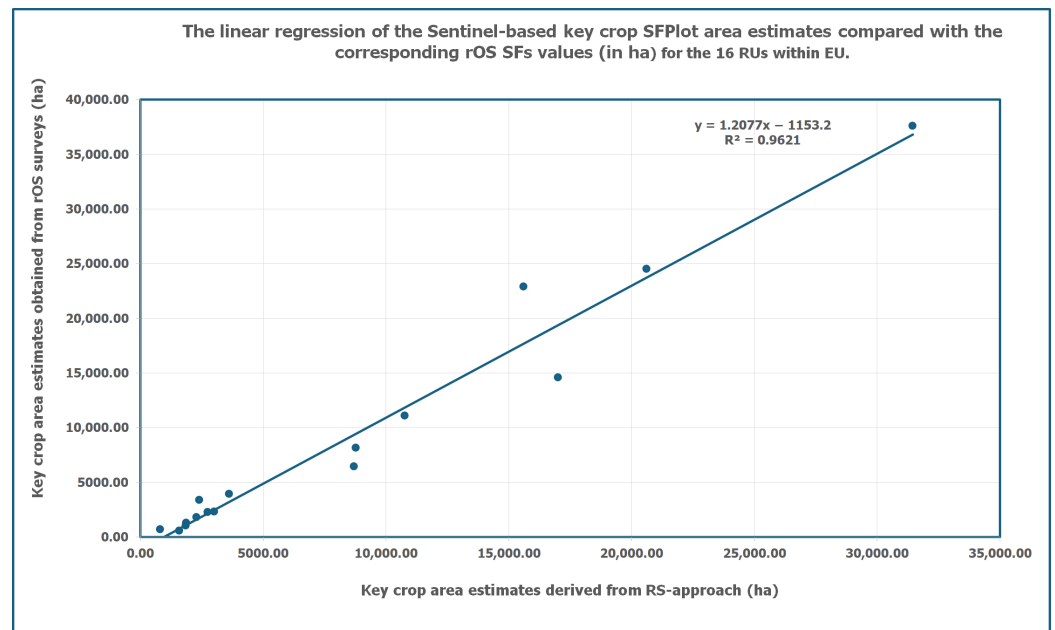


Figure 24. Regression of the Sentinel-based key crop area estimates against the corresponding rOS values of the 16 RUs-EU cases; cultivation year 2017–2018.

To obtain estimates of production per key crop, per RU, the estimated SFPlot key crop areas are multiplied by the corresponding estimated self-reported key crop yields. Regarding the field-level yields acquired from farmers' interviews, it should be noted that for several key crops, such information was lacking, or it was based on relatively few farmers' interviews (see Table 4). This is the case of RU4: Jihocecky Kraj-CZ ("wheat") and RU24: Bistrita-RO ("orchards") (no participating farmers); and of RU1: Montana (one farmer), RU6: Vaucluse-FR (two farmers), RU12: Pisa-IT (five farmers), and RU26: Castellom-SP (four farmers), to name a few (for full details, see Table 4). Thus, only RS-estimated production cases with a non-zero self-reported yield were calculated, with the limitation that cases with a small group interview size need careful consideration. Existing solutions for reducing the influence of probable outliers in self-reported yields and, thereby, lowering error propagation in farmer-provided yield estimates include replacing average yields with a more robust measure, such as Tukey's bi-weight loss function. As such investigations are beyond the scope of this study, we summarize the results already produced from both approaches as follows.

- Although Sentinel-based estimates slightly overestimate or underestimate key crop areas in some cases, the overall results are more comparable to rOS surveys on SFs, demonstrating that crop areas obtained from Sentinel data can be used with confidence, particularly in regions where this information is not available in rOS.
- Tukey's loss function is designed to remove extreme data points that may have an impact on performance assessments. This was the case in this study, when looking at crop production outputs for RUs whose yield estimations were based on only a few farmer interviews per key product [17,18]. However, in order to achieve safe results, relevant issues must be further elaborated.
- In RUs with high accuracy levels ($F\text{Score} > 68.1\%$), the total SFPlot key crop area was estimated to be 194,290.93 ha. This is equivalent to 66,516.75 ha of "cereals" (including "wheat"), 52,067.20 ha of "olive groves", 8333.97 ha of "oil crops" ("sunflower" and "olive groves"), 34,710.71 ha of "fruits" ("apples", "peaches", "pears", "orchards", and "citrus"), 11,653.16 ha of "vegetables" ("potatoes", "tomatoes", "pepper", and "legumes"), and 21,009.53 ha of "vineyards" (see Table 4).

- The “olive groves” in Cordoba (31,449.0 ha) and Ilia (20,618.2 ha); the “citrus” in Castellón (17,016.4 ha); the “cereals” in Giurgiu (17,416.0 ha), Rzeszowski (15,603.1 ha), and Nowosadecki (10,779.9 ha); and the “vineyards” in Vaucluse (10,578.4 ha) are the dominant key crop products in terms of area extent, accounting for 63.5% of the total SFPlot crop area.

Since the estimated SFPlot areas of each key crop case in S-RUs are of interest in this work, we proceed with the regression analysis by splitting the 16 pooled data of the estimated key crop areas into two subsets, one with the 11 S-RUs and the other with the 5 non-S-RUs key crop cases of the RUs-EU. Figure 25 depicts the two regression lines derived from these two datasets. The results deduced from their regression analysis are as follows:

- With $R^2 = 0.9801$, the regression analysis of the 11 estimated key crop areas in S-RUs’ dataset shows an excellent coefficient of determination. Note that the statistic $F_{stat} = 443.2145$ is greater than the $F_{1,9}^{crit} = 8.0215$ at the 1% level of significance ($p\text{-value} = 5.77281 \times 10^{-9} < 0.01$), indicating a very strong linear relationship between the two aforementioned sources of information.
- With a coefficient of determination R^2 value of 0.9201, the regression analysis of the five estimated key crop areas in the non-S-RUs dataset also shows a smaller but still remarkable linear relationship. Again, the statistic $F_{stat} = 34.5366$ is higher than the $F_{1,3}^{crit} = 30.8165$ at the 1% level of significance ($p\text{-value} = 0.0098 < 0.01$). As a result, there is a strong linear relationship between the two sources of information regarding the key crop areas of the five non-S-RUs.

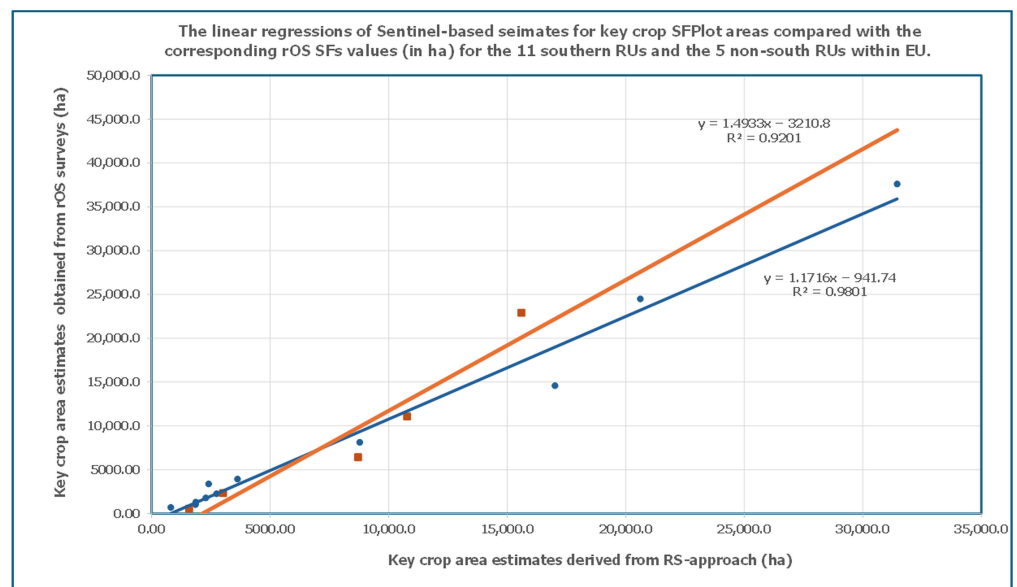


Figure 25. Regressions of the Sentinel-based key crop area estimates against the corresponding rOS values of the 11 S-EU and 5 non-S-EU cases, respectively; cultivation year 2017–2018.

To determine whether the linear regression coefficients and relationships are equal across the two separate datasets of SE and non-SE key crop cases, we perform the Chow test with $n = 16$ (pooled dataset), $n_1 = 11$ (S-RU dataset), $n_2 = 5$ (non-S-RUs dataset), and $dfs = 2$ degrees of freedom. The test result suggests that there is no evidence to reject the null hypothesis since the calculated statistic $F_{stat} = 0.0187$ is smaller than the critical value $F_{2,12}^{crit} = 3.8853$ at a 5% (and also at 1%) significance level. Therefore, the two coefficients of the linear regression are statistically the same in both subsets, and the pooled dataset could be used instead of the two data subsets.

The Chow test presented above has two important limitations to be considered. First, the test can be performed since the available value of $dfs = 2$ is lower than the size of the

smaller sub-sample $n_2 = 5$. Further, the Chow test assumes homoscedasticity, namely that the error variance remains constant in the two subsamples. As it appears, the statistic:

$$F_{stat} = S_2^2/S_1^2 = 8,515,341.33/2,268,582.122 = 3.7536$$

and the critical value is:

$$F_{n_1-2, n_2-2}^{crit} = F_{3,9}^{crit} = 3.8853.$$

Since $F_{stat} < F_{3,9}^{crit}$, there is no evidence to reject the null hypothesis of homoscedasticity at the 5% level (as well as at the 1%). Therefore, the Chow test is valid, and according to its result, fitting a single regression line to a combined dataset can be used instead of fitting two separate lines to the split datasets.

Similar data to Table 4 are presented in Table 5, but they are limited to the 15 cases in S-RUs. Therefore, in addition to the 11 key crop cases of S-RUs, the table also contains the four key crop cases of S-RUs for which rOS surveys have not been reported or made explicit.

Table 5. List of the 15 key crop cases in S-RUs along with their FScore values, the unbiased RS-estimated key crop areas of SFPlots, the rOS-survey SFs values of key crop areas, the number of individuals interviewed, the self-reported yields (in tons/ha), and finally, the RS-estimated crop production (in tons/year). ‘*’: In the absence of field-level wheat information, the mean national wheat yield was utilized. ‘‡’: For Ilia, the information about vineyard yields was obtained in two different ways: 5 out of the 17 farmers surveyed provided data in tons of grapes per hectare, while 12 reported yields as dried grapes per hectare. To ensure comparability with other vineyard regions, the yields in tons of grapes per hectare were utilized. ‘_’: FScore values that are less than 75% are underlined.

Cases	Reference RU	Crop Types	FScore (%)	RS-Estimated SFPlot Key Crop Areas (ha)	OS Surveys on SFs Key-Crop Areas (ha)	Group Interview Size	Self-Reported Yields (tons/ha)	RS-Estimated Production (tons/year)
1	RU06: Vaucluse-FR	Vineyards	92.1	10,578.40	—	2	4.15	43,900.40
2	RU08: Imathia-GR	Peaches	80.4	8782.06	8125.00	38	29.99	263,373.98
3	RU09: Larisa-GR	Vegetables	<u>73.2</u>	814.55	690.43	6	1.48	1205.54
4	RU10: Ilia-GR	Olive grove	85.5	20,618.20	24,488.06	28	1.31	27,009.84
5	RU10: Ilia-GR	Vineyards	77.3	2289.35	1783.21	5 ‡	14.68	33,607.66
6	RU11: Lucca-IT	Olive grove	87.2	2180.83	—	18	3.98	8679.70
7	RU11: Lucca-IT	Vineyards	81.9	789.90	—	13	6.88	5434.53
8	RU12: Pisa-IT	Cereals	75.4	1304.62	—	5	4.06	5296.76
9	RU22: Alentejo C.-PT	Vineyards	87.5	1867.13	1299.00	16	5.63	10,511.94
10	RU23: Oeste-PT	Pears	90.9	2407.64	3379.00	16	9.23	22,222.52
11	RU23: Oeste-PT	Vineyards	83.4	3627.34	3931.00	21	6.66	24,158.08
12	RU26: Castellon-SP	Citrus	88.1	17,016.44	14,583.70	10	35.46	603,402.96
13	RU27: Cordoba-SP	Cereals	87.8	2742.21	2245.00	10	2.95 *	8089.52
14	RU27: Cordoba-SP	Olive grove	87.2	31,449.00	37,599.00	12	0.76	23,901.24
15	RU27: Cordoba-SP	Vineyards	85.5	1857.41	1025.00	10	8.68	16,122.32
Totals				108,325.09	121,364.42			1,096,916.99

In the sequel, we summarize the results already derived and discussed from both approaches.

- Since the R^2 value of the S-RUs dataset is higher than the corresponding R^2 values of the EU-RUs and non-S-RUs datasets, the regression of the S-RUs dataset better presents the SE EU case. However, as it has been pointed out, statistically, we may equally apply the regression of the EU-RUs pooled dataset.
- The results demonstrate that crop area estimates obtained from Sentinel data can be used with confidence, particularly in RUs where this information is insufficient or not available in the rOS, such as the cases of “vineyards” in Vaucluse and Lucca, the “olive groves” in Lucca, and the “cereals” in Pisa.
- Only the “vegetables” key crop in Larisa was classified with accuracy level $FScore = 73.2 < 75.0\%$ without being excluded from the analysis.
- The total key crop area of SFPlots was estimated as 108,384.66 ha, corresponding to 54,248.03 ha of “olive groves”, 21,009.53 ha of “vineyards”, 17,016.44 ha of “citrus”,

4086.43 ha of “cereals”, 8782.06 ha of “peaches”, 2407.64 ha of “pears”, and 814.55 ha of “vegetables” (see Table 4).

- “Olive groves” in Cordoba (31,449.0 ha), Ilia (20,618.2 ha), and Lucca (2180.33 ha); “citrus” in Castellón (17,016.4 ha); and “vineyards” in Vaucluse (10,578.4 ha), Ilia (2289.35 ha), Oeste (3627.34 ha), Cordoba (1847.31 ha), and Lucca (789.90 ha) are the dominant key crop products in terms of area extent, accounting together for 63.5% of the total SFPlots crop area.

Finally, according to the aforementioned statistical analysis, fitting a single regression line to a combined dataset is equivalent to fitting two separate lines to the split datasets.

4. Discussion

The effectiveness of small farming systems is currently constrained by several issues, including the small size of agricultural plots, the fragmented landscapes, the modifications in crop techniques, and the weather fluctuations. As a result, understanding crop growth conditions is essential for enhancing food security by accurately assessing local and regional crop production. This study uses Sentinel imaging to monitor small farming plots (<5 ha) and objectively estimate their relevance in agricultural output. Crop maps created using satellite data can be utilized to estimate crop acreage and production. While it may be simple to identify crop fields in homogeneous agricultural landscapes where large-scale monocultures predominate, getting such information in a highly fragmented and heterogeneous agricultural environment is challenging from the RS standpoint. The current study evaluated Sentinel satellites’ capabilities and efficacy as a data-driven tool for SFPlot monitoring with respect to crop types, crop area extent, and crop productivity, and objectively measured the significance of SFs in agricultural production. Sentinel imagery’s capabilities were tested in various structured farm landscapes using 20 pilot RUs (NUTS-3 level) spread across 11 EU countries and 29 key crop types, yielding valuable insights into their accuracy and effectiveness in terms of SFPlots’ contribution to food security.

The developed RS methodology was based on RF crop identification and classification, which produced the respective LC crop-type maps and captured the many SFPlots of the considered RUs. The produced maps also showed high classification accuracy for many key crops under SFs, as they were proposed by the key informants (experts) in the surveys referenced in [17]. Note that crop identification from RS data has gained significance as free Sentinel data became available, in addition to new algorithms based on machine learning (ML) and deep learning (DL) that have been developed and successfully implemented for classification. Recently, additional studies compared—in detail—the pioneering ML and DL algorithms with respect to their performance as well as their use in applications related to crop identification and LU classification, showing their advantages and limitations ([9,12,13]). These comparisons show that the RF classifier remains one of the top choices, and depending on the application and the available resources, it may be preferred. As a final note on the RS-based framework set, only for the highly accurate key crop products ($F_{Score} > 75\%$) were the corresponding production estimates obtained by multiplying the estimated self-reported crop yields with the unbiased key crop SFPlot area estimations. Therefore, the developed methodology can be extended to other crops in RUs-EU to fully assess the effectiveness of Sentinel data in producing highly accurate crop maps.

We are particularly interested in the 10 RUs distributed in 5 S-EU countries (Greece, Italy, France, Spain, and Portugal) and in 7 key crop types (“olive groves”, “vineyards”, “vegetables”, “cereals”, “citrus”, “peaches”, and “pears”) represented in 15 key crop area estimates. However, to determine each key crop type’s spectral footprint, a representative field dataset should be provided, particularly when SFs dominated the agricultural landscape and crop diversification was the dominant spatial pattern. Table 6 expands Table 2

and Figure 14 by including the basic metrics of *OA*, *Kappa*, and *FScore* to get values for their mean, median, maximum, and minimum for each RU-EU considered.

Table 6. Basic measures of *OA*, *Kappa* and *FScore* values obtained for each RU considered. ‘*’: RU20: Nowosadecki-PL and RU21: Nowotarski-PL use the same Sentinel image.

Metric	EU-RUs	S-RUs	Non-S-RUs
RUs	18 *	10	8 *
$\overline{OA}$	81.41	79.70	83.30
medianOA	83.00	82.90	83.00
maxOA	91.40	91.40	91.40
minOA	59.60	59.60	72.60
<i>Kappa</i>	0.74	0.73	0.75
medianKappa	0.76	0.77	0.76
maxKappa	0.89	0.85	0.89
minKappa	0.50	0.50	0.55
Key crop cases	29	15	14
$\overline{FScore}$	83.27	84.23	82.24
MedianFScore	85.50	85.50	82.40
maxFScore	98.20	92.10	98.20
minFScore	68.10	73.20	68.10

In the sequel, we list the following points that summarize the results presented in Table 6.

- The eighteen EU-RUs’ *OA* values varied from 59.60% to 91.40%, with an average of 81.41% and a median of 83.00%, indicating that nine out of eighteen EU-RUs have values higher than 83.00%. Similarly, the eighteen EU-RUs’ *Kappa* coefficient values varied from 0.50 to 0.89, with an average of 0.74 and a median of 0.76, indicating that nine out of eighteen EU-RUs had a coefficient value larger than 0.76.
- The 10 S-RUs’ *OA* values varied from 59.60% to 91.40%, with an average value of 79.70% and a median of 0.73. It indicates that five of the ten RUs have *OA* values higher than 82.90%. With an average of 0.73 and a median of 0.77, the *Kappa* coefficient varied from 0.50 to 0.85 across S-RUs, indicating that 9 out of 18 S-RUs had a coefficient value higher than 0.77.
- The 8 non-S-RUs’ *OA* values varied from 72.60% to 91.40%, with an average value of 83.30% and a median of 83.00%. It indicates that four out of eight non-S-RUs have *OA* values more than 83.00%. With an average of 0.75 and a median of 0.76, the *Kappa* coefficient varied from 0.55 to 0.89 across non-S-RUs, indicating that four out of eight non-S-RUs had a coefficient value greater than 0.76.
- The EU-RUs with the highest *OA* were RU8: Imathia-GR, *OA*: 91.4% (S-RU); RU24: Bistrita Nasaud-RO, *OA*: 91.4% (non-S-RU); and RU19: Rzeszowski-PL, *OA*: 90.0% (non-S-RU).
- The EU-RUs with the lowest *OA* values were observed in RU26: Castellon-SP, *OA*: 59.6% (S-RU), RU12: Pisa-IT, *OA*: 65.7% (S-RU), and RU14: Latgale-LV, *OA*: 72.6% (non-S-RU).

The above results indicate the suitability of the Sentinel data for producing accurate information about crop area extent in small-scale farming systems, as well as for creating crop maps and estimating crop acreage and production. In fact, this study showed the Sentinel satellites’ capabilities and efficacy as a data-driven tool for SFPlot monitoring with respect to crop types, crop area extent, and crop productivity, and objectively measured the significance of SFs in agricultural production. Sentinel imagery’s capabilities were tested

in various structured farm landscapes using a total of 19 pilot RUs (NUTS-3 level) spread across 11 EU countries (10 S-RUs and 9 non-S-RUs in 5 S-EU and 6 non-S-EU countries, respectively), yielding valuable insights into their accuracy and effectiveness in terms of SFs' contribution to food security (Table 4).

Regarding the crop production estimations created by integrating the unbiased crop area with the field-level crop yields, the data presented herein demonstrated that SFPlots provide a considerable contribution to crop output for the selected key crops. Table 4 demonstrates that from the 29 key crop cases spread across 20 RUs-EU, the RS-estimated SFPlot key crop areas were calculated in all of them. Only in four RUs did all the corresponding *FScore* key crop values not exceed the threshold of 75%. Nevertheless, to include in the analysis at least one key crop per RU, this restriction was relaxed, and the closest to the *FScore* key crop threshold was allowed. This is the case of RU09: Larisa-GR (vegetables); RU14: Latgale-LV (cereals); RU20: Nowosadecki (cereals); and RU21: Nowotarski-PL (cereals), with only one case belonging to the S-RUs and the other three to the non-S-RUs. The following points explain the *FScore* metrics presented in Table 6.

- The *FScore* key crop values of the 29 RUs-EU varied from 68.10% to 98.20%, with an average of 83.27% and a median of 85.50%, indicating that 14 out of 29 RUs-EU have values higher than 85.50%.
- The 15 S-RUs' *FScore* values varied from 73.20% to 92.10%, with an average value of 84.23% and a median of 85.50. This indicates that seven out of 15 S-RUs have *FScore* values higher than 85.50%.
- The 14 non-S-RUs' *FScore* values varied from 68.10% to 98.20%, with an average value of 82.24% and a median of 82.40%. It indicates that seven out of 14 non-S-RUs have *FScore* values higher than 82.40%.

Note that the total regional production was determined in Table 4 using the RS-based approach for 27 out of 29 key crop cases, since for the two excluded cases the self-reported yield values of non-S-RUs were not provided.

In addition to the aforementioned standard accuracy measures used to assess Sentinel data suitability for producing accurate information about crop area extent in small-scale farming systems, Sentinel-based crop area estimates were compared to readily available rOS data. The analysis showed that there was a significant correlation between rOS and Sentinel-based estimates with $R^2 = 0.96$ and $p\text{-value} < 0.001$ (see Figure 24). This suggests that in RUs where no rOS data was available, the Sentinel imagery can be the main source for estimations of crop area extent that are reasonably accurate. Note that from the total of 29 key crop cases, the rOS data exist only for 16 of them. Of these, 11 belong to the S-RUs subset and 5 to the non-S-RU subset. The splitting of the pooled dataset RUs-EU into the two sub-datasets of S-RUs and the non-S-RUs revealed that the linear regressions are statistically the same in both subsets, despite the fact that the regression analysis between the rOS and the Sentinel-based estimates of the RUs-EU sub-datasets showed higher correlation with $R^2 = 0.98$ and $p\text{-value} < 0.001$ compared to the one based on the pooled dataset (see Figure 25). The linear regression over the pooled dataset may therefore be applied equivalently, in statistical terms, in both situations.

Future Research

The present study has achieved its goal of assessing the capabilities and efficiency of Sentinel satellites as a data-driven tool for monitoring SFPlots in terms of crop types, crop area coverage, and crop productivity, and objectively measuring the importance of SFs in agricultural production. Future work may address the following aspects.

- The 75% threshold ensures that crop production estimation is obtained only for the best classified crops. This criterion is not set artificially, but it aligns with the third (upper)

quantile *FScore* values of key crop types in both lists: all the 21 RUs selected and the ten southern RUs. The developed methodology allows extending the research and covering meaningful crop cases by creating an index that will combine the estimated *FScore* values achieved per crop with its relevant importance at the regional level. The introduction of such an index will show the effectiveness of Sentinel data in producing highly accurate crop maps; however, it requires further study on the details of how it will be applied.

- Crop production is obtained by multiplying the remotely estimated area by self-reported yield, while some regional yield estimates are based on only a few farmer interviews. These estimates may contain uncertainty from both crop classification and survey data. Usually, a sensitivity analysis follows; however, given the scale of this project and the unclear nature of how and in what format the survey data will be provided, conducting such an analysis will be complicated. As we stated earlier (see Section 3), existing solutions include replacing average yields with a more robust Tukey's bi-weight loss function measure. Therefore, the problem needs further elaboration in terms of the combined uncertainty coming from both sources.
- An interesting aspect of this study will be to compare the RF with the most prominent SVM and DNN data-driven methods, exploring their performance in modeling and predicting crop yields, particularly under a small/moderate sample data set (at about 100/500 samples per RU). Note that in crop yield prediction, it is difficult to obtain a large amount of real crop yield data, as it takes a lot of workforce, material resources, and time to collect sample data in the field. The RF-based approach may risk overfitting when trained on small or moderate datasets, but it offers an integrated algorithm that can significantly improve the performance of a single classifier. Currently, the three RS-based methods mentioned above are employed in numerous RS-based classification studies. Therefore, the present work leaves room for a comparable analysis of the various RS-based methods utilizing both the S1 and S2 imagery.
- The work contributes to the direction of European Commission (EC) policy regarding the monitoring of crop growth using RS. Similar and likely more advanced RS-based algorithms have been proposed and tested by the Joint Research Center (JRC) of the EC and are currently operating effectively within the framework of the Land Parcel Identification System (LPIS) of the Integrated Administration and Control System (IACS) in the EU [32,33]. This long-term trial is conducted by JRC/EC and Eurostat/EU, and it is being implemented in nearly all member states [34]. With the advancements in RS Sentinel technology and a more precise classification of products, regional modeling and yield evaluation for various crop types can be effectively conducted. In this context, the present work contributes to and promotes similar initiatives.

Author Contributions: Conceptualization, T.A.T. and S.G.; methodology, T.A.T. and S.G.; software, S.G., T.A.T., K.A. and R.M.; validation, S.G., T.A.T., K.A. and R.M.; investigation, T.A.T., S.G., K.A. and R.M.; resources, K.A. and R.M.; visualization, T.A.T.; data curation, S.G. and T.A.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

Acknowledgments: The field campaign took place in autumn 2017. The authors wish to express their gratitude to Nuno Guiomar, Patricia Lourenço, and Teresa Pinto-Correia of the Portuguese team of the SALSA project [19] for their valuable cooperation.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Study Areas

For each region, information is provided for the project period (2017–2020) on the administrative and geographical aspects, on geophysical information related to the main agricultural productivity, and on some basic agro-economic elements of the region related to our study.

Appendix A.1. Regional Unit of Vaucluse (FR)

Provençal Vaucluse (French: Vaucluse [voklyz]; Classical norm: Vauclusa [vaw'kluza] or Mistralian norm: Vau-Cluso [vaw'kluzo]), Department: FRL06 (NUTS-3), is in the southern part of France, with its geographic center at 44°0.00' N latitude and 5°10.00' E longitude. It forms one of the six departments (NUTS-3) of the NUTS-2 province FRL: Provence-Alpes-Côte d'Azur and spans an area of 3520 km², making up about 0.0055% of France's total land area. The province (hereafter referred to as RU-06: Vaucluse (FR)) is an enclosed region bordered by the provinces of Drôme (NUTS-3, FRK23) to the north; Ardèche (NUTS-3, FRK22) to the northwest; Gard (NUTS-3, FRJ12) to the west; Bouches-du-Rhône (NUTS-3, FRL04) to the south; Var (NUTS-3, FRL05) to the southeast; and Alpes-de-Haute-Provence (NUTS-3, FRL01) to the east. Avignon is the capital of Vaucluse, with 572,056 residents as of 2023.

Since the 1980s significant agricultural growth has been made possible due to its Mediterranean environment and the installation of a dense irrigation system. During this time, the "Rhône corridor" developed into a major agricultural production exchange, and several specializations were discovered throughout the region. Specifically, the northern department developed more wine production, whereas the Comtat Venaissin plain is more specialized in producing fruits and vegetables. While livestock, lavender, and cereals are the main products of the Albion plateau situated in the western part of Vaucluse, the Durance valley in the southern part mainly produces grapes, wine, and cherries. As a result of this regional specialization, the average farm size increased, although the number of farms decreased. Recently, due to the aging of farmers, a rise in urbanization and a decline in traditional farming are observed. Compared to 40% in 1970, agriculture has now declined to 34% of Vaucluse's total area. The cause of this abandonment can be found in the market evolution, the decrease in the regulatory tools of the Common Agricultural Policy, and the demographic pressure.

Appendix A.2. Regional Unit of Imathia (EL)

Imathia (Greek: Ημαθία [i.ma.'θi.a], romanized: Imathia [Ima'θia], NUTS-3, EL521) is in northern Greece, with its geographic center at 40°35.00' N latitude and 22°15.00' E longitude. It forms part of the Region of Central Macedonia (NUTS-2, EL53) and spans an area of 1701 km², making up about 1.29% of Greece's total land area. The Imathia province (hereafter referred to as RU-08: Imathia (GR)) is bordered by RU-Pieria (NUTS-3, EL525) to the south, RU-Kozani (NUTS-3, EL531) to the west, RU-Pella (NUTS-3, EL524) to the north, and RU-Thessaloniki (NUTS-3, EL522) to the east. The region's capital is the city of Veria, which lies in the central-northern part of the plain. Imathia is divided into three municipalities. Between the 2011 and 2021 censuses, the total population declined by 6.8%, dropping from 140,611 to 131,001 residents. The population density also decreased slightly, from 83 to 82.66 residents per km². RU-Imathia's GDP per capita was €10,953 in 2017, €11,450 in 2019, and €11,850 in 2021 (provisional data). For context, the GDP per capita in 2010 was €14,093 but saw declines of 16.9% in 2013, 22% in 2016, 22.3% in 2017, 18.8% in 2019, and 15.9% in 2021.

Eurostat reported that in 2016, RU-Imathia had a GDP per capita of \$11,100 and a total population of 141,436, with a population density of 84 people per square kilometer. Small

farms made up 79.42% of all agricultural holdings, out of a total of 12,570. In the same year, there were 13,197 agricultural holdings in total, spanning a UAA of 55,056 hectares, within a broader agricultural area of 64,200 hectares. The average physical farm size was about 4.2 hectares, whereas small farms averaged just 1.79 hectares of UAA, both figures being slightly below the national average of 4.65 hectares. In the northeast of RU-Imathia, the lower course of the Aliakmonas River carves out a broad plain known as Kampania or Roumlouki, which is famous for its abundant fruit cultivation, particularly peaches and strawberries. The Vermio Mountains, rising to 2052 meters above sea level near Naousa, dominate the western reaches of Imathia province. To the south, the Pierian Mountains extend into the region below the Aliakmonas River.

The province is the second-largest producer of peaches in Greece. Extensive irrigated agriculture characterizes the region, with large tracts devoted to peach monoculture, alongside areas producing cherries, pears, apples, cotton, corn, and sugar beets. Vineyards for wine thrive along the mountain foothills, further enriching the agricultural landscape.

Appendix A.3. Regional Unit of Larisa (EL)

Larisa (Greek: Λάρισα [ˈla.ri.sa], romanized: Lárissa [ˈlarisa], NUTS-3, EL612) stands as the second-largest regional unit in central northern Greece. The area is positioned at 39°38.30' N latitude and 22°25.00' E longitude, covering 5387 km²—about 4.09% of Greece's total area—and occupying roughly a third of Thessaly's western region (NUTS-2, EL61). The province (hereafter referred to as RU-09: Larisa (GR)) shares borders with RU-Kozani (NUTS-3, EL531) to the northwest, RU-Pieria (NUTS-3, EL525) to the northeast, the Aegean Sea to the east, RU-Magnesia (NUTS-3, EL613) to the southeast, RU-Phthiotis (NUTS-3, EL644) to the south, RU-Karditsa (NUTS-3, EL611) to the southwest, and RU-Trikala (NUTS-3, EL611) to the west. The capital city, Larisa, is situated in the north-central part of the plain and serves as the administrative center for the unit's seven municipalities. According to the 2011 census, the population totaled 283,727, but by the 2021 census, it had decreased to 268,963—a decline of 5.4%. Correspondingly, population density fell from 53 to 50 inhabitants per km².

Eurostat data from 2016 show that the GDP per capita in the region stood at \$13,600, with a total population of 283,727 and a density of 53 inhabitants per km². That year, small farms made up 51% of all 28,202 farms. Altogether, there were 24,999 holdings spanning 201,555 hectares of UAA out of a total agricultural area of 214,390 hectares. The average farm size was roughly 8.06 hectares, while SFs averaged 2.01 hectares of UAA—both figures larger than the national average of 4.65 hectares. RU-Larisa and Thessaly as a whole are particularly suitable for validating LC mapping due to their diverse landscapes, which include mountains, plains, and mixed land uses. About 45% of RU-Larisa consists of flat terrain, 25% is semi-mountainous, and 30% is mountainous, featuring the famed Olympus and Ossa mountains. Mount Olympus, Greece's tallest peak at 2917 m above sea level, rises in the northeastern part of RU-Larisa. Nearby, Mount Ossa lies to the east, along the Aegean coast. The region's north is forested, and the lower reaches of the Pinios River wind through the scenic Valley of Tempe between Olympus and Ossa. It exemplifies the classic northeastern Mediterranean landscape, both in land cover and climate, characterized by hot summers and mild winters. The region's small farms primarily cultivate fodder, cereals, cotton, olives, fruits, and nuts.

Appendix A.4. Regional Unit of Ilia (EL)

Elis, also known as Ilia (Greek: Ηλεία [i.ˈli.a], romanized as Ilia [Il'ia], NUTS-3, EL633), is in southwestern Greece at approximately 37°40.00' N latitude and 21°30.00' E longitude. As part of the Region of Western Greece (NUTS-2, EL63), this area lies on the western side

of the Peloponnese peninsula and spans 2618 km², making up about 1.99% of the total area of Greece. The province (hereafter referred to as RU-10: Ilia (GR)) shares borders with RU-Messinia (NUTS-3, EL653) to the south, RU-Arcadia (NUTS-3, EL651) to the west, RU-Achaea (NUTS-3, EL632) to the north, and the Ionian Sea to the east. The capital city is Pyrgos, and following the most recent administrative reform, the province is now divided into seven municipalities. According to census data, the population decreased from 159,320 in 2011 to 149,896 in 2021—a decline of 5.9%—with population density dropping from 61 to 57 residents per km². It is also notable that the region's GDP per capita, which stood at €13,485 in 2010, fell by 22.3% by 2013, by 21.7% in 2016, by 18.64% in 2017, by 16.2% in 2019, and by 20.2% in 2021 (provisional data).

According to Eurostat, in 2016, the GDP per capita in the region stood at \$10,557, with a total population of 157,174 and a population density of 60.8 inhabitants per square kilometer. SFs made up 76.9% of all farms, amounting to 23,781 holdings. That same year, there were 26,573 agricultural holdings in total, covering a UAA of 99,100 hectares out of a total agricultural area of 120,100 hectares. The average farm size was 4.95 hectares, notably larger than the national average of 4.65 hectares, while SFs averaged 1.78 hectares of UAA. The region is traversed by three main rivers—Alfios, Pinios, and Neda—which all flow toward the Ionian Sea.

Appendix A.5. Regional Unit of Lucca (IT)

Lucca province (Italian: Lucca [ˈlukka], NUTS-3, ITI12) lies along the northwestern coast of Italy, centered at latitude 43°51.00' N and longitude 10°31.00' E. It is part of the Tuscany region (NUTS-2, ITI1), bordered by the Ligurian Sea to the west. To the northwest, it neighbors Massa e Carrara (NUTS-3, ITI11), while Pisa (NUTS-3, ITI17) lies to the south, Pistoia (NUTS-3, ITI13) to the northeast, and Firenze (NUTS-3, ITI14) to the east. Its northern boundary meets the Emilia-Romagna region (NUTS-2, ITH), specifically the provinces of Reggio Emilia (NUTS-3, ITH53) and Modena (NUTS-3, ITH54). The province (hereafter referred to as RU-11: Lucca (IT)) covers an area of 1773.0 km², making up 0.5869% of Italy's total territory. The city of Lucca serves as its capital, and the province itself is divided into 33 municipalities. Population numbers in Lucca have remained relatively steady, with an average annual change of around −0.15%. Counts have ranged from 372,244 in the 2001 census to 388,327 in 2011, and finally to 382,464 as of the 2021 census.

Based on Eurostat data from 2016, Lucca province recorded a GDP per capita of \$29,200 and a total population of 391,228, resulting in a population density of about 220.66 inhabitants per km². The region's overall agricultural area amounted to 47,200.98 hectares, with 24,343.77 hectares designated as UAA, and 6543 agricultural holdings in total. Farms averaged 3.72 hectares, closely matching the 3.71 hectares of UAA per corresponding farm structure. Lucca's rural territory is divided into three distinct sub-areas: Versilia, the Piana di Lucca, and the Valle del Serchio. Versilia occupies a portion of the Tuscan coast, separated from the inland mountains and plains by a short chain of hills and foothills. The area displays noticeable trends of crop abandonment and natural regrowth of shrubs and trees, especially in the lowlands and on some hillsides. The coastal plain is heavily urbanized, leaving only fragmented rural pockets—primarily comprised of vegetable gardens, small olive groves, and orchards. Farming in the province is dominated by small production units, with a strong focus on intensive land use. The Piana di Lucca, acting as a “hinge area”, is characterized by a broad expanse of flat terrain. Over the past sixty years, the growth of urban development has increasingly marginalized and fragmented agricultural land, resulting in the widespread abandonment of traditional farming practices. Finally, the Valle del Serchio is notable for its mountainous terrain. In the southern part of this area,

land is used intensively, while the higher regions are devoted to sheep grazing—meadows and pastures prevail here—and activities connected to forestry.

Appendix A.6. Regional Unit of Pisa (IT)

The province of Pisa (Italian: Pisa ['pi:za], NUTS-3, ITI17) is located in central-western Italy at approximately 43°43.00' N latitude and 10°24.00' E longitude. It belongs to the Tuscany region (NUTS-2, ITI1) and shares borders with Lucca province (NUTS-3, ITI12) to the north, Florence (ITI14) and Siena (ITI19) to the east, Grosseto (ITI1A) to the south, and Livorno (ITI16) and the Ligurian Sea to the west. The province (hereafter referred to as RU-12: Pisa (IT)) covers 2444.72 km², making up about 0.8% of Italy's total area. Its total population is 421,382, with a population density of 172.36 inhabitants/km². Pisa, the provincial capital, is notable for being Tuscany's second most populous city and ranks fifth in area within the region. The province is divided into 39 municipalities. Population numbers have shown a modest increase, from 384,555 in 2001 to 411,190 in 2011, reaching 417,141 by the 2021 census.

Data from Eurostat indicate that, in 2016, the GDP per capita was \$28,100. The agricultural sector features a total area of 158,576.23 hectares, with 95,754.35 hectares of UAA and 6912 farms. The average farm size stands at 3.72 hectares, which is comparable to the 3.63 hectares of UAA per SF. Within the province, two main local subsystems are identified. The first, known as the coastal subsystem of Pianura dell'Arno (the Arno River plains), occupies the northern part of the region and is distinguished by its significant ecological value and protected areas. The second is the subsystem known as Colline Interne e Meridionali, or the internal and southern hills, which lies mainly in the south-central part of the province and is characterized by extensive forested landscapes. In the proximity of the coastal land, the hills predominantly host traditional woody crops such as specialized olive groves.

Appendix A.7. Regional Unit of Alentejo Central (PT)

The province of Alentejo Central (Portuguese: Alentejo Central [alẽ'tɛʒu sɨ'tɾal] (PT1C4)) is one of the thirty NUTS-3 regions of Portugal located in the interior south of the country, centered at latitude 38°34.0' N and longitude 7°54.0' W. It is part of the NUTS-2 Alentejo region (PT1C), bordered by the Alto Alentejo (NUTS-3, PT1C3) to the north, the Lezíria do Tejo (NUTS-3, PT1D3) to the northwest, the Alentejo Litoral (NUTS-3, PT1C1) to the southwest, the Baixo Alentejo (NUTS-3, PT1C2), and the Badajoz-Extremadura (NUTS-3, ES431) of Spain to the east. The province (hereafter referred to as RU-22: Alentejo Central (PT)) covers an area of 7393 km² and is home to 156,207 inhabitants, resulting in a population density of 21.1 inhabitants per km² (2016). The capital, Évora, is adjacent to the Distrito de Santarém, offering access to the Tagus River basin and connections to the Lisbon metropolitan area. The climate is Mediterranean, with rains distributed throughout the year and summers being hot and dry.

The RU-22: Alentejo Central (PT) mainly has a silvopastoral system, which includes open areas with holm and cork oaks, some scattered shrubs, and large pastures for grazing in large-scale farm units. In a mosaic with Montado, there are open pastures and also localized fields of annual crops. In certain areas, small and medium-sized farms feature vineyards and olive groves. These olive groves include both extensive traditional varieties and intensive, superintensive ones that are irrigated. A small-scale mosaic of different permanent and yearly crops is found around cities and villages. It produces around 21% of the national production of olives, 9.4% of the national production of wine grapes, and 2.3% of the national production of tomatoes, and is an important producer of calves, cattle, pigs, and sheep.

Appendix A.8. Regional Unit of Oeste (PT)

The province of Oeste (Portuguese: Oeste [o'este], NUTS-3, PT1D1), officially the Oeste Intermunicipal Community, is in the western section of the Centro (NUTS-2, PT19) region of interior-south Portugal, at latitude 39°24.00' N and longitude 9°8.00' W. It is bordered by the provinces: Região de Leiria (NUTS-3, PT193) to the north; Lezíria do Tejo (NUTS-3, PT1D3) to the east; Metropolitana de Lisboa (NUTS-3, PT1C1) to the south; and Península de Setúbal (NUTS-3, PT1B0) to the southeast. To the west, the Oeste province (hereafter referred to as RU-23: Oeste (PT)) extends to the Atlantic Ocean. It is one of the three NUTS-3 provinces of the Oeste e Vale do Tejo (NUTS-2, PT1D) region, organized in twelve municipalities, occupying a total area of 2220.16 km² or about 2% of the Portuguese territory. The Oeste landscape is characterized by small hills and mountains separated by fertile fields and mostly sandy and acidic soils. The Mediterranean temperate climate, with cool summers and mild winters, is ideal for the production of many fruits and vegetables all year round. Pine forests lie to the north and northeast, whereas two low mountain ranges and the Tagus Valley are in the interior. Oeste has long been acknowledged as one of Portugal's leading agricultural regions, where agriculture is central to most activities in the primary sector.

Appendix A.9. Regional Unit of Castellon (SP)

The province of Castellon (Spanish: Castellón [kaste'lon]; Valencian: Castelló [kaste'lo]; NUTS-3, ES522) is in the northern part of the Valencian Community (NUTS-2, ES52) region of southern Spain. It is located near the eastern coastline of the Iberian Peninsula by the Mediterranean Sea, at latitude 39°24.00' N and longitude 9°8.00' W. It is bordered by the provinces of Valencia (NUTS-3, ES523) to the south, Teruel (NUTS-3, ES5242) to the west, Tarragona (NUTS-3, ES514) to the north, and the Mediterranean Sea to the east. The western side of the province (hereafter referred to as RU-26: Castellón (SP)) is in the mountainous Sistema Ibérico area. It is a relatively small province in Spain (6662 km²), situated on the Mediterranean coast with a declining population of 594,423 inhabitants and an average population density of 89.23 people per km². This population is unevenly distributed in the region and mainly concentrated on the coast, where most economic activities are located. Castellón's capital is Castellón de la Plana.

Castellón's territory is clearly divided between the coast and inland due to its physical features, and this divide is reflected in the socio-economic profile of the two distinct areas and in the distribution of the different types of agricultural activity. The lowland strip along the coast is where the main economic sectors are located, including tourism and industry, and the largest urban centers (Castelló de la Plana, Vila-real, and Vinarós), hence concentrating most of the region's population. The agricultural activity in this area is mostly irrigated and is dominated by citrus crops, mainly clementines, with some horticulture in the north of the region. The majority of Castellón inland is mountainous with low population densities and relatively low elevations (less than 1000 m). Agriculture in this area is mostly rain-fed, dominated by almond and olive trees, which often are part of mixed farms that combine different crops depending on the area (cereals, pastures, and other tree crops) or animal farms (mainly intensive pigs or poultry and some semi-extensive cattle and sheep). The farming systems located in both areas were affected by these overall transformations (in terms of land-use competition, available labor, part-time opportunities, and access to markets).

Appendix A.10. Regional Unit of Cordoba (SP)

The province of Cordoba (Spanish: Córdoba ['ko.ɾðoβa], NUTS-3, ES613) belongs to the autonomous community of Andalusia (NUTS-2, ES61). It borders the provinces of Seville

(NUTS-3, ES618) to the west, Málaga (NUTS-3, ES617) to the southwest, Granada (NUTS-3, ES614) to the south, and Jaén (NUTS-3, ES616) to the east. It also borders the province of Badajoz (NUTS-3, ES431) to the northwest, which belongs to the region of Extremadura (NUTS-2, ES43), and the province of Ciudad Real (NUTS-3, ES422) to the northeast, which is part of the autonomous community of Castile-La Mancha (NUTS-2, ES42).

The province of Cordoba (hereafter referred to as RU-27: Cordoba (SP)) is located in the south of the Iberian Peninsula at latitude 37°53.24' N and longitude 4°46.48' W, and has an area of 13,769 km², which represents 2.7% of the total area of Spain. Its total surface area has 802,575 inhabitants, with a population density of 58.28 inhabitants per km², lower than the Spanish average (93.17 inhabitants/km²). A characteristic of the productive sector's structure is the importance of the agricultural sector and the industrial sector linked to agrifood. The province is mainly divided into three geographic zones that have been decisive in the socio-economic and demographic configuration of the territory: Sierra Morena in the north, the Guadalquivir Valley in the center, and the Cordilleras Béticas in the south. The territories located in the north, with a high natural capital, have traditionally been characterized by a weak connection to the provincial capital and by demographic problems associated with aging and depopulation. However, the relevance of livestock breeding and the use of endogenous resources have led to a major food industry with quality products such as Iberian ham, milk, dairy products, and ecological olive oil. On the other hand, the central and southern areas present a major demographic potential and rural-urban balance. They are characterized mainly by the importance of the agricultural sector and agrifood industry, in which olive-growing and the production of olive oil acquire great relevance.

Two types of development models can be observed in the rural environment of the region: The central and southern territories have followed a local development model similar to so-called 'agglomerations', facilitating the installation of diverse economic activities; the northern territories have followed a process of 'specialization' of livestock-breeding products (milk and meat), i.e., a model of local development that can be identified as 'development by product' and is based on certain comparative advantages in the area. The farming systems located in both areas were affected by these overall transformations (in terms of land-use competition, available labor, part-time opportunities, and access to markets).

References

1. Lowder, S.K.; Skoet, J.; Raney, T. The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Dev.* **2016**, *87*, 16–29. <https://doi.org/10.1016/j.worlddev.2015.10.041>.
2. Guiomar, N.; Godinho, S.; Pinto-Correia, T.; Almeida, M.; Bartolini, F.; Bezák, P.; Biró, M.; Bjørkhaug, H.; Bojnec, S.; Brunori, G.; et al. Typology and distribution of small farms in Europe: Towards a better picture. *Land Use Policy* **2018**, *75*, 784–798. <https://doi.org/10.1016/j.landusepol.2018.04.012>.
3. Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* **2020**, *236*, 111402. <https://doi.org/10.1016/j.rse.2019.111402>.
4. Jänicke, C.; Wesemeyer, M.; Chiarella, C.; Lakes, T.; Levers, C.; Meyfroidt, P.; Müller, D.; Pratzer, M.; Rufin, P. Can we estimate farm size from field size? An empirical investigation of the field size to farm size relationship. *Agric. Syst.* **2024**, *220*, 104088. <https://doi.org/10.1016/j.agry.2024.104088>.
5. Costa Freitas, M.d.B.; Teixeira, M.D.; Antunes, C.R.; Ribeiro, H.C.; Partidário, M.d.R. Beyond Cadaster: Landowners and Land Fragmentation—Insights from a Case Study. *Geographies* **2026**, *6*, 57. <https://doi.org/10.3390/geographies6020057>.
6. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. <https://doi.org/10.1038/nature14539>.
7. Freeman, E.A.; Moisen, G.G.; Coulston, J.W.; Wilson, B.T. Random forests and stochastic gradient boosting for predicting tree canopy cover: Comparing tuning processes and model performance. *Can. J. For. Res.* **2016**, *46*, 323–339. <https://doi.org/10.1139/cjfr-2014-0562>.

8. Liu, T.; Abd-Elrahman, A.; Morton, J.; Wilhelm, V.L. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GISci. Remote Sens.* **2018**, *55*, 243–264. <https://doi.org/10.1080/15481603.2018.1426091>.
9. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. <https://doi.org/10.1016/j.rse.2018.11.032>.
10. Dang, C.; Liu, Y.; Yue, H.; Qian, J.; Et., A. Autumn Crop Yield Prediction using Data-Driven Approaches:- Support Vector Machines, Random Forest, and Deep Neural Network Methods. *Can. J. Remote Sens.* **2020**. <https://doi.org/10.1080/07038992.2020.1833186>.
11. Tetteh, G.O.; Schwieder, M.; Erasmi, S.; Conrad, C.; Gocht, A. Comparison of an Optimised Multiresolution Segmentation Approach with Deep Neural Networks for Delineating Agricultural Fields from Sentinel-2 Images. *PFG J. Photogramm. Remote Sens. Geoinf. Sci.* **2023**, *91*, 295–312. <https://doi.org/10.1007/s41064-023-00247-x>.
12. Wu, F.; Maleki, R.; Oubara, A.; Gómez, D.; Eftekhari, A.; Yang, G. Machine Learning Approaches for Crop Identification from Remote Sensing Imagery: A Review. In *Proceedings of the 14th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2022)*; Abraham, A.; Hanne, T.; Gandhi, N.; Manghirmalani Mishra, P.; Bajaj, A.; Siarry, P., Eds.; Series Title: Lecture Notes in Networks and Systems; Springer Nature Switzerland: Cham, Switzerland, 2023; Volume 648, pp. 325–336. https://doi.org/10.1007/978-3-031-27524-1_31.
13. Papadopoulou, E.; Mallinis, G.; Siachalou, S.; Koutsias, N.; Thanopoulos, A.C.; Tsaklidis, G. Agricultural Land Cover Mapping through Two Deep Learning Models in the Framework of EU's CAP Activities Using Sentinel-2 Multitemporal Imagery. *Remote Sens.* **2023**, *15*, 4657. <https://doi.org/10.3390/rs15194657>.
14. Campos-Taberner, M.; García-Haro, F.J.; Martínez, B.; Sánchez-Ruiz, S.; Gilabert, M.A. A Copernicus Sentinel-1 and Sentinel-2 Classification Framework for the 2020+ European Common Agricultural Policy: A Case Study in València (Spain). *Agronomy* **2019**, *9*, 556. <https://doi.org/10.3390/agronomy9090556>.
15. Godinho, S.; Guimar, N.; Machado, R.; Lourenço, P.; Tsiligridis, T.; Ainali, K.; Pinto Correia, T. *Report on the Assessment and Characterization of Small Farms Distribution and Spatial Characteristics Obtained from SENTINEL-2 Data*; Document WP2.4, Horizon2020 Project; EU-CORDIS: Luxembourg, 2019.
16. Tsiligridis, T.A.; Ainali, K.A. Mapping of small farming plots using remote sensing and regional statistical surveys. In *Proceedings of the Remote Sensing for Agriculture, Ecosystems, and Hydrology XXVII*; SPIE: Cergy-Pontoise, France, 2025; Volume 13666, p. 241. <https://doi.org/10.1117/12.3069600>.
17. Rivera, M.; Guarín, A.; Pinto-Correia, T.; Almaas, H.; Mur, L.A.; Burns, V.; Czekaj, M.; Ellis, R.; Galli, F.; Grivins, M.; et al. Assessing the role of small farms in regional food systems in Europe: Evidence from a comparative study. *Glob. Food Secur.* **2020**, *26*, 100417. <https://doi.org/10.1016/j.gfs.2020.100417>.
18. Guimar, N.; Godinho, S.; Rivera, M.; Pinto-Correia, T.; Machado, R.; Czekaj, M.; Tyran, E.; Puchała, J. Assessing food availability: A novel approach for the quantitative estimation of the contribution of small farms in regional food systems in Europe. *Glob. Food Secur.* **2021**, *30*, 100555. <https://doi.org/10.1016/j.gfs.2021.100555>.
19. Salsa, H. *Small Farms, Small Food Businesses and Sustainable Food Security | SALSA Project | Fact Sheet | H2020*; CORDIS | European Commission, 2016–2020. Available online: <https://ec.europa.eu/project/id/677363> (accessed on 19 November 2024).
20. Tsiligridis, T.; Ainali, K. Remote sensing Big AgriData for food availability. In *Proceedings of the 2018 International Conference on Image and Video Processing, and Artificial Intelligence*; Su, R., Ed.; SPIE: Shanghai, China, 2018; p. 17. <https://doi.org/10.1117/12.2327014>.
21. Tsiligridis, T.; Ainali, K. Integrating Statistical and Earth AgriData in Small Farming Systems for Food Security. *AgriEngineering* **2025**, *7*, 54. <https://doi.org/10.3390/agriengineering7030054>.
22. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; R Core Team; et al. caret: Classification and Regression Training, 2023. Available online: <https://CRAN.R-project.org/package=caret> (accessed on 19 November 2024).
23. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. <https://doi.org/10.1023/A:1010933404324>.
24. Tatsumi, K.; Yamashiki, Y.; Canales Torres, M.A.; Taipe, C.L.R. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. *Comput. Electron. Agric.* **2015**, *115*, 171–179. <https://doi.org/10.1016/j.compag.2015.05.001>.
25. Rodríguez-Galiano, V.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sánchez, J. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>.
26. Zhong, L.; Gong, P.; Biging, G.S. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 1–13. <https://doi.org/10.1016/j.rse.2013.08.023>.
27. Liaw, A.; Wiener, M. Classification and Regression by RandomForest. *R News* **2002**, *2*, 18–22.
28. Canters, F. Evaluating the Uncertainty of Area Estimates Derived from Fuuy Land-Cover Classification. *Photogramm. Eng. Remote. Sens.* **1997**, *63*, 403–414.

29. Tsiligrirides, T. Remote sensing as a tool for agricultural statistics: A case study of area frame sampling methodology in Hellas. *Comput. Electron. Agric.* **1998**, *20*, 45–77. [https://doi.org/10.1016/S0168-1699\(98\)00011-8](https://doi.org/10.1016/S0168-1699(98)00011-8).
30. Gallego, F.J. Remote sensing and land cover area estimation. *Int. J. Remote Sens.* **2004**, *25*, 3019–3047. <https://doi.org/10.1080/01431160310001619607>.
31. Lambert, M.J.; Traoré, P.C.S.; Blaes, X.; Baret, P.; Defourny, P. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* **2018**, *216*, 647–657. <https://doi.org/10.1016/j.rse.2018.06.036>.
32. Lawson, G.; Den Hererder, M.; Ronn Anderson, K.; Kay, S.; Machine, A.; Palma, J.H.; Tomás, A.; Corbacho, J. *Availability and Use of IACS Geospatial Data in the EU*; Zenodo: Geneva, Switzerland, 2025. <https://doi.org/10.5281/zenodo.15712678>.
33. Leonhardt, H.; Wesemeyer, M.; Eder, A.; Hüttel, S.; Lakes, T.; Schaak, H.; Seifert, S.; Wolff, S. Use cases and scientific potential of land use data from the EU's Integrated Administration and Control System: A systematic mapping review. *Ecol. Indic.* **2024**, *167*, 112709. <https://doi.org/10.1016/j.ecolind.2024.112709>.
34. Jänicke, C.; Petersen, K.A.; Schmidts, P.; Müller, D.; Jepsen, M.R. Field and farm-level data on agricultural land use for the European Union. *Sci Data* **2025**, *12*, 1050. <https://doi.org/10.1038/s41597-025-05210-6>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.