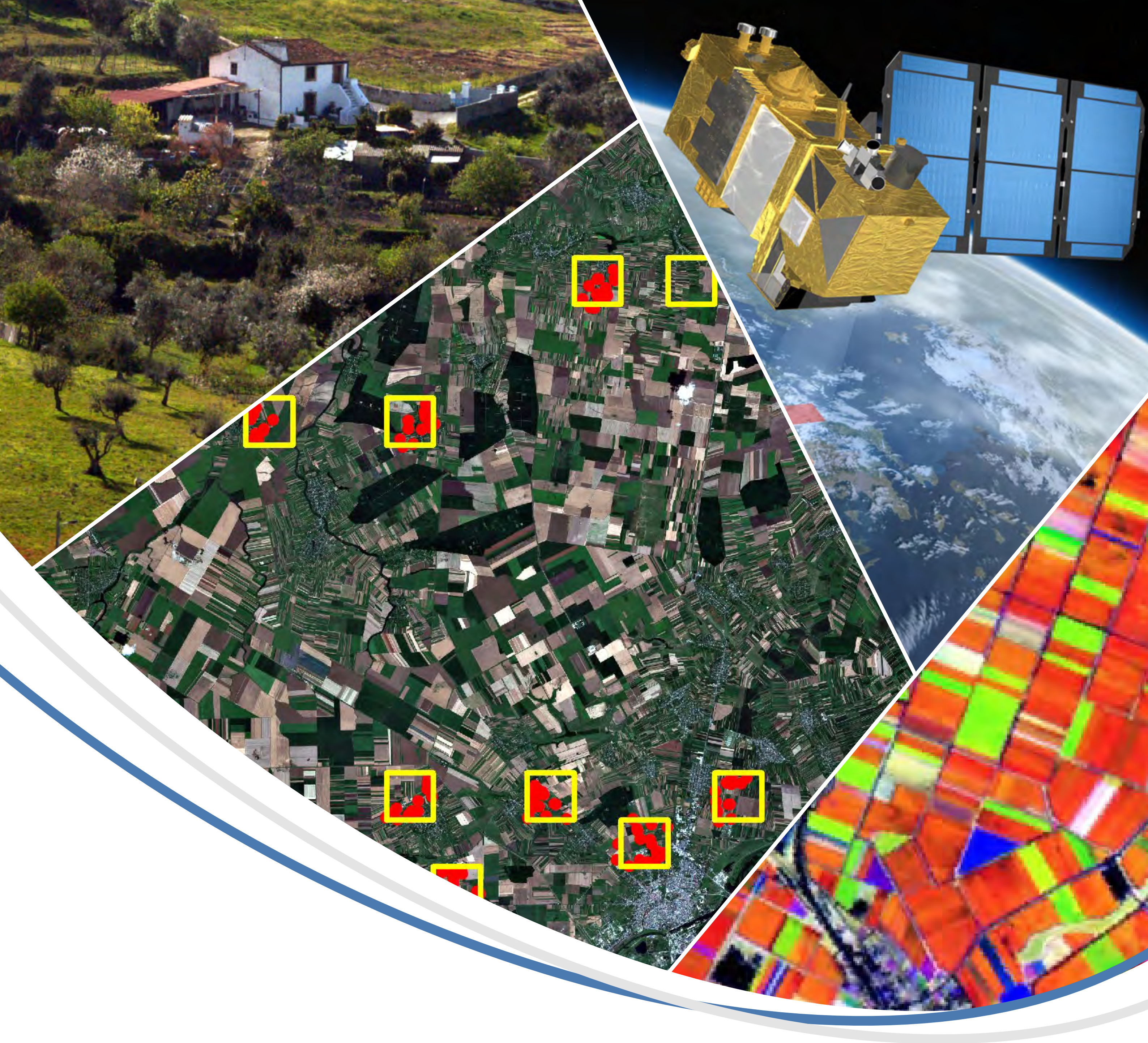


From Space to Farm: Characterization of Small Farms Using Remote Sensing Data

salsa
small farms
small food businesses and
sustainable food security

Acknowledgements

This booklet is part of the H2020 SALSA project publications that the Food and Agriculture Organization of the United Nations (FAO) has supported as co-leader of the work package on Communication and Joint Learning (WP7).

This publication would have not been possible without the contributions of Matheus Lima, Giulia Palestini, Cristiano Consolini and Nevena Alexandrova-Stefanova, from the Research and Extension Unit of the Office of Innovation (OIN) of the Food and Agriculture Organization of the United Nations (FAO), and support of Teresa Pinto-Correia and Patrícia Lourenço, from the University of Évora. Special recognition goes to Alessandro Mannocchi, FAO graphic designer, for his continuous support of this publication and SALSA's work.

Many thanks are owed to Pietro Gennari and Yakob Mudesir Seid, of the Office of the Chief Statistician of FAO, and Benjamin Koetz, of European Space Agency, for their contribution to the section Use of Earth Observation Data and the Sustainable Development Goals: Insights from FAO-ESA Partnership.

We greatly appreciate the comments that were provided by the SALSA partners Sandra Šūmane, from the Baltic Studies Centre, and Marta Czekaj and Monika Szafrańska, from the University of Agriculture in Krakow. We would also like to thank all the other SALSA partners who participated in the activities described in this booklet, as without their achievements this work would not have been possible. In particular, we would like to acknowledge that the remote sensing work of SALSA, which is the subject of this publication, was led by Sergio Godinho, Nuno Guiomar, Rui Machado, Teresa Pinto-Correia and Patricia Lourenço of the University of Évora, with the support of Theodore Tsiligiridis and Katerina Ainali of the Agricultural University of Athens.

This publication was realised with the technical support of

Food and Agriculture Organization of the United Nations

and with the collaboration of the University of Évora

salsa
small farms
small food businesses and
sustainable food security

From Space to Farm: Characterization of Small Farms Using Remote Sensing Data

Foreword

The Food and Agriculture Organization of the United Nations (FAO) believes that innovation, through which our most adventurous scientific endeavours can serve our most basic needs, is the central driving force for achieving a world free from hunger and malnutrition. FAO Director General, Mr QU Dongyu has highlighted innovation in agriculture as "a way to enhance effectiveness, competitiveness and resilience with limited land and resources".

Remote sensing applications in agriculture are one such well-recognized innovation with ever expanding potential. They are used every day – from monitoring land use for crops, to assessing the condition of vegetation and soil, to measuring yields and spotting pests and diseases. They also have agrometeorological applications, and are used for analysis of forest vegetation and agro-ecological zoning.

The greatest advantage of using remote sensing data is the ability to generate information across time and space which is essential for informed decision making. Use of remote sensing technology in agriculture involves a large amount of data, its analysis, interpretation and management. In the past three decades, much progress has been made in developing state-of-the-art analytical procedures and tools to process and interpret space-based remote sensing data and to combine it with ground-based measurements for agricultural applications. However, there is still larger need to invest in technological and human capital.

The EU-funded project on **Small farms, small food businesses and sustainable food and nutrition security (SALSA)** pioneered the development of innovative approaches, tools and procedures for using the latest satellite technologies, transdisciplinary approaches, food systems mapping and participatory foresight analysis. Most importantly, the project is replicable and scalable as it uses open source information and tools. In this booklet, we provide an interesting example of the application of analytical process for using satellite-based information for monitoring and assessment of crop types, crop area extent and crop production in small scale farming systems. We truly believe that this publication captures the extent of the opportunity for what modern technologies and methodologies can do for small scale farms worldwide to contribute to sustainable food security and nutrition.

Selvaraju Ramasamy
Head of the Research and Extension Unit
Office for Innovation

Contents

Glossary

Cluster Analysis: A cluster analysis is a statistical classification technique in which a set of objects or points with similar characteristics are grouped together in clusters. It encompasses a number of different algorithms and methods that are all used for grouping objects of similar kinds into their respective categories. The aim of cluster analysis is to organise observed data into meaningful structures to gain further insight from them.

Confusion Matrix: In the field of machine learning classification, a confusion matrix is also known as an error matrix. A confusion matrix is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known. It allows for visualizing the performance of an algorithm.

Crop Richness: In this publication, the term refers to the number of total unique crops cultivated in one same region or in the same farm size cluster.

EDORA typology: A European typology developed by EPSON that attempts to capture the aspects of rural differentiation. It is divided in three typologies: Rurality/accessibility, Economic Restructuring and Performance.

European Size Units: Abbreviated as ESU, this is a standard gross margin of EUR 1 200 that is used to express the economic size of an agricultural holding or farm. For each farm activity or 'enterprise', such as wheat production, dairy cows or the output from a vineyard, the standard gross margin (SGM) is estimated based on the area used for the particular activity (or number of heads of livestock) and a regional coefficient. The sum of all such margins derived from activities on a particular farm is its economic size, which is then expressed in European size units (this is done by dividing the total SGM by euro 1 200).

EUROSTAT: Eurostat is the statistical office of the European Union. Its mission is to provide high quality statistics for Europe. Eurostat offers a whole range of important and interesting data that governments, businesses, the education sector, journalists and the public can use for their work and daily life.

Farm Structure Survey: This is also known as the Survey on the Structure of Agricultural Holdings, which is carried out by all European Union (EU) Member States. The FSS is conducted consistently throughout the EU with a common methodology on a regular basis and therefore provides comparable and representative statistics across countries and time, at regional levels (down to NUTS 3 level). Every three or four years the FSS is carried out as a sample survey, and once in ten years as a census.

F-Score: The F-Score is used to measure the accuracy of a test and can provide a more realistic measure of a test's performance by balancing the use of precision and recall. The F-score is often used for information retrieval for measuring search, document classification, and query classification performance.

Key Crops: Key crops, also called relevant crops, are the crops that were selected in this study for each reference region based on two criteria: 1) Importance in terms of production and consumption; and 2) Largely produced in the region.

k-means: This is one of the simplest unsupervised learning algorithms that solves the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. Since it is important to pay special attention to where these centroids are placed, as different locations cause different results, it is preferable to place them as far away from each other as possible.

NUTS Level: The Nomenclature of Territorial Units for Statistics, abbreviated NUTS (from the French version Nomenclature des Unités Territoriales Statistiques) is a geographical nomenclature subdividing the economic territory of the European Union (EU) into regions at three different levels (NUTS 1, 2 and 3 respectively, moving from larger to smaller territorial units).

Official Statistics: This term refers to the data collected from EUROSTAT and Agricultural Censuses from the reference regions of the study.

Overall Accuracy: The overall accuracy is calculated by summing the number of correctly classified values and dividing by the total number of values. The correctly classified values are located along the upper-left to lower-right diagonal of the confusion matrix. The total number of values is the number of values in either the truth or predicted-value arrays.

Producer's Accuracy: This is the accuracy of the map from the map maker's point of view (the producer). This shows how real features on the ground are correctly indicated on the classified map or what probability there is that a certain land cover of an area on the ground is classified as such.

p-value: This is the probability of obtaining the observed results of a test, assuming that the null hypothesis is correct. It is the level of marginal significance within a statistical hypothesis test representing the probability of the occurrence of a given event. The p-value is used as an alternative to rejection points to provide the smallest level of significance at which the null hypothesis would be rejected. A smaller p-value means that there is stronger evidence in favour of the alternative hypothesis.

Random Forest: Random Forest is a machine-learning classifier and regression that consists of a large number of individual decision trees that operate as an ensemble that will be used to make automatic classification. Each decision tree node uses a subset of attributes that are randomly selected from the whole original set of attributes.

randomForest package: This is a complied code and sample data used for applying the Random Forest algorithm using the R software. It is stored under a directory called "library" in the R environment which was developed by Breiman and Cutler in 2018.

Reference Regions: The European regions that were selected as representative of Small Farms inside the SALSA project countries. The regions were selected based on the combination of size and economic indicators, EDORA typologies and expert consultations.

RGB: RGB stands for "Red Green Blue." It refers to three hues of light of the visible spectrum that can be mixed to create different colours (true or false) in an image. Combining red, green, and blue light is the standard method of producing colour images on screens such as TVs, computer monitors and smartphone screens.

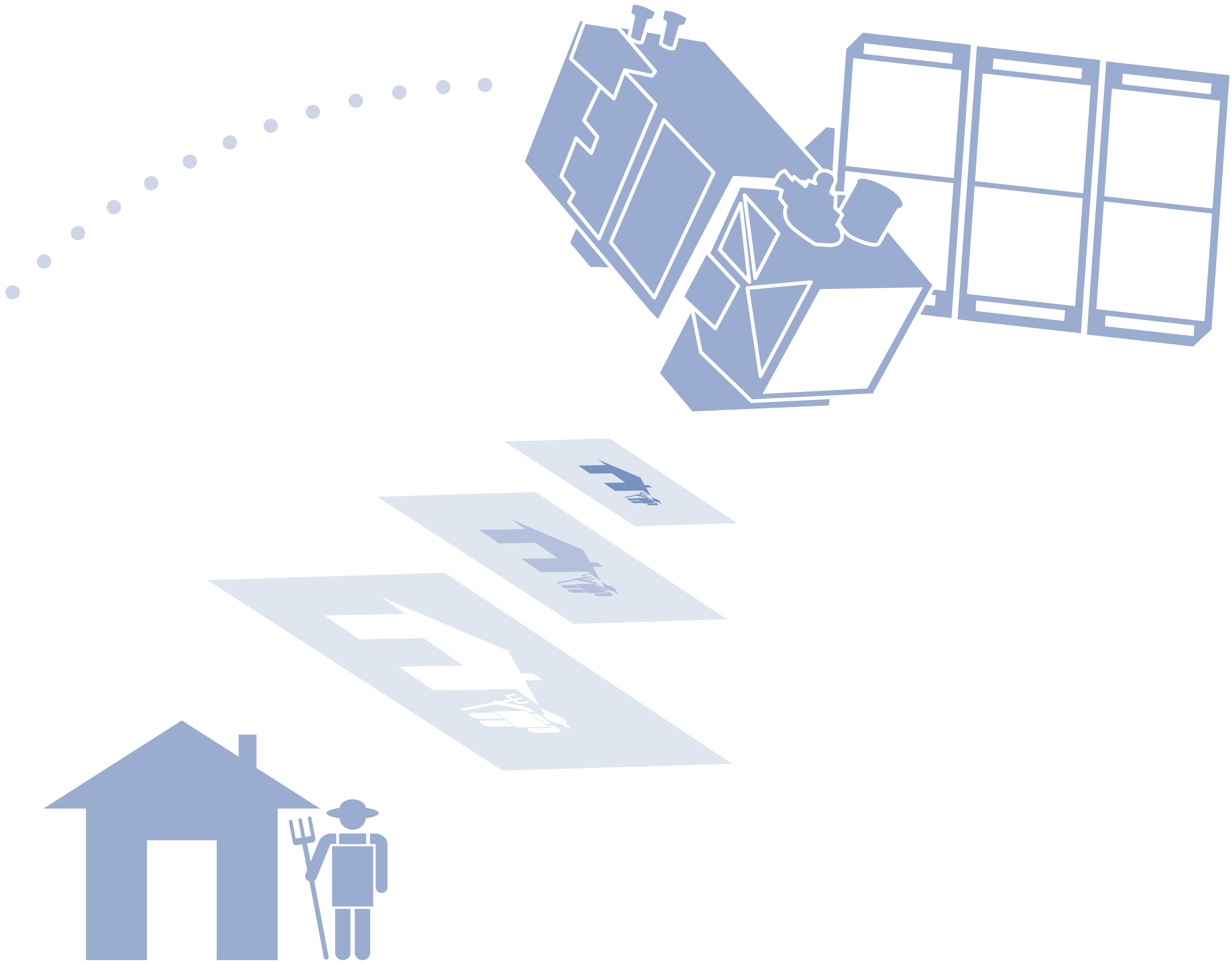
Sentinel 1-SAR: Sentinel-1 is a Synthetic Aperture Radar (SAR) mission that provides continuous all weather, day-and-night imagery at C-band (centre frequency: 5.405 GHz), operating in four exclusive imaging modes with different spatial resolutions and coverages. SAR imaging is used for monitoring sea-ice zones and the polar environment; mapping in support of humanitarian aid in crisis situations; surveilling marine environments; monitoring land surface motion risks; and mapping land surfaces: forests, water, soil and agriculture.

Sentinel-2: SENTINEL-2 is a European wide-swath, high-resolution, multispectral imaging mission. SENTINEL-2 carries an optical instrument with 13 spectral bands: four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution. Multispectral imaging can be used for land cover; land usage and land-use-change detection maps; geophysical variable maps (leaf chlorophyll content, leaf water content and leaf area index); risk mapping; and fast images for disaster relief efforts.

Spectral Signature: Different surface types such as water, bare ground and vegetation types reflect radiation differently in various channels. The radiation reflected as a function of the wavelength is called the spectral signature of the surface. The spectral signature is specific to each type of surface like a fingerprint.

Standard Gross Margin: The standard gross margin, abbreviated as SGM, is a measure of the production or the business size of an agricultural holding. It is based on the separate activities or 'enterprises' of a farm and their relative contribution to overall revenue. For each separate activity (for instance wheat, dairy cows or a vineyard), an SGM is estimated based on the area (for crop output) or the number of heads of livestock (for animal output) and a standardised SGM coefficient for each type of crop and livestock, calculated separately for different geographical areas to allow for differences in profit.

Sustainable Development Goals: The 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015, provides a shared blueprint for peace



and prosperity for people and the planet, now and into the future. At its heart are the 17 Sustainable Development Goals (SDGs), which are an urgent call for action by all countries - developed and developing - in a global partnership. They recognize that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality and spur economic growth – all while tackling climate change and working to preserve our oceans and forests.

Training Data: Training Data is a sizable body of known data used for the training set. The training set is the data set on which the machine learning model is built. The training set is usually collected manually, and the classification model follows the exact same rules and definitions given in the training set.

User's Accuracy: The accuracy from the point of view of a map user, not the map maker. The User's Accuracy essentially tells the user how often the class on the map will be present on the ground. This is referred to as reliability.

Utilized Agricultural Area: Utilized Agricultural Area, abbreviated as UAA, is the total area taken up by arable land, permanent grassland, permanent crops and kitchen gardens used by the holding, regardless of the type of tenure or whether it is used as a part of common land. It excludes: mushrooms, unutilised agricultural land (NUAA), woodland (WA), other land, occupied by buildings, farmyards, tracks, ponds, etc.

Vegetation Indices: A vegetation index is a single value that quantifies vegetation biomass and/or plant vigour. The vegetation index measures chlorophyll absorption through combinations of the red portion of the spectrum relative to reflectance or radiance in the near infrared. It is calculated based on two or more bands (e.g. red and near infrared bands) to improve the contribution of vegetation properties in an image through remote sensing.

Executive Summary

This publication describes the analytical process carried out under the European Union-funded SALSA project which enabled the development of a European map of the distribution of small farms at the NUTS-3 level and assessed the capabilities and usefulness of Copernicus Sentinel-1 and Sentinel-2 satellites for small farms monitoring, specifically in providing information about crop types, crop area extent and crop production.

The European map of small farms distribution was developed through a stepwise approach that combined diverse datasets and information gathered from key experts. The criteria used to classify small farms was defined by the SALSA partners and Expert panels, based on physical size (farms with less than 5 hectares of Utilized Agricultural Area) and economic size (farms with less than 8 Economic Size Units, ESU) of the Standard Gross Margin (SGM).

These indicators were then analysed using k-means clustering technique and combined with the European Development Opportunities for Rural Areas (EDORA) typologies. Experts evaluated the results, and one map was selected as the most meaningful representation of the small farms typology across Europe.

The capabilities and usefulness of Copernicus Sentinel-1 and Sentinel-2 satellites as a method for small farms monitoring was assessed to objectively quantify the importance of small farms in terms of crop production.

By considering a gradient of 21 Reference Regions (NUTS-3 level) distributed over eleven European countries (Bulgaria, Czech Republic, France, Greece, Italy, Latvia, Lithuania, Poland, Portugal, Romania, and Spain) and one African country (Tunisia), Sentinel images were tested in very differently structured farm landscapes, which allowed for a better understanding of the accuracy and effectiveness of these satellites for small farms assessment.

Field data was collected *in situ* for calibration of the image classification model and validated across the reference regions.

Approximately 500 points were used to collect crop information in each reference region. Farmer surveys were conducted to acquire reliable data regarding the productivity of key crops in each region. With this information the crop type diversity was assessed and compared among different farm sizes. The results confirmed the importance of small farms for the conservation and maintenance of agrobiodiversity.

The crop classification method using satellite imagery presented different levels of efficacy over the regions and among the crops. Among the analyzed crops, cereals showed to be more reliably mapped when applying the described methodology, followed by meadows, pastures and forage crops (MPFC), vineyards and orchards. The least well-classified class of crops in this study was that of the vegetables which comprised a high diversity of crops (e.g. carrots, legumes, onions and potatoes).

This study also showed the consistency of Sentinel-derived crop areas against documented crop areas from official statistics at the regional level. The relation between crop areas from both data sources (official statistics and Sentinel data) shows a significant and very high correlation with an R^2 value of 0.96 ($p<0.001$), demonstrating that there is no significant difference between the Sentinel-based crop area and the official regional statistics.

In summary, this report clearly showed that Sentinel-1 and Sentinel-2 missions open a new era of opportunities towards the development of more robust tools and methodologies based on remote sensing data to accurately assess small scale farming systems.

From Space to Farm: Characterization of Small Farms Using Remote Sensing Data

Introduction

It is widely recognized that small farms play a relevant role in terms of food production and food security, as well as holding great significance in providing important social, cultural and environmental services in rural areas.

The importance of small farmers and their role in food security has been highlighted on many occasions in international fora, such as the Committee on World Food Security (CFS), which is the foremost inclusive international and intergovernmental platform aiming to ensure food security and nutrition for all.

The **Sustainable Development Goals (SDGs)** address the global importance of small-scale farming systems in the first instance under *Goal 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture; and Target 2.3: By 2030, double the agricultural productivity and incomes of small-scale food producers*. However, the extent and distribution of crop types, crop production, and its spatial distribution in a small-scale farming context still remain uncertain or unknown, as a large extent of small farms tends to be excluded from the official statistical surveys – especially in developing countries.

Therefore, considering the importance of small farms, and the ambition of SDGs Goal 2, policymakers need tools that can provide accurate and timely information on the crop area extent, types and yield estimates to objectively quantify the crop production capabilities of small farms.

The **project "Small Farms, Small Food Businesses and Sustainable Food Security" (SALSA)** was developed taking this context into account. SALSA aimed to provide a better understanding of the role of small farms and food businesses in sustainable food and nutrition security. Supported by the European Union's Horizon 2020 program, SALSA consisted of a coalition of 16 European and African partners (including universities, research institutions and farmers' organizations) which collaborated between 2016 and 2020. SALSA was implemented through an integrated multi-method approach in 30 regions in Europe and Africa, using the most recent satellite technologies, transdisciplinary approaches, food systems mapping and participatory foresight analysis to produce evidence-based policy recommendations.

As small farms production patterns were among the key concepts to be addressed, one of the main goals of SALSA was to estimate the current and potential production of small-scale farming systems at a regional level. An important first step in the analysis of SALSA was to develop and test methods and tools that can produce accurate and useful information about small farms.

Definitions of small farms involving only the criterion of farm size have universal appeal, as they are relatively easy to apply and allow for simple comparisons across countries and world regions. However, they do not capture all the complexities of farming.

In October 2016, the Food and Agriculture Organization of the United Nations (FAO), a partner in SALSA, hosted a first email conference titled "Exploring the contribution of small farms to achieving food security and improved nutrition"[1], focusing the attention of researchers, educators and a wide spectrum of food chain/food system actors and entrepreneurs, as well as policymakers and administrators on multiple levels, on the role of small farms within a larger context of food security.

1 FAO summary document of the 1st FAO SALSA email conference. Available at http://www.fao.org/3/a-bq691e.pdf

The aim was to take a fresh look at the contribution of small farms to food security and nutrition, allowing stakeholders worldwide to share their experiences and up-to-date knowledge regarding this important topic.

In April 2018, a second FAO-SALSA email conference was held with the title: "Small Farms, Small Food Businesses and Sustainable Food Security"[2]. The conference covered six overarching topics. The results provided further feedback on the work in the SALSA project through shared examples while also identifying key knowledge gaps that contribute to building the SALSA empirical base.

Among other themes, the first FAO-SALSA email conference provided guidelines for defining small-farms in regional contexts and research projects.

Definitions involving additional criteria to farm size were considered more meaningful, particularly those including indicators of the farms' economic output, but data availability is often a limitation. At the same time there is an increasing need to better understand the relative importance and role of small farms in different regional contexts in order to support the design of public interventions. This is crucial for many regions in Europe, where changes in the farm sector are occurring at an exceptionally fast pace.

The main objective of this booklet is to describe the analytical process which enabled the development of a **European map of the distribution of small farms at the NUTS-3 level** and the capabilities and usefulness of **Copernicus Sentinel-2 satellite as a data-based method for small farms assessment**, specifically in providing information on the distribution (location) of small farms, crop types (crop diversity), crop area extent (crop acreage) and yield estimates (crop production).

This process considered different small farm configurations through the combination and analysis of different datasets with information obtained from key experts.

Box 1 – The SALSA Project - Small Farms, Small Food Businesses and Sustainable Food and Nutrition Security

The project "Small Farms, Small Food Businesses and Sustainable Food and Nutrition Security" aims to provide a better understanding of the current and potential contribution of small farms and food businesses to sustainable Food and Nutrition Security (FNS). It is supported by the EU Horizon 2020 program, a coalition of 16 European and African partners that are collaborating to assess the role of small farms and small food businesses in delivering a sustainable and secure supply of affordable, nutritious and culturally adequate food.

The four-year **SALSA** project began in April 2016. The partners have adopted a novel, transdisciplinary, multi-scale approach across 30 regions in Europe and Africa that builds on and connects relevant theoretical and analytic frameworks within a food system approach. Using this perspective, the project is looking beyond production capacity, and investigating food security in terms of the availability of nutritious and safe food, food access and control (including affordability), food utilisation and food stability.

SALSA is paying particular attention to effectively foster stakeholder involvement, knowledge exchange and joint learning at local, regional, national and international levels.

SALSA is unravelling the complex interrelationships between small farms, small food businesses and FNS. It is also shedding light on the role played by small farms in (a) the balance between the different dimensions of sustainability, (b) maintaining more diverse production systems, (c) supporting the urban/rural balance in terms of labour and (d) facilitating territorial development in countries facing strong rural population growth.

Since its beginning, **SALSA** has convened two workshops at international conferences (XXVII European Society for Rural Sociology Congress and the 2017 Global Food Security Conference) to raise awareness and solicit input into the project's research.

Regional workshops in many of the reference regions are building **SALSA** communities of practice (CoP).

SALSA effectively engages with stakeholders and decision-makers regarding small farms and food and nutrition security, and facilitates a dialogue that cuts across classical boundaries in research, policy and practice.

As a place of meeting and knowledge exchange between different stakeholders, the CoPs help to integrate SALSA work with existing policy and practice discussions, and identify and support new market and policy arrangements.

2 Background document to the FAO e-conference on "The Role of Small Farms Within a Larger Context of Food Security" http://www.fao.org/3/BU493en/bu493en.pdf

Subjects, Materials and Methods

European Distribution of Small Farms

The **first stage** of the analytical process involved the elaboration of a European map showing the distribution of small farms at the NUTS-3 level.

This process was developed through a stepwise approach that combined diverse datasets and information gathered from key experts. The criteria that were used drew on the thresholds defined in the project's Conceptual Framework that classifies small farms in two ways: by physical size (farms with less than 5 hectares of Utilized Agricultural Area) and by economic size (farms with fewer than 8 Economic Size Units, ESU) of the Standard Gross Margin (SGM).

National Experts of the regions of interest and SALSA partners comprised the expert group that provided key information during the initial and final phases of the first stage of the analytical process.

The stepwise process is simplified in Figure 1

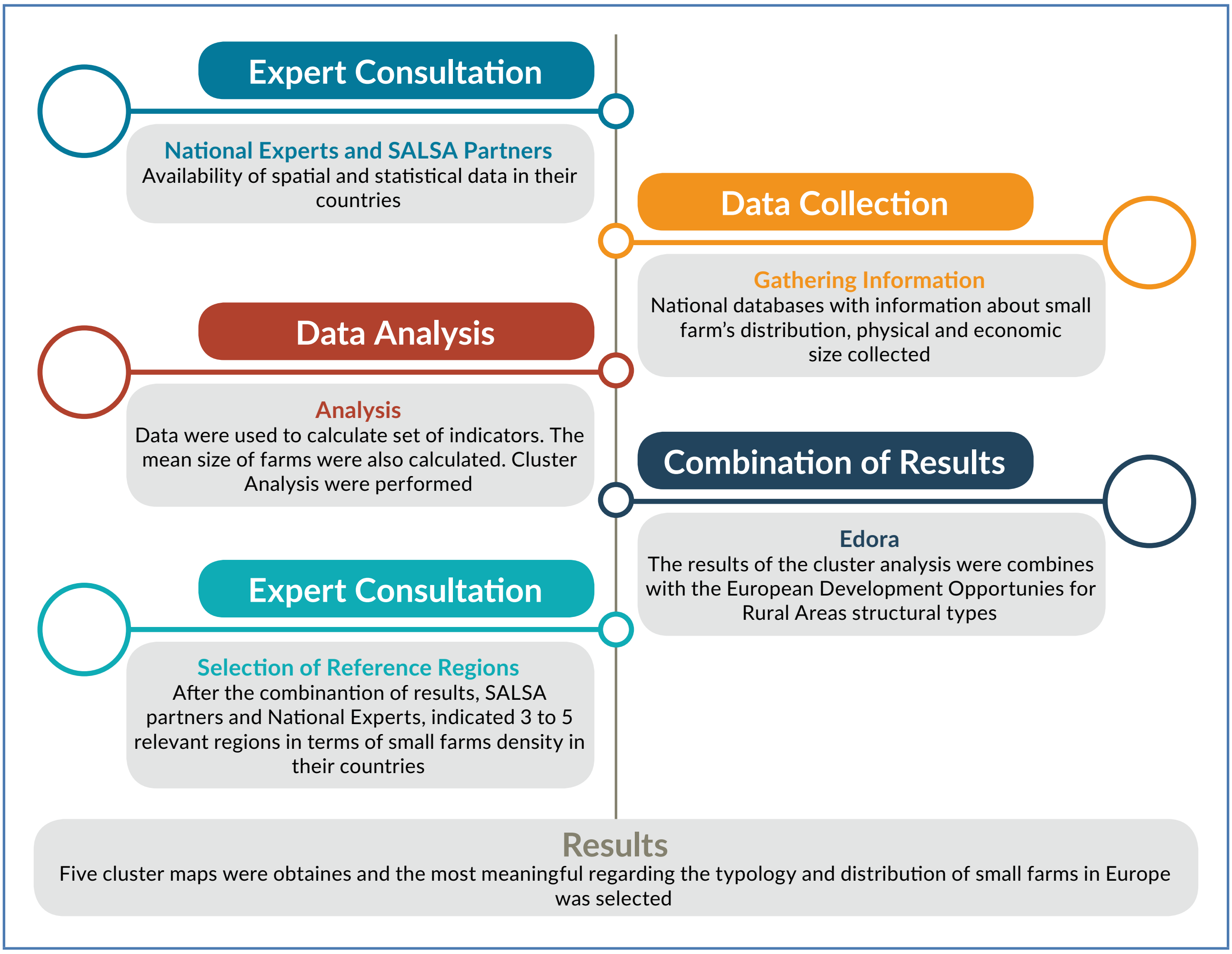


Figure 1 – Flowchart describing the stepwise process
to identify and characterize small farms across Europe

During the initial phase, the group of experts provided key information regarding the availability of datasets with spatial distribution information and statistical data of small farms in the national/ regional databases. This information was gathered, and statistical analysis was carried out.

The main sources of data for the initial phase were **EUROSTAT**, the latest National **Agricultural Censuses** and the **European Farm Structure Survey (FSS)**.

Data analysis

According to the definition of small farms agreed upon in the SALSA Conceptual Framework and expert panels, the occupational and economic variables were selected, and maps were created, based on the NUTS-3 georeferenced data.

The selected variables (indicators) used to identify areas with a high density of small farms were:

- density of farm units (number/ha) with less than 5 hectares;
- percentage of the region occupied by Utilized Agricultural Area (UAA);
- percentage of the UAA occupied by small farms defined as farms with less than 5 ha of UAA;
- percentage of farm units (number) with less than 8 ESU/SGM; and
- percentage of the region occupied by small farms defined as farms with less than 8 ESU/SGM.

The utilised agricultural area by farm size class, as well as the economic and physical size of the rural properties, were collected across Europe with high resolution at NUTS-3 level. According to the SALSA partners and Expert panels, these variables reflect the main characteristics of the spatial distribution of small farms in Europe (Figure 2).

Figure 2 – Maps with the selected indicators for spatial distribution of small farms across Europe at NUTS-3 level (except for Germany, at NUTS-2 level)

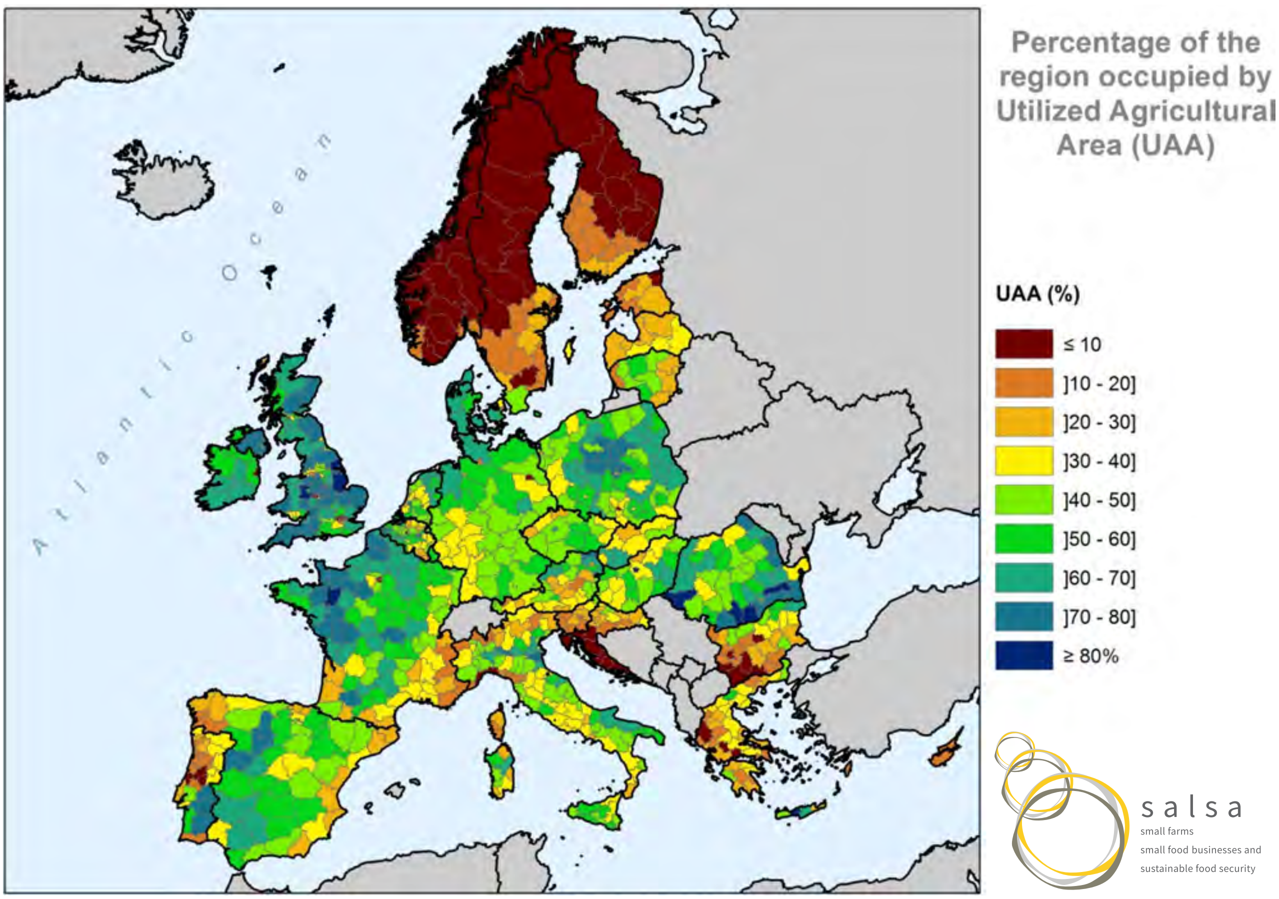
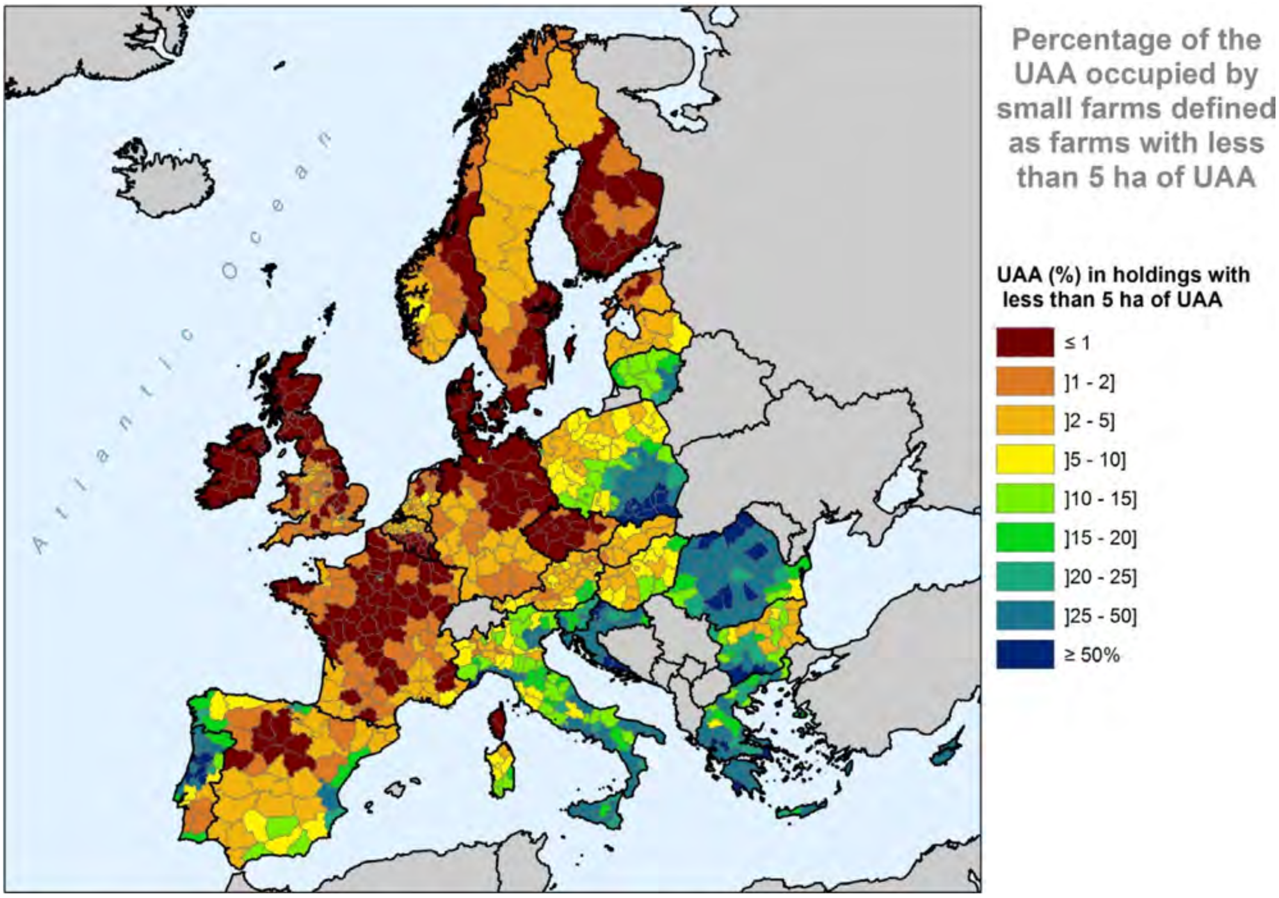
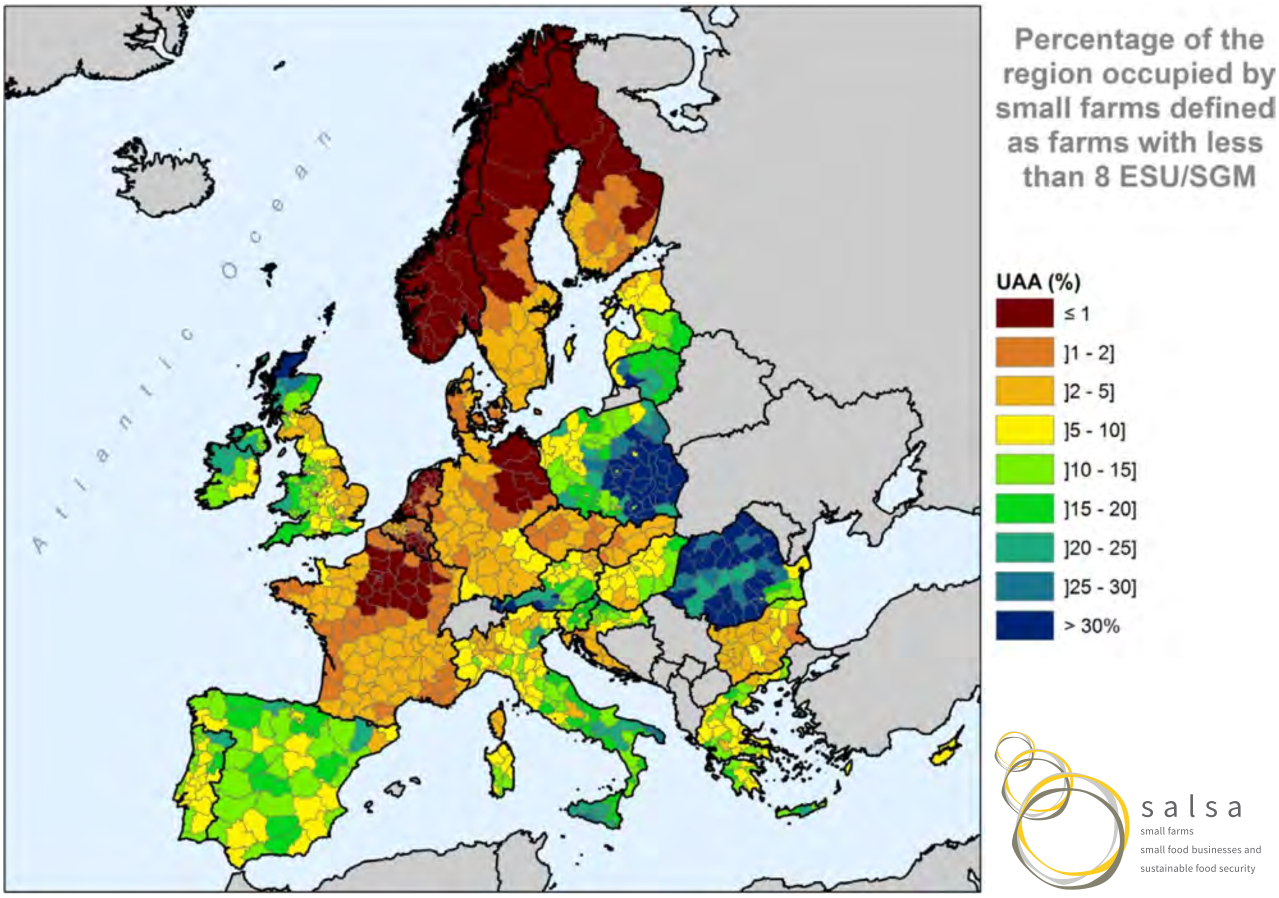


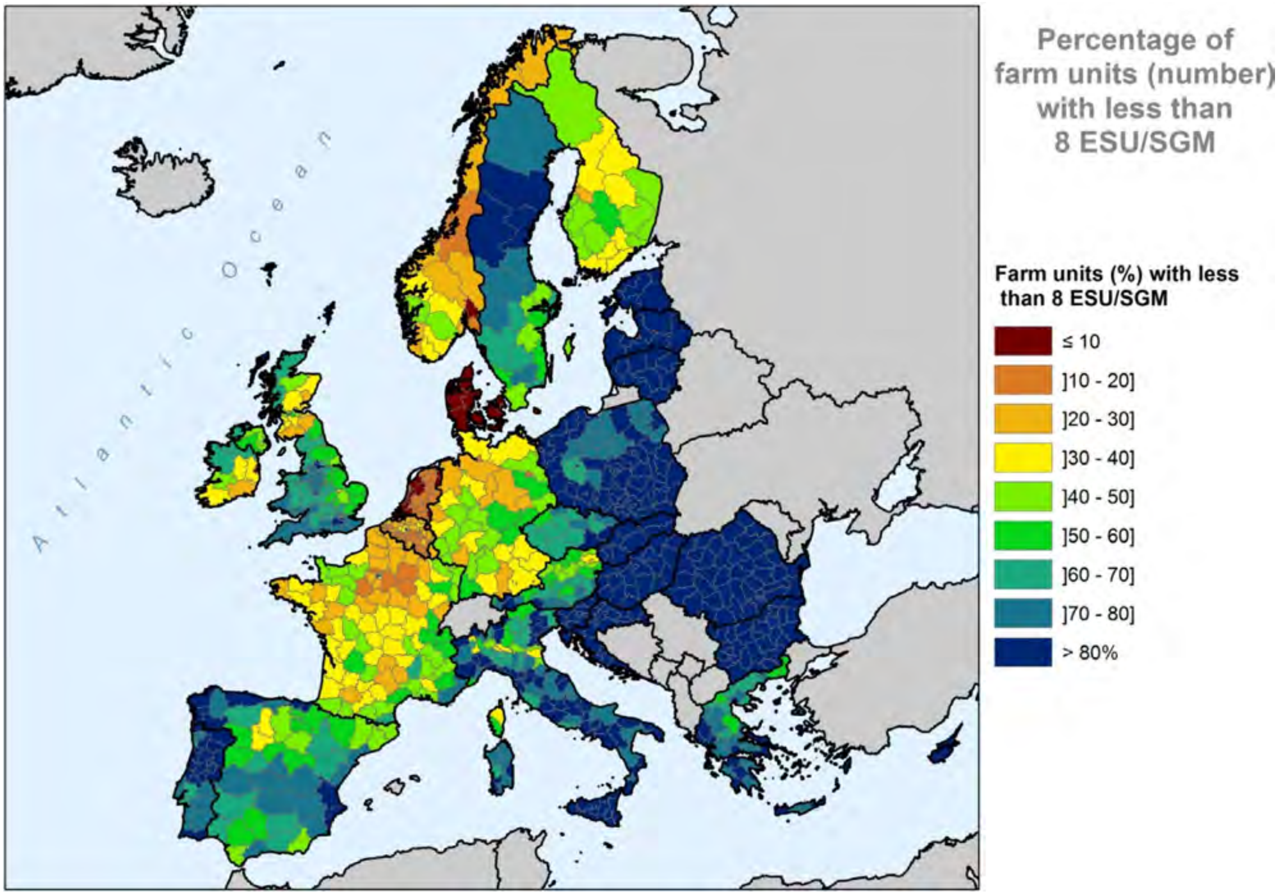
Figure 2 – Maps with the selected indicators for spatial distribution of small farms across Europe at NUTS-3 level (except for Germany, at NUTS-2 level) (cont.)

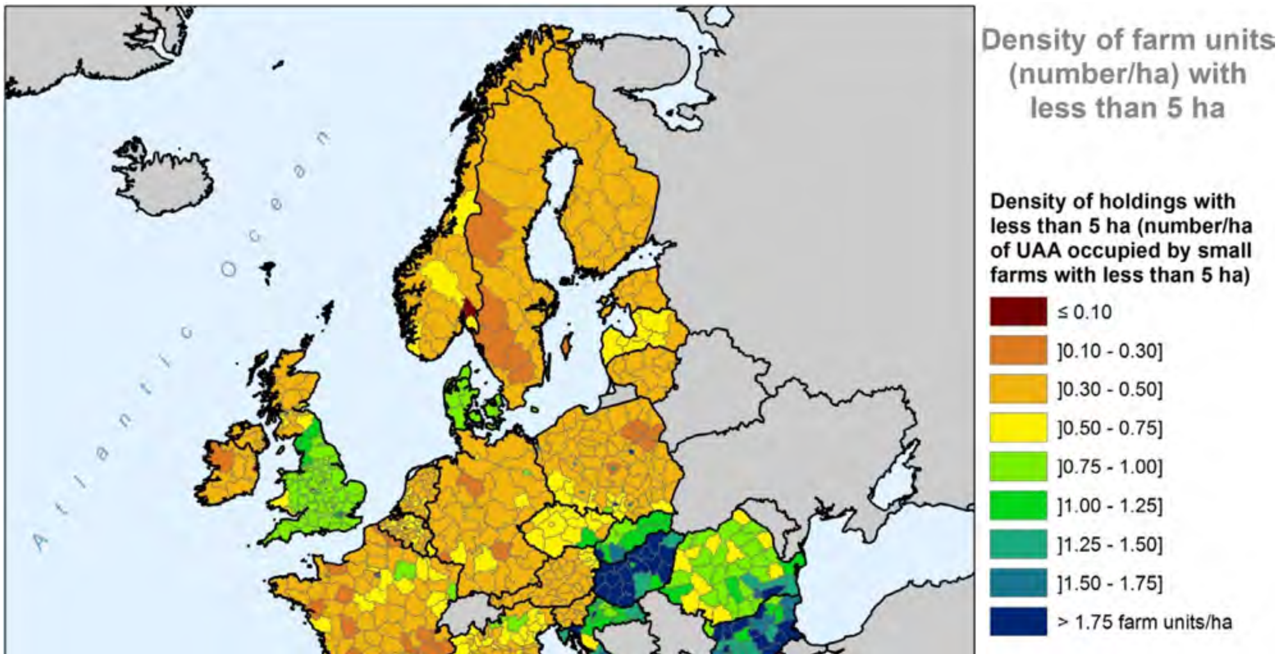




salsa
small farms
small food businesses and
sustainable food security

Figure 2 – Maps with the selected indicators for spatial distribution of small farms across Europe at NUTS-3 level (except for Germany, at NUTS-2 level) (cont.)





salsa
small farms
small food businesses and
sustainable food security

The spatial distribution and classes of the selected indicators are not homogenous across the European region and cannot reflect singly the complexity of small farms. In order to combine all spatial information of the farms' variables and incorporate it in one single map that reflects the real distribution of the small-scale farms in Europe**,** a **cluster analysis** was performed using k-means.

Due to the complexity of some of the indicators, it was decided to use only this small set of non-correlated variables and reduce the number of clusters (between 4 and 6), in order to improve the comprehensibility of the results.

After the cluster analysis, a final expert consultation with SALSA partners was carried out to identify the most reliable map for the distribution of small farms across Europe. The map based on the cluster with five variables was identified as the one that better displays the distribution and characterization of small farms.

This map grouped small farms into three categories that were defined according to the characteristics of the groupings retrieved by the k means: Predominantly agricultural regions; Regions with a balanced distribution between agriculture and other land uses; Regions with little agricultural land surface; and in six sub-categories, based on farm size and income.

The final map information was then combined with the European Development Opportunities for Rural Areas (EDORA) typologies, and based on this information, the experts of each country selected 25 specific and representative regions of the small farms in 13 countries to perform more in-depth studies (Figure 3).

Figure 3. Methodological process implemented for mapping the distribution and characterization of small farms in Europe and selecting Reference Regions

Results

The cluster analysis provided a meaningful map regarding the typology and distribution of small farms across Europe. Six regions with similar characteristics were grouped and classified as follows (see also Figure 4):

a) Predominantly agricultural regions

- Cluster 1. An extremely high number of small farms with very low incomes. These are the core regions for small farms in Europe. They are predominantly farming regions (where the UAA = >50 percent), of which a very significant part of the farming area (20 percent or more) is occupied by small farms (defined by area) and almost all farm units (more than 90 percent of all farms) are small farms in terms of economic size. These farms are mostly extremely small, on average the smallest small farms in Europe, both in surface area and economic size. These regions have few larger farms.

- Cluster 2. Regions with few, relatively small farms, that have medium incomes. These are predominantly large-scale farming regions, where small farms (both in terms of area and economic size) occupy just a small part of the farming area. The large farms occupy most of the agricultural land and are probably responsible for the majority of the production.

- Cluster 3. Regions with a low proportion of small farms that are close to the upper size threshold and have high incomes. These are large-scale, specialized and market-oriented farming regions. The proportion of small farms is very low, and those that do exist are on average larger than in other regions of Europe, with higher incomes.

b) Regions with a balanced distribution between agriculture and other land uses

- Cluster 4. Regions with a low proportion of small farms, which are relatively small and have low incomes. These are regions where farming occupies only a small part of the territory (mean value is close to 35 percent), but where 70 percent of the farm units are small farms, in economic terms, with very low incomes. Small farming is thus relatively important, although the farmers are mostly poor. These regions also have other types of farms.

c) Regions with little agricultural land surface

- Cluster 5. Small farms exist in large numbers which are extremely small and have low incomes. These regions are either dominated by forestry or are primarily urban. The large majority (80 percent) of farm units in these regions are extremely small in area (average < 2 ha) and economic size.

- Cluster 6. Small parts of the region are occupied by small farms, which are close to the upper size threshold and have medium incomes. These are mostly regions dominated by forests, which contain the lowest proportion of agricultural land in Europe. However, almost half of the existing farm units are small in economic terms, yet still relatively important in the farming landscape.

Merging the cluster information with the EDORA typologies, the national experts and SALSA partners were asked to indicate 3 to 5 relevant regions within their countries, considering the diversity and distribution of small farms. The regions indicated by both a) the group of National experts and b) by SALSA partners were selected as Reference Regions. The 21 selected Reference Regions in Europe are shown in Figure 5.

In conclusion, it is rather difficult to capture the diversity, heterogeneity and/or variability in each country in a European-wide analysis with the same (and limited) dataset. The analysis is particularly complex in regions where extreme situations occur. For example, Romania can be considered a hotspot of small farms, both in terms of structural and economic size. However, at European scale and with a simplistic and small set of variables, it is impossible to capture the Romanian asymmetries in terms of small farms distribution since they show, in this context, lower spatial variability. If we increase the number of variables, or number of clusters, we lose readability, whereas with a low number of clusters or a reduced number of variables we are not able to capture this variability.

One of the major constraints in the development of an analysis on this scale and with this particular complexity is related to the quality of the data and particular elements that are considered relevant drivers of farm size in each country (e.g. biophysical constraints). The use of different size thresholds for including farms in national agricultural censuses is also a major limitation, constraining a European-level analysis. Davidova et al. (2013) highlight the lack of adequate data concerns in specific subsistence farming. Subsistence farms tend to be excluded from official statistical surveys as they fall below size thresholds for data collection. Finally, continuous changes in administrative boundaries (e.g. Poland) constrain the spatial integration of statistical data.

Yet the SALSA approach allows for:

- differentiating the Scottish uplands (croft regions) from the lowlands or the sheep farming systems of the north-western Ireland, which follow a tradition of multiple activities;
- identifying the three main farm structures characterized by Kostrowicki (1970) in Poland;
- separating the plots subjected to intense afforestation in Portugal, but where the small agricultural plots remain;
- differentiating the mountain or Apennine areas in Italy; and
- distinguishing the Southern and Alpine areas from the rest of the regions in France.

Moreover, the cluster map only represents the farm structure in terms of its structural and economic sizes. It is not possible to infer similarities or dissimilarities between regions concerning farm types or land use intensity. In view of the above-mentioned limitations, and considering the tremendous importance of small farms in many European regions, one of the SALSA project recommendations is that more efforts should be made to:

- improve official statistical data on small farms (e.g. harmonise minimum farm size thresholds);
- further elaborate the typologies of small farms presented in this report; and
- further develop methods and technologies based on remote sensing for assessing the spatial distribution of small farms.

The high variability in the spatial distribution and types of small farms means that better information is needed to support policy development.

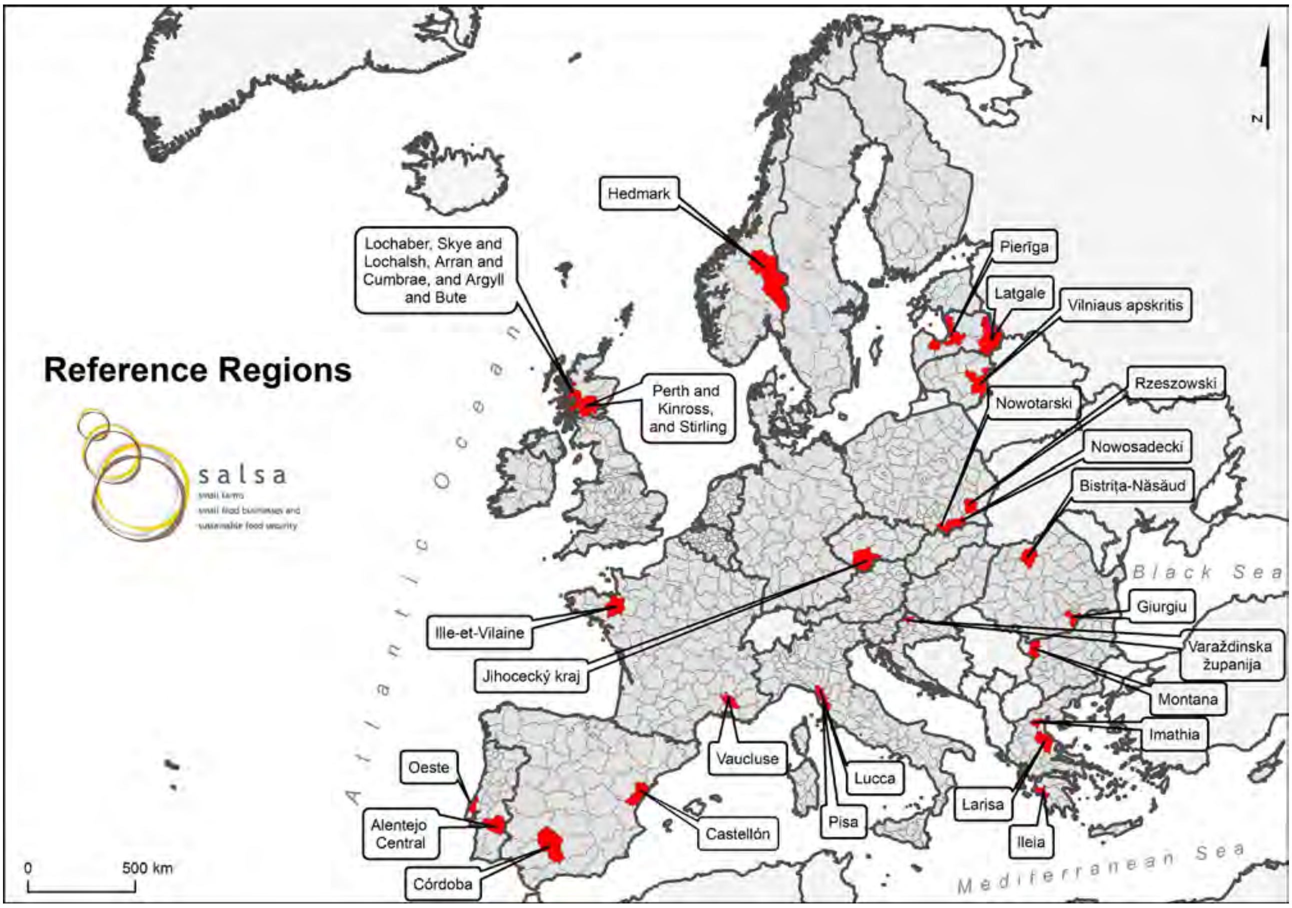


Figure 4 – Typology and distribution of small farms in Europe

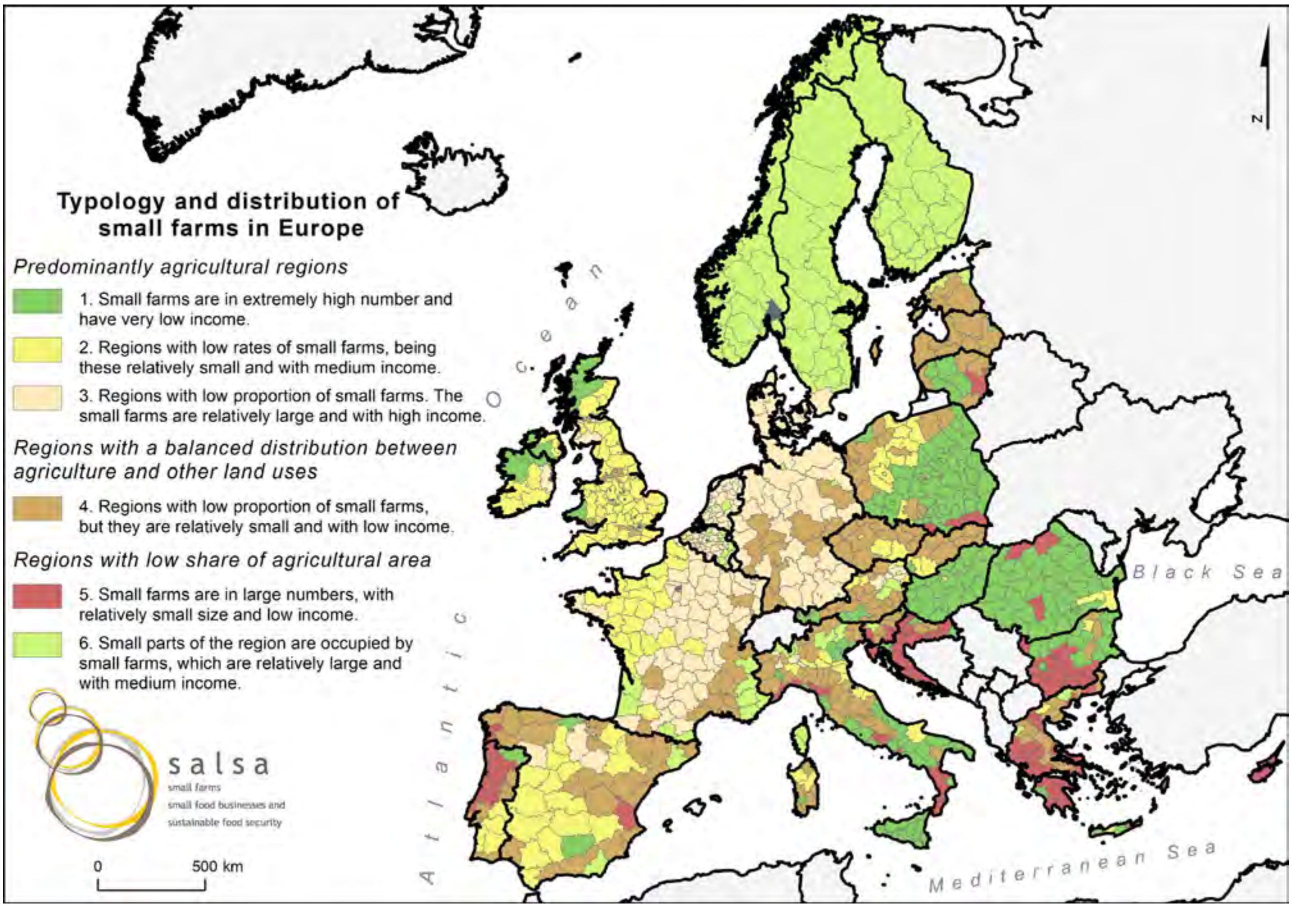


Figure 5 – Geographic distribution of the Reference Regions within countries

Spatial Characteristics and Assessment of Small Farms Distribution Based on Sentinel-2 Data

This **stage** of the analysis focused on the development and testing of remote sensing methods based on Sentinel-2 images to generate useful information about small farms in very diverse landscapes and crops, thus allowing for a better understanding of the accuracy and effectiveness of this tool for assessing and monitoring small farms in Europe and Africa with validated methodological guidelines.

The framework that was developed had the objective to establish remote sensing protocols and methods to produce maps that reflect the crop types cultivated by the small farmers across 21 Reference Regions in 12 countries across Europe and Africa (Table 1).

This data was then combined with information obtained through a Farmer's Survey to estimate the amount of food produced by small farms as well as the potential of smallholders' production increase in the Reference Regions. The information about yield gap is crucial to understanding the potential contribution of small farms for food security (Beddow *et al*. 2014).

Mapping Crop Types

In order to make future investments, policies and logistical decisions that address food security, accurate and reliable information about the location and distribution of croplands and crop types is needed (Fritz 2013).

Country	Reference region
Bulgaria	Montana
Czech Republic	Jihočeský kraj
France	Vaucluse
Greece	Imathia
	Larisa
	Ileia
Italy	Lucca
	Pisa
Latvia	Latgale
	Pierigia
Lithuania	Vilniaus Apskritis
Poland	Rzeszowski
	Nowosadecki
	Nowotarski
Portugal	Alentejo Central
	Oeste
Romania	Bistrița-Năsăud
	Girgiu
Spain	Castellón
	Córdoba
Tunisia	Haouaria

Table 1. SALSA Reference Regions, plus Tunisia (Haouaria), under analysis to obtain crop field data and satellite images[3] are reported in Figure 5

3 For more details on why other SALSA Reference Regions were excluded from the analysis, see the SALSA Report: *Deliverable 2.4. Report on the assessment and characterization of small farms distribution and spatial characteristics obtained from SENTINEL-2 data*, page 6.

A methodological approach based on three main steps was implemented under SALSA for obtaining crop type maps in the Reference Regions.

The initial step was the collection of reference data in situ (field work) to create and calibrate the classification method. The final step consisted in the image analysis and classification applying machine learning techniques.

A summary of each step is described below.

Collection and Quality Control of Reference Crop Data in Each Region

The first step of the methodology consisted in intense fieldwork for crop type data collection.

The field data collection (training data) must be representative of the output classes and their subclasses and must have a sufficient number of samples to allow for pattern recognition to occur (Muchoney and Strahler 2002).

The dataset obtained during the fieldwork was used for calibration and validation of the **image classification procedure.**

Due to the high costs of data collection at field level, a methodological approach was developed to select a minimum representative number of points for each region. Details about the methodology are described in the **SALSA Deliverable 2.3** (Godinho *et al*. 2019).

Approximately 500 points were used to collect crop information in each reference region. The sampling method used 2 km x 2 km squares with an average of 20 sampling points/square. More than 12 230 points were collected and checked by the different SALSA teams across the 21 Reference Regions.

The crop type (training data) of each field point was verified, photographed and numbered, and the geographical coordinates were registered.

In order to ensure the standardization of the data collection process, and therefore the quality of the field data, a document summarizing the guidelines for field crop data collection was elaborated and distributed to all teams working on the field. A training session was also organised to provide practical instructions on how to carry out the survey in the field (work with GPS etc.).

Figure 6 shows one example of field sampling. Figure 7 shows examples of the crops used in the study.

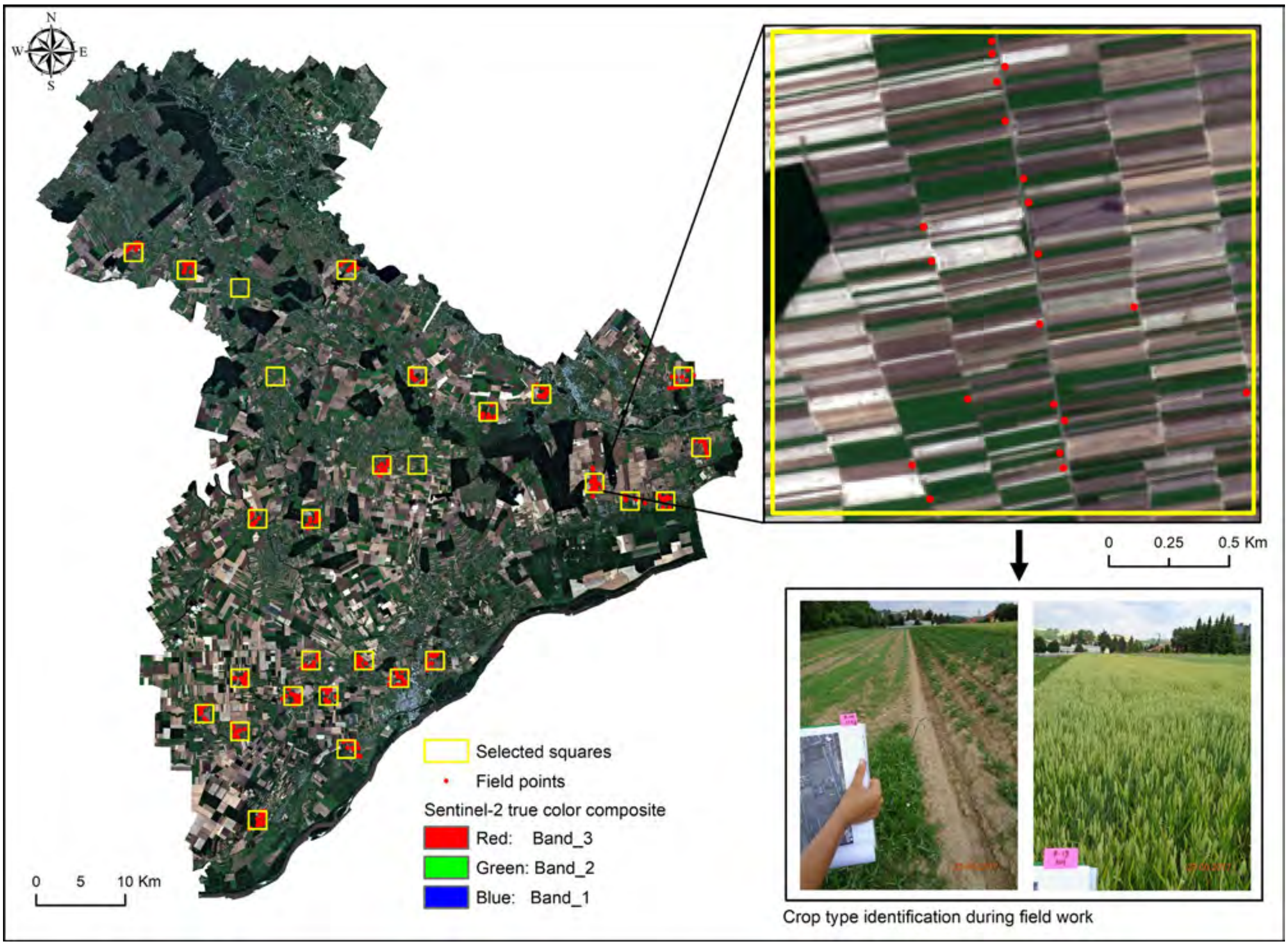


Figure 6 – Example of selected squares and spatial distribution of field points for crop type identification in one reference region

Figure 7 – Example of crop samples collected during the field survey campaign

Satellite Image Processing and Classification

The satellite data (Sentinel-2 images) used for the crop type map classification was obtained from ESA's Sentinel Scientific Hub (ESA 2017). All the available images for each reference region between April and September 2017 were downloaded. Only the images with cloud coverage inferior to 10 percent were used. Level-2A image products were selected since they are already atmospherically corrected.

A minimum of two Sentinel-2 images were established for crop type classification, ideally one image in spring and another in summer. However, considering the Sentinel-2A and 2B sensors, it was not possible to obtain the minimum cloud-free (<10 percent) Sentinel-2 images for 6 regions: Montana (Bulgaria); Jihočeský kraj (Czech Republic), Latgale (Latvia), Pieriga (Latvia), Nowosadeki (Poland), and Nowotarski (Poland). Therefore, the solution was to use the Sentinel-1 SAR images for the crop type mapping for those regions.

The Sentinel-1 satellite is a radar system that is able to register the Earth's surface despite the cloud cover. Sentinel-1A and 1B are able to obtain images of the same place over a 6-day interval, greatly improving the ability to identify crop species. Nevertheless, a higher number of available Sentinel-1 images imposes more computational power resulting in a time-consuming process. Therefore, a monthly composite using at least one image every 6 days was built using the Google Earth Engine cloud computer platform.

More details about technical characteristics of Sentinel-1 and Sentinel-2 images are described in the **Box 2**.

Box 2 – The Copernicus Constellation: Sentinel Satellites

The **Copernicus Programme** is a cornerstone of the European Union's efforts to monitor the Earth and her many ecosystems, whilst ensuring that her citizens are prepared and protected in the face of crises and natural or man-made disasters. Building on the foundations of deeply rooted scientific knowledge and decades of EU investment in research and technological development, the Copernicus Programme is exemplary of European strategic cooperation in space research and industrial development.

European Space Agency is currently developing a family of missions, called Sentinels, that are specifically designed for the operational needs of the Copernicus Programme.

Each Sentinel mission is based on a constellation of two satellites to fulfil revisit and coverage requirements, providing robust datasets for Copernicus Services.

These missions carry a range of technologies, such as radar and multi-spectral imaging instruments for land, ocean and atmospheric monitoring.

Sentinel Satellite Family:

- **Sentinel-1** is a polar-orbiting, all-weather, day-and-night radar imaging mission for land and ocean services (launched).
- **Sentinel-2** is a polar-orbiting, multispectral high-resolution imaging mission for land monitoring to provide, for example, imagery of vegetation, soil and water cover, inland waterways and coastal areas. Sentinel-2 can also deliver information for emergency services (launched).

- **Sentinel-3** is a multi-instrument mission to measure sea-surface topography, sea- and land-surface temperature, and ocean and land colour with high-end accuracy and reliability. The mission will support ocean forecasting systems, as well as environmental and climate monitoring (launched).
- **Sentinel-5 Precursor** – also known as Sentinel-5P – is the forerunner of Sentinel-5 that provides timely data on a multitude of trace gases and aerosols affecting air quality and climate (launched).
- **Sentinel-4** is a payload devoted to atmospheric monitoring that will be embarked upon a Meteosat Third Generation-Sounder (MTG-S) satellite in geostationary orbit.
- **Sentinel-5** is a payload that will monitor the atmosphere from polar orbit aboard a MetOp Second Generation satellite.
- **Sentinel-6** carries a radar altimeter to measure global sea-surface height, primarily for operational oceanography and climate studies.

Source: European Space Agency - ESA

Methodological workflow for crop type mapping

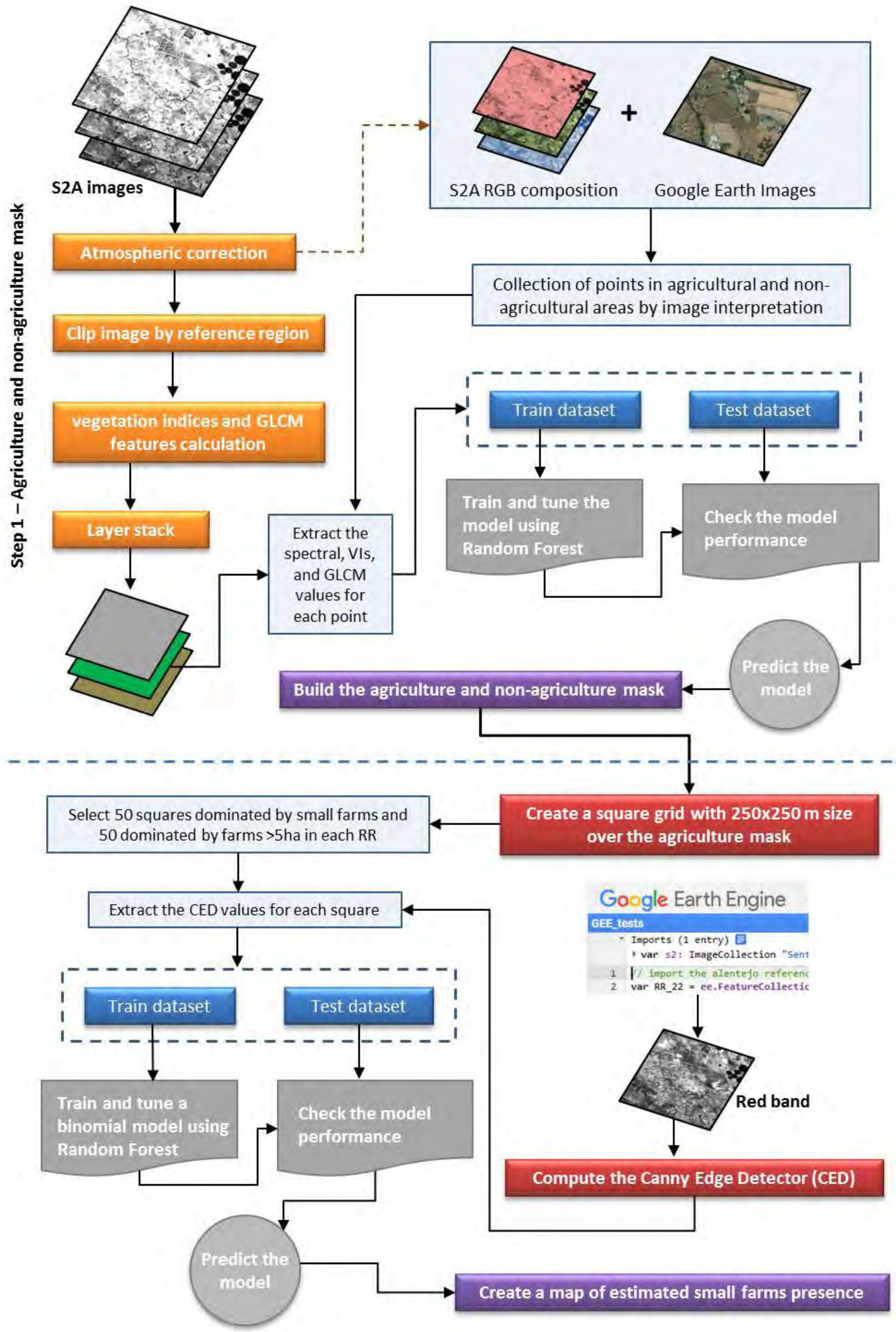
This section will briefly describe the workflow carried out to produce the crop type maps. The methodological framework consisted in 10 main stages as follows:

1. downloading of multi-temporal Sentinel images and preprocessing;
2. vegetation indices computation to be used as auxiliary information in the classification;
3. creation of multi-temporal NDVI stack for the segmentation process;
4. image segmentation;
5. preparation of the crop dataset to be used in the classification;
6. selection of the segments intersected by the crop field points;
7. extraction of all the pixels within the segments to be used in the classification procedure;
8. image classification (pixel-based) using Random Forest machine learning algorithm;
9. accuracy assessment of the classification; and
10. building of the small farms crop type classification and creation of the maps.

A detailed description about the methodological framework is described in **Deliverable 2.3** of the Salsa Project (Godinho *et al*. 2019).

Figure 8 presents a flowchart illustrating the image analysis process used to estimate the distribution of small farms across the European reference regions.

Figure 8 – Flowchart illustrating the processing scheme used to create distribution maps of small farms based on Sentinel-2A and Google earth images (Godinho et al. 2019)



Crop Type Dataset

A final dataset for each region was built based on the crop data which had been collected from field work. The dataset was checked point-by-point, through visual inspection of the digital photos taken by the SALSA teams on the field, as well as by superimposing the points over the Sentinel-2 and high-resolution images from the Google Earth Platform.

This process allowed for checking the thematic and geographic accuracy, which was essential to ensure data quality. Points that did not represent the correct crop data or that were not collected properly were excluded. Moreover, all the points registered as "non-identified crop type", "tillage lands" or "plowed lands" were removed from the final dataset.

After this process, the distribution of all validated points was combined with the segment boundaries (they represent the boundaries of the crop plot identified during the field work and are generated by the image segmentation process). In this way it was possible to identify the segments (polygons with the crop plots) related to each point.

Selection of the crops

In order to understand and assess the regional food systems, and particularly the contribution of small farms and related small food business to Food Nutrition and Security (FNS) in each reference region, a set of relevant crops were selected based on two criteria: 1) Importance in terms of production and consumption and 2) Largely produced in the region. Both criteria are related to the spatial representativeness of the crops within the region and thus potentially easier to obtain enough field information (field points).

In this way, the crop types that covered a residual percentage in terms of cultivated area, or not considered alimentary products (such as ornamental flowers and fiber crops) were not included in the analysis, reducing the classification errors. Nevertheless, in some specific cases, crop types that fitted both criteria were excluded from the analysis due to the lack of representative points in the dataset.

Image classification and accuracy assessment

For the satellite imagery classification, a pixel-based supervised Random Forest (RF) machine-learning algorithm was applied. The effectiveness of this pattern recognition algorithm has been demonstrated in several studies, with multiple applications in science, such as: crop classification (Li *et al*. 2020), mapping of soil contamination (Tan *et al*. 2020), prediction of river pollution (Victoriano *et al*. 2020) and human cancer diagnosis (Wang *et al*. 2020).

This method presents multiple benefits: it can be run efficiently on large databases; it can handle thousands of input variables without variable deletion; it is able to estimate the importance of variables to the model; it generates an internal unbiased estimate of the generalization error; it is relatively robust to outliers and noise, and it is computationally lighter than other tree ensemble methods (Rodriguez-Galeano *et al*. 2002).

The Random Forest classification in this study was implemented using the *randomForest* package (Law and Wiener 2002) from R software.

For the Random Forest classification, the crop segments dataset from each region were split into training (75 percent) and testing (25 percent) subsets. The training subset was used to train the RF model by using 1 000 trees to obtain stabilized variable importance estimation (Law and Wiener 2002). The test or validation subset in turn was used to evaluate the model performance through **confusion matrix** analysis. This method allows for calculating diverse accuracy elements of the model: **overall accuracy (OA), producer's accuracy (PA) and user's accuracy (UA).**

Furthermore, the F-score (harmonic mean) was computed for each crop. This score is often used to measure the accuracy of a test and classification performance, providing a more realistic measure using both precision and recall. The F-score ranges from 0 to 1; the higher the F-score, the higher the accuracy of the classification.

Results can be observed in Table 2.

Crop area and production estimation

The estimation of crop areas plays an important role, as it makes it possible to make a worldwide assessment of the total area, and with local data, estimate the production potential for different crop commodities. It is generated by combining the unbiased crop area with the field-level crop yields.

The understanding of the spatial patterns of agricultural production in small farms can reveal untapped opportunities regarding regional marketing, intensification and diversification, processing and trade and/or might uncover significant levels of regional inequality that could be helpful in shaping spatially-strategic responses to such opportunities and challenges. The more reliable the spatial information is, regarding the location (area) and performance (yield), the more cost-effective the formulation and targeting of appropriate policy and investment actions can be (You *et al*. 2014).

The classic way to estimate crop area through remote sensing techniques (classified images) is by simply counting the number of pixels allocated to each crop and multiplying by the pixel area. However, this method of crop area estimation is biased due to classification error (Canter 1997).

In order to avoid this error and reduce the uncertainties in crop production estimation, a direct calibrator estimator of the area based on ground truth information was used (Gallego 2004; Lambert *et al*. 2018).

After the estimation of the crop areas (unbiased), the crop production was calculated by using the estimated crop yields multiplied by the corresponding crop area of the small plots (<5 ha). The information of the crop yields was obtained through the farmers survey described in the **Salsa Deliverable 3.1**[4]**.**

For each of the 21 Reference Regions, the average number of farmers interviewed was 32; the farmers were asked about the crops that they produce, and when the selected key crops were present, they were asked to provide an estimated yield of the crop, based on the average value in the last five years. Based on the farmers survey, the average yields per key crop were computed for each reference region.

In order to ensure that the crop production estimation was precise, the unbiased crop area computation as well the subsequent crop production estimates were performed only for the key crop products that presented *F-score* classifications greater than 75 percent, guaranteeing a reliable estimation only for the best classified crops.

This approach reduced the error propagation, when excluding the areas where the crop classification presented low, or very low accuracy levels which could have reduced the strength of the overall conclusions about the main contributions of small farms for the FNS.

4 For more details, see the SALSA Report: *Deliverable 3.1 Set of 30 regional reports with the results of the validated in-depth analysis of regional food systems and the contribution of small farms and related small food businesses to FNS (reports based on a common reporting template).*

Box 3 – Machine Learning- Automated Supervised Classification and Random Forest Algorithm

In the Machine Learning - Automated Supervised Classification, the user or image analyst "supervises" the pixel classification process.

The user assigns the various pixel values or spectral signatures that should be associated with each category (class) of sites (objects). This is done by selecting representative sample sites (object) of a known crop (class) called Training Data. Each class is characterized by a specific spectral signature (fingerprint).

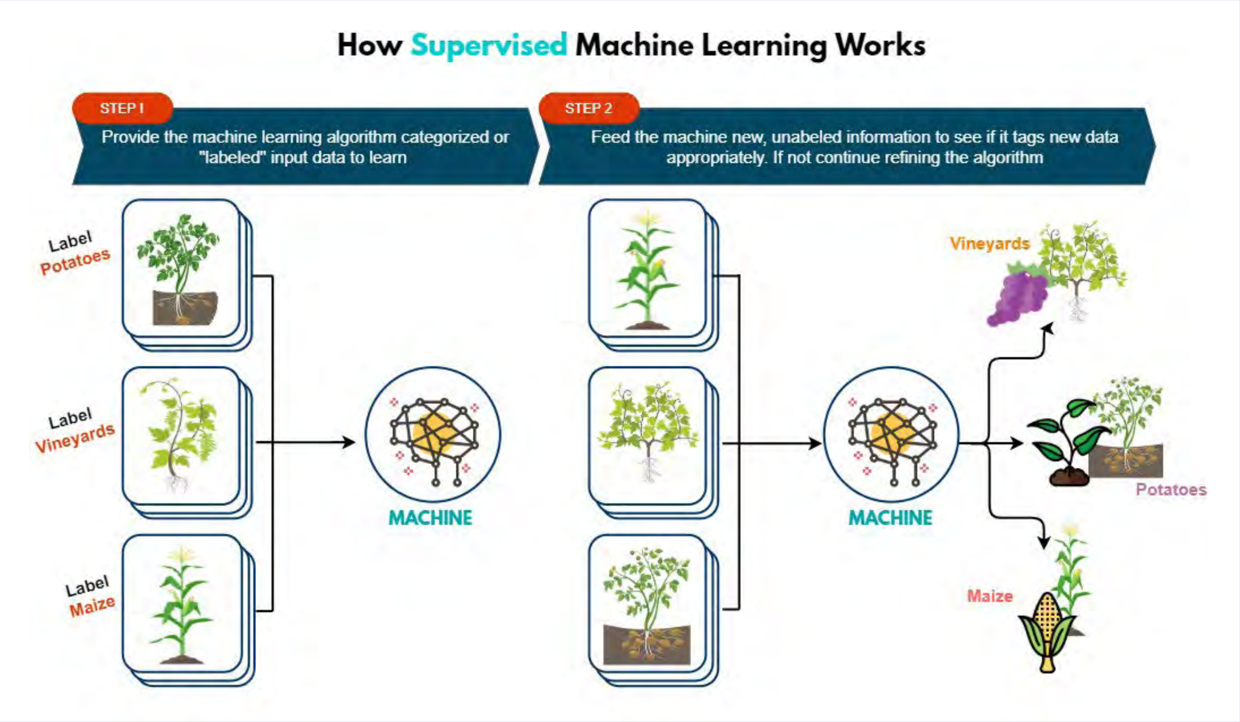
The computer algorithm then uses the spectral signatures from these training areas to classify the whole image. For an accurate classification, either the classes must be different from each other and not overlap, or they should only minimally overlap with other classes.

Random Forest Algorithm in turn is a supervised machine learning classification and regression technique that creates a vast number of uncorrelated decision trees at training time.

A forest is comprised of trees. It is said that the more trees a forest has, the more robust it is.

Random forest creates decision trees on randomly selected data samples, gets a prediction from each tree and then selects the best solution by voting in the most accurate decision tree. It also provides a ranking of the importance of variables.

Random forest algorithm has the advantage of dealing with noise and large datasets, and it provides better estimation of performance than traditional regression tree approaches.



Results

Crop types classification

Assessment of crop types diversity based on field data sampling

Crop diversity has been considered an important part of the agricultural landscapes, by benefiting a set of ecosystems services. Crop diversity enhances the resilience of the field to be attacked by pests and diseases, increases the suppressive effect of soils, provides shelter and food for beneficiary insects that can act as pollinators and biological control agents, reduces erosion and nutrient depletion, enhances the biodiversity of the soil, and has the potential to stabilize the national food production systems (Renard and Tilman 2019).

The collected crop field data represents a sample of all features present in each reference region in terms of its agroecosystem composition. Thus, aside from its importance in the image classification phase, the field data have the potential to be used as a reliable data source for a first assessment about the crop diversity that exists in each region.

Therefore, based on the crop plots visited in each region (max: 693, min: 316, average: 509), a total of 124 crop types were registered over the 21 regions. The highest crop diversity was registered in Imathia (Greece) with ± 35 crop types, and the lowest crop diversity was observed in Nowotarski (Poland) and Haouaria (Tunisia) with only 11 crop types (Figure 9).

In order to understand the relationship between crop diversity and farm size, all the 21 Reference Regions were grouped into four farm size categories (0-5 ha, 5-10 ha, 10-20 ha and > 20 ha) and for each category the unique crop types were counted.

The results revealed that the regions exhibiting the lowest mean farm size (0-5 ha) presented greater crop richness (diversity) when compared to the farmers with larger sized farms (>20 ha) - Figure 10.

These results suggest a negative relationship between crop diversity and mean farm size at the regional level, which is in agreement with the study conducted by Ricciardi *et al*. (2018) that indicates that agricultural landscapes with a predominance of small farms are more diversified, and thus have the potential to improve the nutrient adequacy and food security in such regions (Herrero *et al*. 2017).

The SALSA project also concluded that the contribution of small farms to increase crop diversity is more evident when considering the agricultural landscape rather than the crop diversity within the small farms.

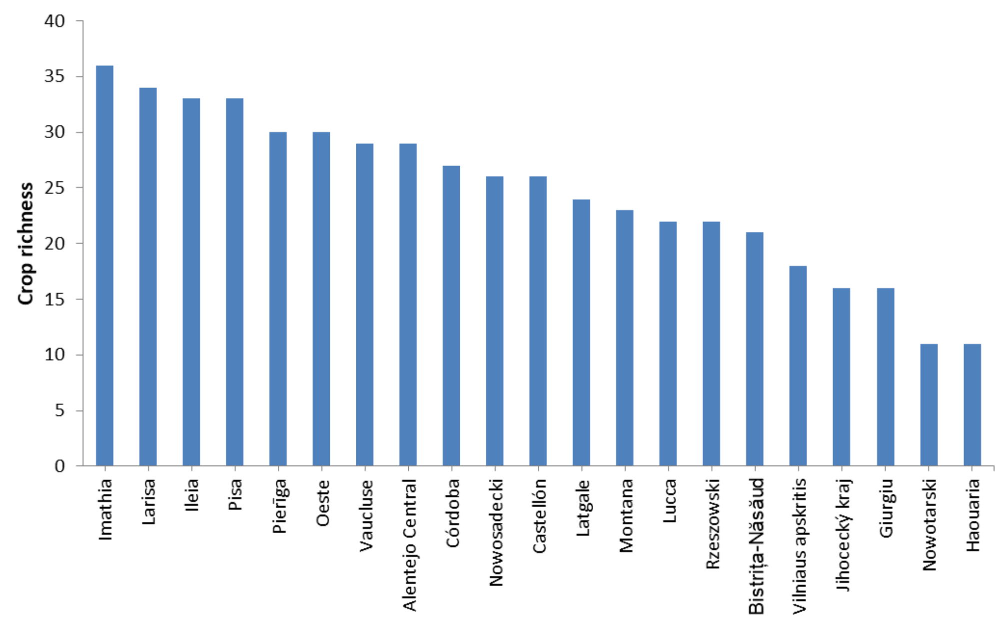


Figure 9 – Distribution of crop richness across SALSA Reference Regions

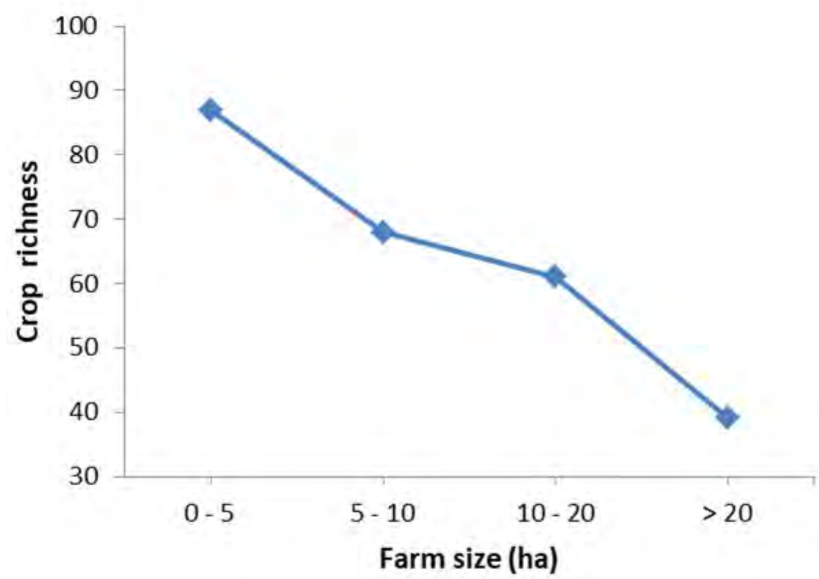


Figure 10 – Distribution of crop richness values over farm size categories.
Total unique crop species were counted for all Reference Regions with a mean farm size within one of the four categories. Regional mean farm size was obtained from the official statistics.

Crop type mapping and accuracy assessment

The satellite image supervised random forest classification performed for the selected data showed that the accuracy of crop types identification varied based on the geographical region, ranging from 59.5 percent to 91.4 percent with a mean overall accuracy of 81.6 percent (Table 2). Of the 21 regions in the study, 14 showed overall accuracy values higher than the mean value (in bold).

Imathia (Greece), Bistrita-Nasaud (Romania) and Rzeszowski (Poland), presenting an accuracy of 91.4 percent, 91.4 percent and 90 percent, respectively. The lowest accuracy values were observed in Castellón (Spain), Pisa (Italy) and Latgale (Latvia) showing an overall accuracy value of 59.6 percent, 65.7 percent and 72.6 percent, respectively.

Country	Reference Region	Satellite - derived map	Overall accuracy	Kappa
Bulgaria	Montana	Sentinel - 1	83,0%	0,76
Czech Republic	Jihočeský kraj	Sentinel - 1	82,8%	0,73
France	Vaucluse	Sentinel - 2	88,3%	0,85
Greece	Imathia	Sentinel - 2	91,4%	0,83
	Larisa	Sentinel - 2	86,8%	0,77
	Ileia	Sentinel - 2	78,3%	0,73
Italy	Lucca	Sentinel - 2	87,4%	0,83
	Pisa	Sentinel - 2	65,7%	0,55
Latvia	Latgale	Sentinel - 1	72,6%	0,55
	Pierigia	Sentinel - 1	81,9%	0,72
Lithuania	Vilniaus Apskritis	Sentinel - 2	85,6%	0,76
Poland	Rzeszowski	Sentinel - 2	90,0%	0,89
	Nowosadecki + Nowotarski	Sentinel - 1	78,9%	0,64
Portugal	Alentejo Central	Sentinel - 2	73,8%	0,65
	Oeste	Sentinel - 2	82,5%	0,77
Romania	Bistrița-Năsăud	Sentinel - 2	91,4%	0,88
	Girgiu	Sentinel - 2	83,6%	0,79
Spain	Castellón	Sentinel - 2	59,6%	0,50
	Córdoba	Sentinel - 2	83,2%	0,78
Tunisia	Haouaria	Sentinel - 2	86,0%	0,82
		Mean	81,6%	0,74

Table 2 – Overall accuracy and kappa values obtained for each crop type map among the 21 Reference Regions

Four main reasons can be considered to be responsible for the differences in terms of accuracies observed over the regions: i) different number of satellite images used per region; ii) the crop season of the images; iii) the spatial and spectral heterogeneity of each agricultural landscape and within crop type classes analysis (ex. Figure 11); and iv) the availability and representativeness of the field dataset – crop type polygons to train the classification models – (Torbick *et al.* 2018; Lobell, 2013; Jain *et al.* 2013, 2016; Lambert *et al.* 2018 ; Lebourgeois *et al.* 2017 and Teluguntla *et al.* 2018).

Figure 11 – Example of how different two agricultural landscapes can appear. The images here are seen from spatial and spectral heterogeneity points of view. Both images represent an NDVI RGB (R: NDVI - April; G: NDVI - June; B: NDVI - August). Upper: Imathia (Greece); Lower: Castellón (Spain).

There is no hierarchical importance in the sequence of the four main reasons presented above. In fact, detailed research that is truly focused on the study of the main impacts of each one of these factors on the accuracy levels of satellite-derived crop maps (and in general land cover/use) is still lacking in the literature. Therefore, future efforts towards a better understanding of this issue should be promoted. In general, the results reported here are clearly in agreement with some recent studies that have been proving the accuracy in using Sentinel-1 and Sentinel-2 data to produce crop type/ land cover maps in such complex small-scale agricultural areas (Lambert *et al*. 2018; Kenduiynwo *et al*. 2018 and Clerici *et al*. 2017).

The crop classification levels derived from the confusion matrix and produced for each reference region show that F-scores vary over the regions and among the crops. In order to better understand the different accuracy levels between crop types, the most common crop types were selected for comparison purposes. Among the analyzed crops, cereals showed to be more reliably mapped following the described methodology, with an average F-score of 82.8 percent, followed by meadows, pastures and forage crops (MPFC) with an F-score of 78.5 percent, vineyards (F-score = 78.4 percent), and orchards (F-score = 69.9 percent). The least well-classified class of crops in this study was that of the vegetables. This class comprised a high diversity of crops (e.g. carrots, legumes, onions and potatoes), with an average F-score of 52.7 percent. This weak result can be explained due to the different crop architecture of the vegetables, and high heterogeneity on the spectral signature of the plants (Figure 12). On the contrary, the maize crop class was identified as the more homogenous crop in terms of spectral signature.

However, the results reported here showed a high variation among maize F-scores, with values ranging from 97.1 percent to 11.1 percent (average of 64.4 percent). These results may indicate that the number and size of the data sample used for maize classification, applying the random forest model, are determining the classification accuracy for this crop. These results corroborate with the main conclusions founded by Champagne *et al.* (2014) that showed that classification accuracy increases with an increase in the number of field samples used to train the classification model.

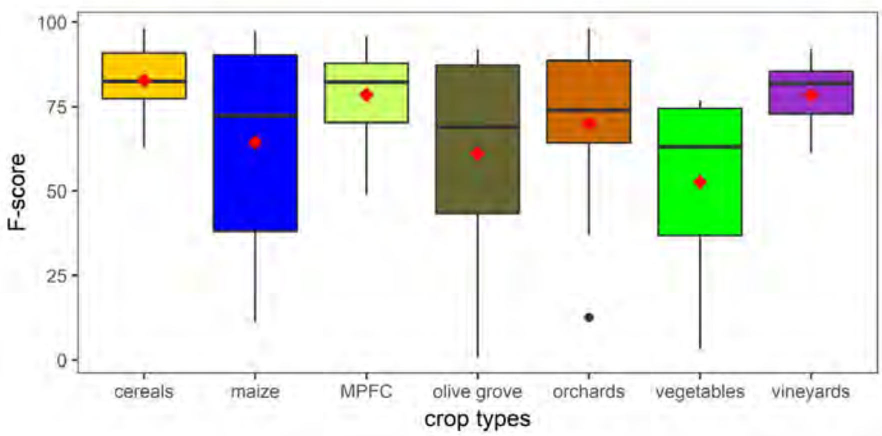


Figure 12 – Comparison of the F-score values obtained for each dominant crop type. Red diamonds represent the mean F-score value, while the black lines inside the box represent the median.

An example of the random forest classification (for Pisa, Italy) can be observed in Figure 13, which highlights the small agricultural plots (< 5 ha).

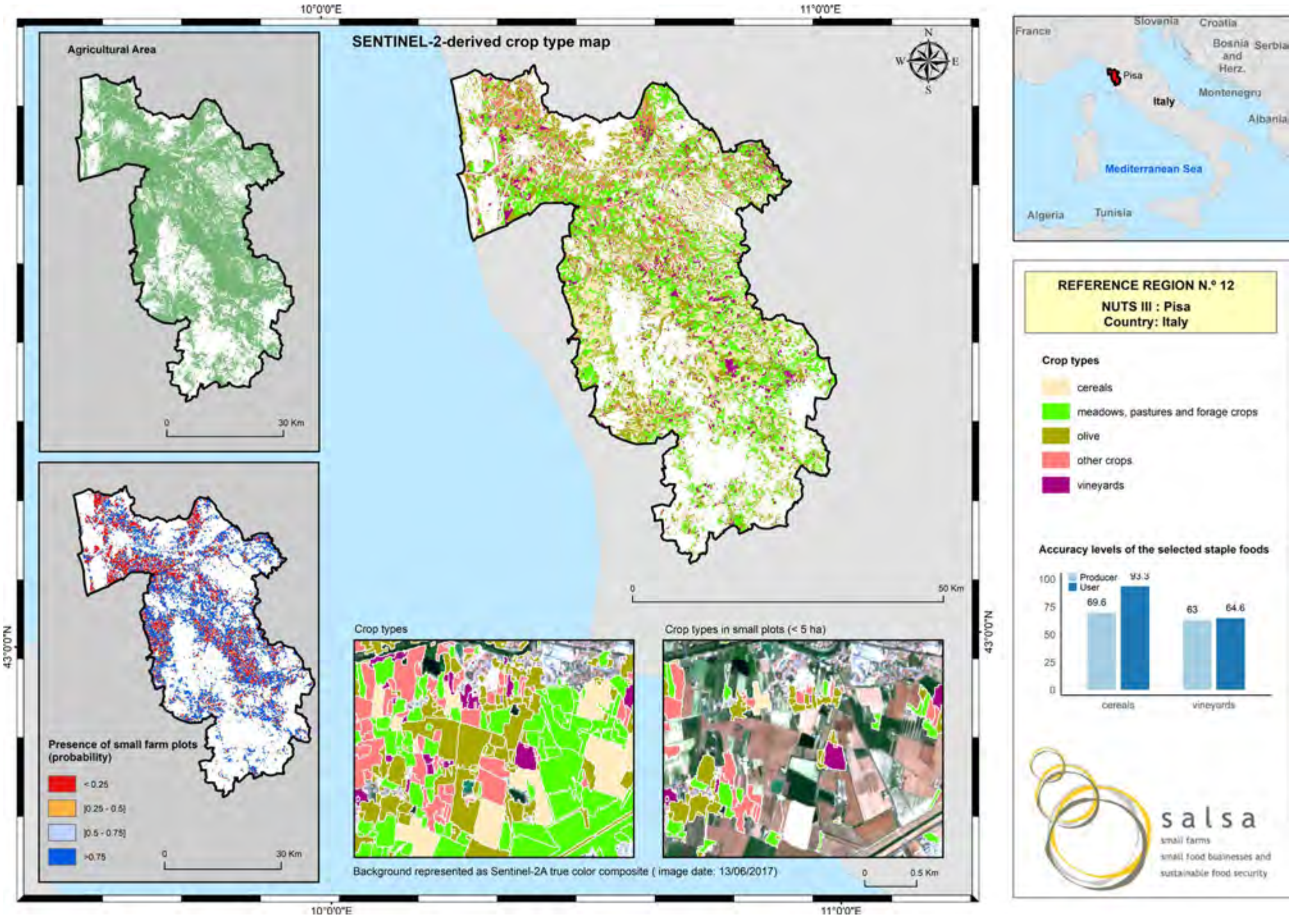


Figure 13 – Example of crop types in the small farms assessment using Sentinel 2 satellite imagery

Crop area and crop production estimation

Crop area estimation

As mentioned in the previous sections, in order to compute an unbiased area for the key crops, only the crops with classification accuracy superior to F-score 75% should be used to guarantee more confidence when drawing conclusions. However, in order to make it possible to estimate the crop production for at least one key crop per selected region, an F-score lower than 75 percent was accepted for Larisa (F-score for vegetables = 73.2 percent), Latgale (F-score for wheat = 68.1 percent), Nowosadecki and Nowotarski (F-score for cereals = 70.8 percent). For Pieriga (Latvia), the key crops products that were mapped were wheat, vegetables, and orchards, with F-scores of only 18.4 percent, 17.9 percent and 68.0 percent, respectively.

For this region, a class of cereals (wheat, oat, barley, and rye) was also mapped and presents a high accuracy level (F-score = 88.8 percent). It was thus selected as a key crop product instead of wheat. Regarding the field-level yield data obtained from the farmers' interviews, it was found that such information was absent for some key crop products or for some regions it was calculated based on very few farmers' interviews (Table 3). Therefore, caution must be employed before interpreting the crop production estimations for those regions with limited field-level yield data.

Table 3 - Unbiased crop area for small plots of <5 ha (output 4) and production estimations (output 5) for each key crop product

Reference region	Key crop types mapped in WP2	Fscore (%)	crop area (ha)	# of farmers interviews	yield estimation (ton/ha)	crop production (ton/year)
Montana	Cereals	83,4	3523,98	1	4	14095,92
Jihocecky kraj	Wheat	85,5	2140,48	0	6.1*	13056,94
Vaucluse	Vineyards	92,1	10578,4	2	4,15	43900,4
Imathia	Peaches	80,4	8782,06	38	30,82	270663,09
Larisa	Vegetables	73,2	814,55	6	3,13	2549,55
Ileia	Olives Groves	85,5	20618,2	28	6,51	134224,28
	Vineyards	77,3	2289,35	5**	13,6	31135,1
Lucca	Olives Groves	87,2	2180,83	18	4,23	9224,9
	Vineyards	81,9	789,9	13	7,35	5805,79
Pisa	Cereals	75,4	1304,62	5	4,4	5740,32
Latgale	Wheat	68,1	6593,26	18	2,98	19656,85
Pierigia	Cereals	88,8	3389,91	7	2,17	7356,1
Vilniaus Apskritis	Vegetables	76,5	1592,14	11	3,9	6209,34
Rzeszowski	Cereals	91,7	15603,1	34	4,79	74738,9
	Potatoes	86,9	8714,34	29	18,53	161476,72
Nowosadecki	Cereals	70,8	10779,93	38	3,31	35681,56
	Apples	81,4	1705,5	9	33,33	56844,32
Nowotarski	Cereals	70,8	3020,26	26	2,89	8728,55
Alentejo Central	Vineyards	87,5	1867,13	16	7,89	14731,66
Oeste	Pears	90,9	2407,64	16	14,01	33731,04
	Vineyards	83,4	3627,34	21	7,26	26334,49
Bistrița-Năsăud	Vegetables	75,2	532,13	0	-	-
	Orchards	98,2	4799,07	21	1,59	7630,52
Girgiu	Cereals	98,2	17416	14	4,39	76456,24
	Sunflower	75,9	6152,74	4	1,69	10398,13
Castellón	Citrus	88,1	17016,44	10	34,2	581962,3
Córdoba	Cereals*	87,8	2742,21	10	2,95	8089,52
	Olives Groves	87,2	31449	12	5,15	161962,35
	Vineyards	85,5	1857,41	10	8,62	16010,87
Haouaria	Tomato	95,5	852,33	11	65,91	56117,37
	Pepper	88,8	1332,44	10	6,95	9260,45
		Total	196475,7		Total	1903833,35

Note: * Due to the absence of field-level wheat information the mean national wheat yield was used.
** For Ileia, the information about vineyard yields was obtained in two different ways: 5 out of the 17 inquired farmers delivered the information in tonnes of grapes per hectare, while 12 give this information in dried grapes per hectare. In order to be comparable with other regions with vineyards, the information in tonnes of grapes per hectare was used.

The consistency of the Sentinel-derived crop areas was evaluated against documented crop areas from official statistics at the regional level. For this purpose, the area covered by each key crop product cultivated by small farms in each region was extracted from the regional statistics and regressed against the unbiased crop area (plots < 5 ha) estimated by Sentinel using linear regression (Figure 14).

The relation between crop areas from both data sources (official statistics and Sentinel data) shows a significant and very high correlation with an R^2 value of 0.96 ($p<0.001$), demonstrating there is no significant difference between the Sentinel-based crop area and the official regional statistics.

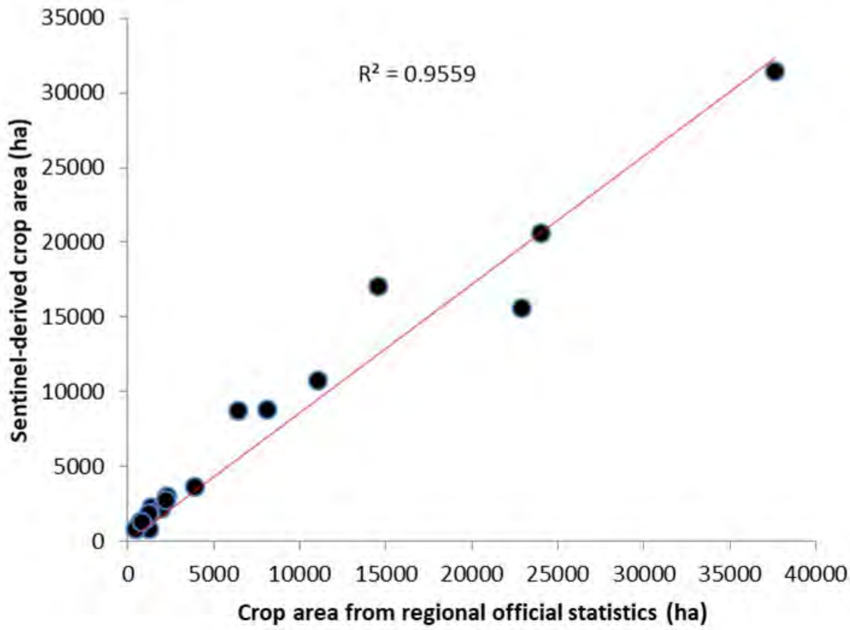


Figure 14 – Comparison between Sentinel-based crop area estimations and those from the official regional statistics

These results were surprising for the SALSA project team, since it was expected that a considerable number of small farms were not considered for the statistic records. On the one hand, it was possible to conclude that the official regional statistics information concerning small farms are more accurate than expected. On the other hand, this high correlation can also confirm the suitability of the use of Sentinel images in providing accurate and reliable information about crop area extent in complex agricultural systems.

Among the Sentinel-based crop area estimations obtained for the 21 Reference Regions, only potatoes in Rezewoski (Poland) and citrus in Castellón (Spain) showed overestimation in comparison with the official statistics. Cereals in Rezewoski (Poland) presented underestimation compared with the official data. In these cases, the official statistics may not cover all the information of small productive farms areas per crop.

Even if some of the Sentinel-based area estimations differed slightly from the official records, the overall results clearly demonstrate that the crop area obtained from the Sentinel imagery can be used with confidence, in particular for those regions where the information is not fully represented by the official statistics.

To have a better comprehension of the relative importance of the small farms for each crop type, the percentage of the small farms' contribution in terms of area in relation to the total area (all farm ranges) per crop was calculated. For this, the total area covered by specific crop product over the region was extracted from the official statistics at a regional level (Figure 15).

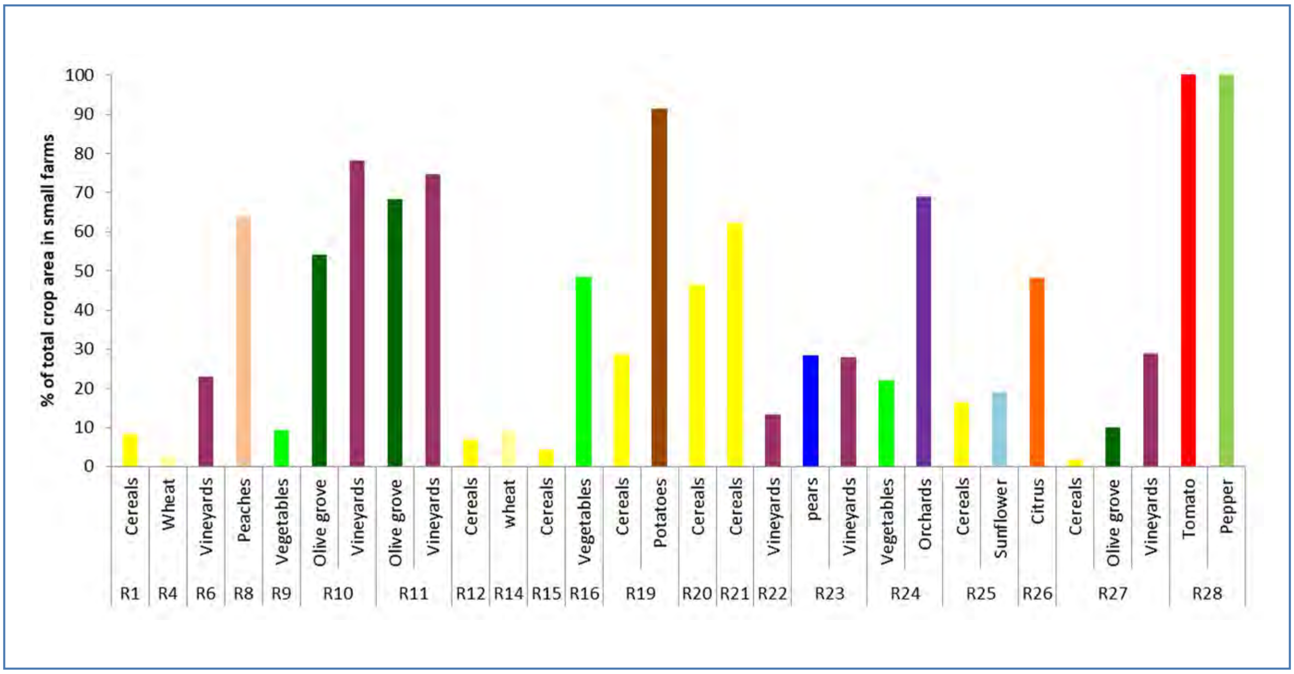


Figure 15 – Percentage of the total regional crop type area cultivated in small farms. In some regions, small farms correspond to almost 100 percent of the total crop area, which highlights the relevance of small farms in these contexts.

Crop production estimation

The unbiased cultivated area estimations merged with the field-level yields assessment of the key crops resulted on crop production estimations for the small farms across the reference regions. The results showed a total of 1 903 833.35 tonnes of agricultural products produced by small farms over the 21 Reference Regions of which 1 088 749.56 tonnes of fruits (grapes, apples, pears, peaches and other orchards), 315 809.66 tonnes of oil crops (olives and sunflower grains), 263 600.90 tonnes of cereals (wheat, barley, oats and rye), and 235 673.43 tonnes of vegetables (eg. potatoes, tomatoes, peppers, pulses, beans, etc.).

These results indicated that the small farms present in the reference regions have the potential to produce an average of 19.5 tonnes/ha/year of fruits, 18.0 tonnes/ha/year of vegetables, 5.2 tonnes/ha/year of oil crops, and 4.0 tonnes/ha/year of cereals.

Given the considerable difference in crop area estimations, and yield levels due to weather conditions, agronomic management and market orientation, different production levels were observed within key crops across reference regions (e.g. cereal: min 2.17 tonnes/ha; max= 4.79 tonnes/ha).

Considering absolute values, the highest production levels were from citrus (581 962.30 tonnes) in Castellón (Spain); peaches (270 663.09 tonnes) in Imathia (Greece); olives (161 962.35 tonnes) in Córdoba (Spain); and potatoes (161 476.72 tonnes)in Rzeszowski (Poland). The lowest production estimations were obtained for vegetables (2 549.55 tonnes) in Larissa (Greece), cereals (5 740.32 tonnes) in Pisa (Italy), vineyards (5 805.79 tonnes) in Lucca (Italy) and vegetables (6 209.34 tonnes) in Vilniaus Apskritis (Lithuania).

The importance of small farms in terms of crop production were evaluated and the results proved that small farms are responsible for a very high percentage of the total regional production for a set of crops (Figure 16).

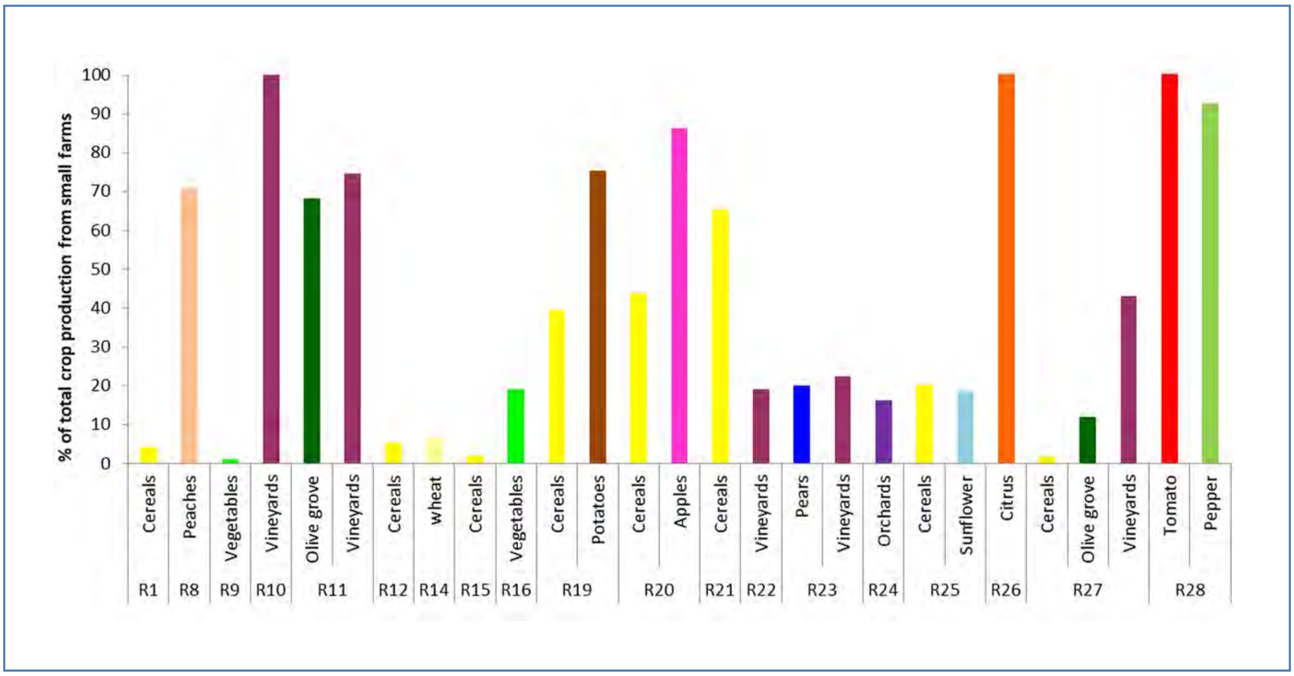


Figure 16 – Percentage of the total regional crop type production produced by small farms. Here, there are also products that are almost 100 percent produced by small farms.

In Haouaria, for example, small farms are responsible for producing 92.6 percent of the peppers. In Nowosadecki, the small farmers contribute with 86.4 percent of the apple production; in Rzeszowski the small farms are responsible for 75.3 percent of the potato production; small farms also are responsible for 70.9 percent of the peach production in Imathia.

The percentage of small farms' production that was higher than the total estimation was found in two cases: tomatoes in Haouaria (102.1 percent) and citrus in Castellón (108.1 percent). These results can be explained due to the differences observed in terms of area and yield estimations obtained from Sentinel data, field interviews and official statistics.

For example, from the Sentinel crop map the total area covered by citrus in small scale plots was estimated as 17 016.74 hectares, while the official statistics present an area of 14 583.70 hectares, resulting in a difference of 2 433.04 hectares. In addition,, the field-level information showed different results from those reported on official data regarding productivity (yield per area): from the interviews it was estimated that there was a yield of 34.2 tonnes/ha while in the official statistics this value was only 15.3 tonnes/ha for the Castellón region.

However, none of these values seem to be accurate, since the national average citrus yield is 19.1 tonnes/ha (Navarro *et al*. 2015). Therefore, with more area covered by citrus, as well as a higher yield value used to compute the citrus production, it is expected that the relative importance of small farms in this region exceeds 100 percent in terms of citrus production.

Regarding the tomato production estimations, the area covered by this crop, which was obtained with a very high accuracy level (F-score=95.5 percent) from the Sentinel data, was 852.3 hectares which contrasts with the 480.0 hectares given by the official statistics. The difference between the tomato yields obtained from field interviews (65.9 tonnes/ha) and the ones from official statistics is very low (68.8 tonnes/ha), suggesting that the main difference in tomato production estimations is much more connected to some inaccuracies in the official statistics of Haouaria region, specifically concerning the tomato area statistics.

Final Remarks

The obtained results derived from Sentinel-1 and Sentinel-2 images acquired during the spring-summer season of 2017 produced good classification accuracies (mean values OA = 81.6 percent and F-score = 70.2 percent for several crop types under small scale farming systems with different environmental and territorial conditions. Appling the methodology to the 21 Reference Regions, it was possible to understand the effectiveness of Sentinel data combined with field data in producing high-resolution crop maps.

The field data is needed to capture the spectral signature (fingerprint) of each main crop type across different agricultural season periods. This process is especially needed when small farms dominate the agricultural landscape and crop diversity is the main characteristic in the spatial pattern.

Aside from the classical accuracy metrics used to access the suitability of Sentinel data in producing accurate and useful information about crop area extent in small-scale farming systems, the results obtained through remote sensing were compared with the official statistics of each reference region. The results presented a strong correlation (R^2 = 0.96, p-value < 0.001) between Sentinel-based estimations and official statistics.

This result leads to two main conclusions:

a) These images can be used as the main source to provide fairly accurate estimations on crop area extent for regions where there is no up-to-date information, or where no information exists. Furthermore, this opens up the possibility of monitoring changes, avoiding the very heavy procedure of data collection for statistical data sets.

b) When statistical data sets concerning small farms and differentiating crops exist, the available data has shown to be accurate and can thus be used as quality information about small farms.

Regarding the crop production estimations generated by combining the unbiased crop area with the field-level crop yields, the results hereby reported highlighted the fact that small farms make an important contribution in terms of crop production of the selected crops.

In summary, this report clearly shows that Sentinel-1 and Sentinel-2 missions open a new era of opportunities towards the development of more robust tools and methodologies based on remote sensing data to accurately assess food security in small scale farming systems and monitor changes using an accelerated method. The current unseen drivers of change linked to climate change, global drivers and markets highlight the importance of straightforward evaluation and monitoring methods on farm systems and production variations, thus the use of these tools hereby exploited offer a range of possibilities that are worth exploring further in the future.

References

Godinho, S., Guiomar, N.,Machado, R. & Lourenço, P. 2019. Report on the assessment and characterization of small farms distribution and spatial characteristics obtained from SENTINEL-2 data - Deliverable 2.4 – WP2 – Salsa Project: http://www.salsa.uevora.pt/wp-content/uploads/2019/08/D-2.4-Report_April2019_final_version160819.pdf

Theodore Tsiligiridis, Katerina Ainali, Teresa Pinto CorreiaBeddow, J. M., Hurley, T. M., Pardey, P. G., & Alston, J. M. 2014. Food security: Yield gap. Encyclopedia of agriculture and food systems, 3, 325-365.

Canters, F. 1997. Evaluating the Uncertainty of Area Estimates Derived from Fuzzy Land-Cover Classification. Photogrammetric Engineering & Remote Sensing. 63:403-414.

Champagne, C., McNairn, H., Daneshfar, B. & Shang, J. 2018. A bootstrap method for assessing classification accuracy and confidence for agricultural land use mapping in Canada. International Journal of Applied Earth Observation and Geoinformation. 29:44-52

Clerici, N., Calderón, C.A.V. & Posada, J.M. 2017. Fusion of Sentinel-1ª and Sentinel-2ª data for land cover mapping: a case study in the lower Magdalena region, Colombia. Journal of Maps. 13:718-726.

Davidova, S., Bailey, A., Dwyer, J., Erjavec, E., Gorton, M. & Thomson, K., 2013. Semisubsistence Farming − Value and Directions of Development Policy Department B: Structural and Cohesion Policies. European Parliament, Brussels.

European Space Agency, 2017. Sentinel Scientific Hub. Retrieved over 2017, from https://scihub.copernicus.eu/dhus/#/home

Fritz, S., See, L., You, L., *et al* 2013. The need for improved maps of global cropland. Eos, Transactions American Geophysical Union, 94(3), 31-32.

Gallego, F.J. 2004. Remote sensing and land cover area estimation. Int. J. Remote Sens. 25 (15), 3019-3047.

Herrero *et al.* 2017. Farming and the geography of nutrient production for human use: a transdisciplinary analysis. Lancet Planet Health. 1: 33-42.

Jain, M., Srivastava, A. K., Singh, B., Joon, R.K., McDonald, A., Royal, K., Lisaius, M. C. & Lobell, D. B. 2016. Mapping Smallholder Wheat Yields and Sowing Dates Using Micro-Satellite Data. Remote Sensing. 8, 860.

Jain, M., Mondal, P., DeFries, R.S. & Small, C., 2013. Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. Remote Sensing of Environment. 134: 210–223.

Kenduiywo, B.K., Bargiel, D. & Soergel, U. 2018. Crop-type mapping from a sequence of Sentinel 1 images. International Journal of Remote Sensing. 39:6383-6404.

Kostrowicki, J. "Types of agriculture in Poland. A preliminary attempt at a typological classification." Geographia Polonica 19 (1970): 99-110.

Lambert, M.J., Traoré, P.C.S., Blaes, X., Baret, P. & Defourny, P. 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. Remote Sensing of Environment. 216:647-657.

Lebourgeois, V., Dupuy, S., Vintrou, E., Ameline, M., Butler, S. & Bégué, A. 2017. A Combined Random Forest and OBIA Classification Scheme for Mapping Smallholder Agriculture at Different Nomenclature Levels Using Multisource Data (Simulated Sentinel-2 Time Series, VHRS and DEM). Remote sensing. 9, 259.

Li, H., Zhang, C., Zhang, S., & Atkinson, P. M. 2020. Crop classification from full-year fully-polarimetric L-band UAVSAR time-series using the Random Forest algorithm. International Journal of Applied Earth Observation and Geoinformation, 87, 102032.

Liaw and M. Wiener 2002. Classification and Regression by randomForest. R News 2(3), 18-22.

Lobel, D.B. 2013. The use of satellite data for crop yield gap analysis. Field Crops Research. 143: 56-64.

Muchoney, D.M. & Strahler A.H. “Pixel-and site-based calibration and validation methods for evaluating supervised classification of remotely sensed data.” Remote Sensing of Environment 81.2-3 (2002): 290-299.

Renard, D. & Tilman, D. 2019. National food production stabilized by crop diversity. Nature, 571(7764), 257-260.

Ricciardi, V., Ramankutty, N., Mehrabia, Z., Jarvisa, L. & Chookolingo, B. 2018. How much of the world’s food do smallholders produce? Global Food Security. 17: 64 – 72.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sanchez, J.P. 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, 67: 93-104.

Tan, K., Wang, H., Chen, L., Du, Q., Du, P., & Pan, C. 2020. Estimation of the spatial distribution of heavy metal in agricultural soils using airborne hyperspectral imaging and random forest. Journal of hazardous materials, 382, 120987.

Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K. & Huete, A. 2018. A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. ISPRS Journal of Photogrammetry and Remote Sensing. 144:325-340.

Torbick, N., Huang, X., Ziniti, B., Johnson, D., Masek, J. & Reba, M. 2018. Fusion of Moderate Resolution Earth Observations for Operational Crop Type Mapping. Remote Sensing. 10, 1058.

Victoriano, J. M., Lacatan, L. L., & Vinluan, A. A. 2020. Predicting River Pollution Using Random Forest Decision Tree with GIS Model: A Case Study of MMORS, Philippines. International Journal of Environmental Science and Development, 11(1).

Wang, S., Wang, Y., Wang, D., Yin, Y., Wang, Y., & Jin, Y. 2020. An improved random forest-based rule extraction method for breast cancer diagnosis. Applied Soft Computing, 86, 105941.

You, L., Wood, S., Wood-Sichra, U., & Wu, W. 2014. Generating global crop distribution maps: From census to grid. Agricultural Systems, 127, 53-60.

Earth Observation Data and the Sustainable Development Goals: Insights from FAO-ESA Partnership

In order to monitor the commitment by world leaders to achieve the agreed upon Sustainable Development Goals and targets (SDGs), countries are expected to report on the progress made towards achieving the SDGs on a regular basis. However, five years after the implementation of the 2030 agenda, the data and capacity gaps for monitoring SDGs, including data related to food and agriculture, are still huge. For instance, in 2019, countries were able to report on average at least one data point for just 43 percent of the SDG indicators under FAO custodianship. One of the main tenets in improving data generation for monitoring progress related to agriculture is to bridge the existing data gaps in many countries while reducing the cost of data production and dissemination. Addressing this seemingly paradoxical problem requires adopting new approaches and investing in innovative, modern and cost-effective tools for data collection, analysis and dissemination.

Satellite Earth Observation (EO) data can support countries to generate the information needed to monitor the SDG targets and basic agricultural statistics more efficiently by reducing the cost of collecting and analysing data. Land-based SDG indicators under FAO custodianship (including those under Goals 2, 6 and 15, and more specifically the physical component of such indicators) can be computed by analysing EO data[1] and/or EO products[2] that are developed and maintained by space agencies, research institutes and academia, and in most cases, freely available. However, the use of EO has its own challenges as its access, storage, preprocessing and analysis demand technical capabilities that can limit their uptake by countries.

FAO aims to tackle these issues through short- and medium-term projects of technical assistance to overcome these technical barriers, and support countries in using EO data to improve agricultural statistics in general, and in particular, to monitor SDGs.

In the short term, FAO is establishing a collaboration with European Space Agency (ESA) and Catholic University of Louvain (UCL) to implement a user-friendly and open-source solution, namely the Sen2Agri[3] tool box, to create national crop maps which can be used to generate some crop-related statistics.

FAO is committed to delivering specific in-country technical assistance in the uptake of the Sen2Agri tool as one of the cost-effective methods to improve the coverage, quality and timeliness of agricultural statistics, therefore enabling timely country SDG reporting. Furthermore, FAO is committed to building capacity on top of the Sen2Agri crop maps to extract crop acreage statistics for early crop yield assessment and forecasting.

The expected outcome of these planned projects, that were implemented in two pilot countries (Senegal and Uganda), is to ensure that the technical capacities for monitoring crops and generating agricultural statistics in these two countries are enhanced, including the generation of:

- crop maps
- crop area statistics
- crop acreage estimates
- crop yields map and statistics.

1 Earth Observation (EO) data refers to measurements from a range of satellites such as the Copernicus Sentinels or the Landsat missions.

2 EO product refers to classified remote sensing data (e.g. land cover map, crop map, water bodies etc.), biophysical parameters (e.g. Leaf Area Index, evapotranspiration) or derived indexes (e.g. vegetation and water indexes: Normalised Difference Vegetation Index, NDVI; Enhanced Vegetation Index, EVI; Normalised Difference Water Index, NDWI etc).

3 Sen2Agri open source toolbox for national crop monitoring: http://www.esa-sen2agri.org

While crop maps, crop acreage estimates, crop area statistics and yield maps will be provided through the production line of Sen2Agri and further developed tools, this EO information will be integrated in the working environment of National Statistics Offices through the FAO and ESA joint implementation of the ongoing Sen4Stat project[4]. By utilising the results from these pilot countries, this project can be scaled up to include a larger number of countries, in particular those where the capacity to generate crop statistics is relatively low.

FAO's long term vision is to enable the National Statistics Offices (NSOs) to integrate the use of open EO data and tools (produced by space agencies, research institutes and academia) into their national statistics programme by validating the EO outputs using low cost in situ data collection mechanisms. This is expected to contribute to strengthening the NSO's statistical reporting systems to produce accurate and timely data on SDG indicators as well as data related to agricultural statistics. To this end, the FAO and ESA are working towards establishing a long-term collaboration on the use of EO data from ESA for improving agricultural statistics and SDG monitoring.

In conclusion, the long-term vision of FAO is to use the EO data for three main applications: a) direct measurement of agricultural statistics; b) provision of subnational disaggregated estimates for some SDG indicators; and c) improvement of the statistical efficiency of field and household surveys. The following table describes the use of the EO data under the three categories described above for selected SDG indicators related to food and agriculture.

SDG	Indicator Number	Indicator	Direct Measure	Disaggregation	Survey Design
2 ZERO HUNGER	2.1.1	Hunger		✔	✔
	2.1.2	Severity of food insecurity		✔	✔
	2.3.1	Productivity of small-scale food producers		✔	✔
	2.3.2	Income of small-scale food producers			✔
	2.4.1	Agricultural sustainability	✔	✔	✔
5 GENDER EQUALITY	5.a.1	Women's ownership of agricultural land		✔	✔
	5.a.2	Women's equal rights to land ownership		✔	
6 CLEAN WATER AND SANITATION	6.4.1	Water use efficiency	✔	✔	✔
	6.4.2	Water stress	✔	✔	✔
12 RESPONSIBLE CONSUMPTION AND PRODUCTION	12.3.1	Global food losses		✔	✔
14 LIFE BELOW WATER	14.4.1	Fish stock sustainability			
	14.6.1	Illegal, unreported undregulated fishing			
15 LIFE ON LAND	15.1.1	Forest area	✔	✔	✔
	15.2.1	Sustainable forest management	✔	✔	✔
	15.4.2	Mountain Green Cover	✔	✔	✔

4 Sen4Stat – EO for national agricultural statistics: https://www.esa-sen4stat.org

Acronyms

EDORA: European Development Opportunities in Rural Areas

ESPON: European Spatial Planning Observation Network

ESU: Economic Size Units

FAO: Food and Agriculture Organization of the United Nations

FNS: Food and Nutrition Security

FSS: European Farm Structure Survey

GPS: Global Positioning System

ha: hectare

MPFC: Meadows, Pasture and Forage Crops

NDVI: Normalized Difference Vegetation Index

NUTS: Nomenclature of Territorial Units for Statistics

OA: Overall Accuracy

PA: Producer's Accuracy

R2: Coefficient of determination

RF: Random Forest Algorithm

RGB: Red, Green, Blue

SALSA: Small Farms, Small Food Business and Sustainable Food Security Project

SDGs: Sustainable Development Goals

SGM: Standard Gross Margin

UA: User's Accuracy

UAA: Utilized Agricultural Area

This publication was realised with the technical support of

Food and Agriculture Organization of the United Nations

and with the collaboration of the University of Évora

salsa
small farms
small food businesses and
sustainable food security

Stay Connected

www.salsa.uevora.pt/en/

www.fao.org/in-action/small-farms-businesses-sustainable-food-nutrition

Contacts: Research and Extension Unit, Office of Innovation, Food and Agriculture Organization of the United Nations (FAO) AGDR-Chief@fao.org

Teresa Pinto-Correia, Universidade de Évora, Portugal mtpc@uevora.pt

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement No. 677363