

Chapter 20

Copernicus Data and CAP Subsidies Control



Olimpia Copăcenaru, Adrian Stoica, Antonella Catucci, Laura De Vendictis, Alessia Tricomi, Savvas Rogotis, and Nikolaos Marianos

Abstract This chapter integrates the results of three pilots developed within the framework of the Horizon 2020 DataBio project. It aims to provide a broad picture of how products based on Earth Observation techniques can support the European Union's Common Agricultural Policy requirements, whose fulfillments are supervised by National and Local Paying Agencies operating in Romania, Italy and Greece. The concept involves the use of the same data sources, mainly multitemporal series of Copernicus Sentinel-2 imagery, but through three different Big Data processing chains, tailored to each paying agency's needs in terms of farm compliance assessment. Particularities of each workflow are presented together with examples of the results and their accuracy, calculated by validation against independent sources. Business value aspects for each use case are also discussed, emphasizing the way in which the automation of the CAP requests verification process through satellite technologies has increased the efficiency and reduced cost and time resources for the subsidy process. We end the chapter by highlighting the benefits of continuous satellite tracking as a substitute, but also complementary to the classical field control methods, and also the enormous potential of Earth Observation-based products for the agri-food market.

20.1 Introduction, Motivation, and Goals

In the framework of European Union (EU) common agricultural policy (CAP), farmers can have access to subsidies that are provided through paying agencies and authorized collection offices operating at national level or regional level [1]. For the

O. Copăcenaru (✉) · A. Stoica
TERRASIGNA SRL, 3 Logofatul Luca Stroici, 020581 Bucharest, Romania
e-mail: olimpia.copacenaru@terrasigna.com

A. Catucci · L. De Vendictis · A. Tricomi
e-GEOS, 965 Via Tiburtina, 00156 Rome, Italy

S. Rogotis · N. Marianos
NEUROPUBLIC SA, Methonis 6 and Spiliotopoulou, 18545 Piraeus, Greece

provision of the subsidies, paying agencies must operate several controls in order to verify the compliance of the cultivation with EU regulations. At present, the majority of the compliance controls are limited to a sample of the whole amount of farmers' declarations due to the increased costs of acquiring high and very high-resolution satellite imagery [2]. Moreover, they are often focused on a specific timeframe, not covering the whole lifecycle of the agricultural land plots during the year.

However, EU Regulation No. 746 of 18 May 2018 [3, 4] introduced the option for member states, starting from the 2018 campaign, to use an alternative methodology to that of field controls, using information from Copernicus Sentinel satellites, possibly supplemented by those of EGNOS/Galileo. Thus, paying agencies in several countries have set strategic targets to implement CAP subsidies control systems based on cost-efficient collection and processing of earth observation data [5] and efficiently converting them into added value operational services that can be embedded into the already existing workflows and integrated with the information already available in several institutional registers.

Therefore, the aim of the CAP support initiatives within the DataBio project was to provide products and services tuned in order to fulfill the requirements for the 2015–20 CAP [6], improve the CAP effectiveness, leading to a more accurate, and complete farm compliance evaluation provided to paying agencies operating in three EU countries: Greece, Italy, and Romania.

The technological core competency lies mainly in the implementation of specialized highly automated big data processing techniques, particularly based on multi-temporal series of Copernicus Sentinel-2 data, and directly addresses the CAP demands for agricultural crop-type identification, systematic observation, tracking, and assessment of eligibility conditions over the agricultural season.

The final products are tailored to the specific needs of the end-users and demonstrate the implementation of functionalities that can be used for supporting the subsidy process in verifying specific requests set by the EU CAP.

20.2 Pilot Set-Up

As the main goal of the approach was to provide services in support to the national and local paying agencies and the authorized collection offices for a more accurate and complete farm compliance evaluation, the pilot included trial stages in three different areas of interest.

- In **Romania**, TERRASIGNA ran CAP support monitoring service trials for a 10,000 km² area of interest (AOI) in the southeastern part of the country, thus aiming to provide crop-type maps for a large area, characterized by geographical variability and a broad number of crops, distributed over diverse locations and including small and narrow plots, making use of the Copernicus Sentinel-2 spatial and temporal resolution. Initially, the selection of the 10,000 km² AOI was done by performing a multi-criteria analysis based on three main elements: plots' size

(a minimum degree of land fragmentation was desirable in order to properly test the methodology), crops diversity (the selected area included a large selection of agricultural crop types), and accessibility (any point or parcel within the area had to easily be accessed during field campaigns to collect observations needed for validation). For the 10,000 km² area of interest, intersecting three Sentinel-2 granules (35TLK, 35TMK, 35TNK), more than 150,000 plots of different sizes have been analyzed during each agricultural season. The analysis performed included parcels of over 0.3 ha, regardless of shape. Of course, the 10-m spatial resolution made the narrower parcels difficult to properly label. Starting from the 2018 agricultural season, TERRASIGNA has extended its CAP-related services and has monitored the declarations for the entire agricultural area of Romania, exceeding 9 million ha and corresponding to more than 6 million plots of various sizes and shapes, distributed across the 41 Sentinel-2 scenes, projected in 2 UTM zones, that intersect the territory of Romania (Fig. 20.1). The main end user was APIA—the Romanian National Paying Agency.

- In **Italy**, e-GEOS sets up a methodology that has been tested and applied for a 50,000-ha area of interest in the region of Veneto, Verona Province (Fig. 20.2), where the land parcel identification system (LPIS) 2016 data was available. The approach was based on the computation of markers, in relation to predefined scenarios in terms of crop types and reference periods for agricultural practices. It aimed to demonstrate and detect LPIS anomalies concerning crop types or crop families, with respect to the last update of the farmer's declaration integrated in the geospatial aid application (GSAA), and to re-classify the parcel itself. The end user, in this case, was AVEPA Paying Agency (Agenzia Veneta per I Pagamenti in Agricoltura), operating at regional level in one of the most important agricultural regions in Italy.
- In **Greece**, NEUROPUBLIC tested and evaluated a set of EO-based services designed to support specific needs of the CAP value chain stakeholders, for an area of interest covering 50,000 ha of annual crops with an important footprint in the Greek agricultural sector (rice, wheat, cotton, maize, etc.), located in the greater area of Thessaloniki (Fig. 20.3). The main stakeholders of the pilot activities were the farmers from the engaged agricultural cooperatives in the pilot area and GAIA EPICHEIREIN, that had a supporting role in the farmers' declaration process through its farmers service centers (FSCs). CSEM and FRAUNHOFER were also involved in the pilot, providing their long-standing expertise in the technological development activities. The pilot aimed at supplying EO-based products and services designed to support key business processes, including the farmer decision-making actions during the submission of aid application, and more specifically leading to an improved "greening" compliance in terms of crop diversification, which acts as a driver toward more sustainable ecosystems. Greening conditions dictate that farms with more than 10 ha of arable land should grow at least two crop types, while farms with more than 30 ha are required to maintain more than three crop types. The main crop type is not expected to cover more than 75% of the arable land. The ambition of this pilot case was to effectively deal with CAP demands for agricultural crop-type identification,

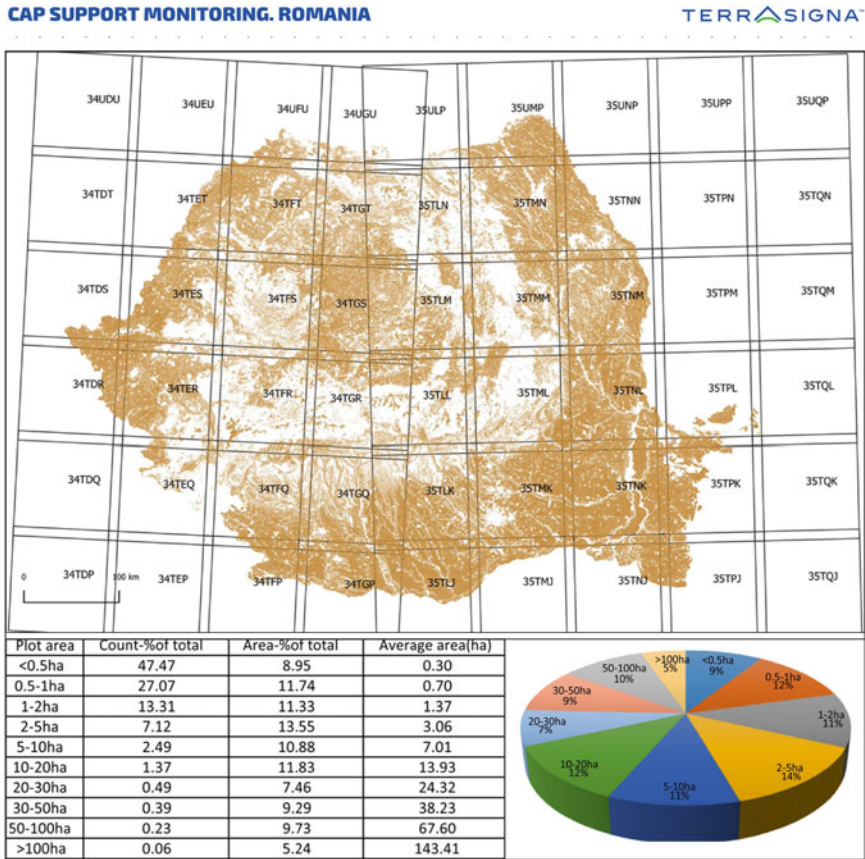


Fig. 20.1 Romania—total declared area and number of plots registered for CAP support (2019). Alphanumeric in the cells represent Sentinel-2 tiles. *Data source* Agency for Payments and Intervention in Agriculture (APIA), Romania

systematic observation, tracking, and assessment of eligibility conditions over a period of time.

20.3 Technology Used

20.3.1 Technology Pipeline

While the overall objective was similar, providing CAP-related services tailored to the specific needs of different stakeholders in charge of agricultural subsidies management, the three different approaches were based on technologies that have both similarities and important differences. Therefore, while the data collection and data

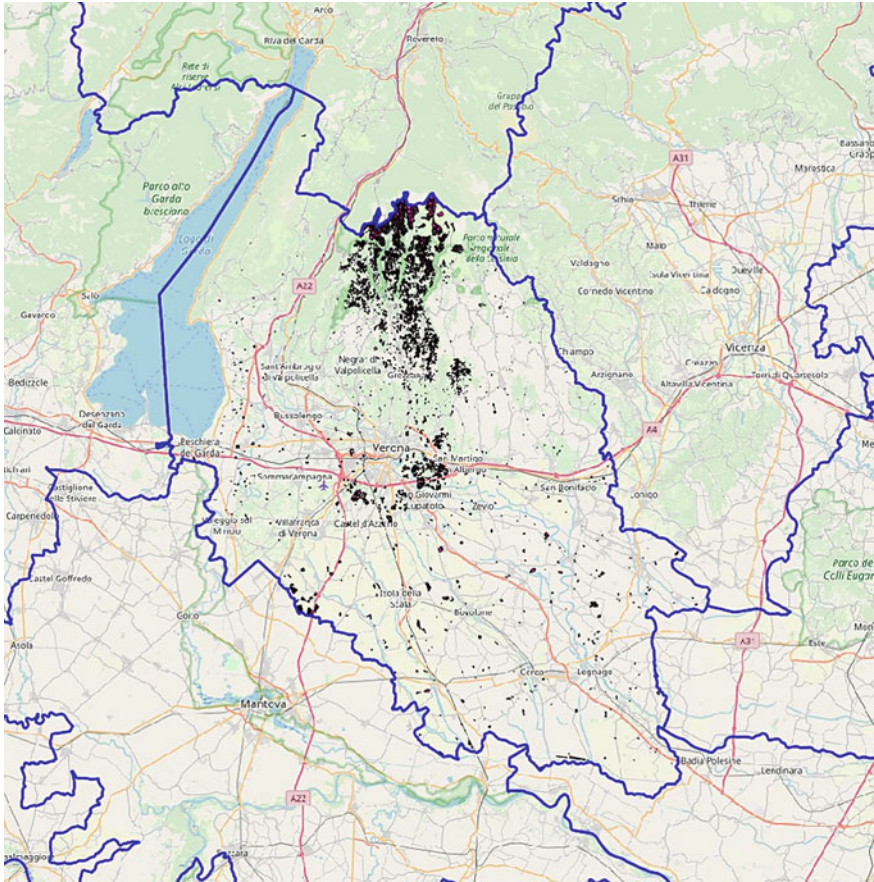


Fig. 20.2 Geographical distribution of the parcels analyzed within the trial stage in Italy (highlighted in black)

preparation phases follow very similar workflows, the data processing and analysis are based on separate technology pipelines (Fig. 20.4).

Technology Pipeline for the Trial Stage in Romania

For the trial stage in Romania, TERRASIGNA proposed an in-house developed fuzzy-based technique for crop detection and monitoring, based on combined free and open Sentinel-2 and Landsat-8 Earth Observation data image processing, data mining, and machine learning algorithms, all integrated in a toolbox for crop identification and monitoring [7].

The processing chain involves a series of well-defined steps:

- image preprocessing (numerical enhancements for Sentinel-2 and Landsat-8 scenes, ingestion of external data, clouds and shadows masking);



Fig. 20.3 Geographical distribution of the parcels that take part to the Greek pilot activities (highlighted with yellow color)

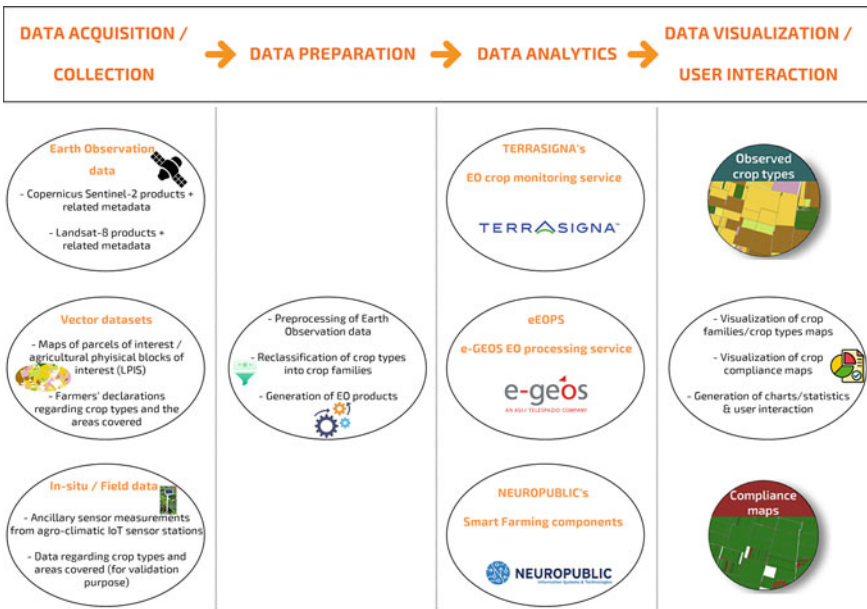


Fig. 20.4 Generic technology pipeline for the three CAP support trial stages

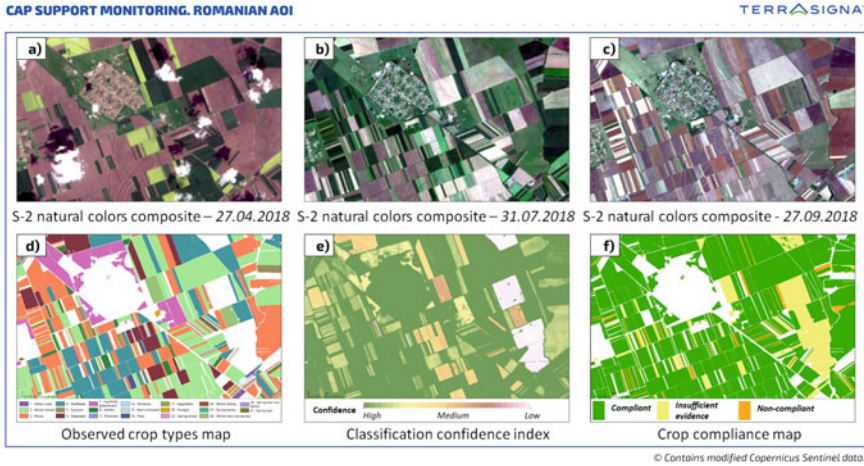


Fig. 20.5 Romania—example of CAP support analysis results. **a–c** Sentinel-2 natural color mosaics (27.04.2018, 31.07.2018, 27.09.2018); **d** observed crop types map; **e** classification confidence index; **f** crop compliance map

- individual scene classification;
- the use of unsupervised machine learning techniques in order to obtain the crop probability maps at scene level;
- time series analysis, making the system capable of recognizing several types of crops, of the order of several tens and allowing the generation of overall crop probability maps and derived products.

The developed toolbox allows the automatic calculation of the following products (Figs. 20.5 and 20.6):

- (1) Maps with the main types of crops, for a completed annual agricultural cycle;
- (2) Intermediate maps with the main types of crops, during an ongoing annual agricultural cycle (which may serve as early alarms for non-observance of the declared crop type);
- (3) Layers of additional information, showing the classification confidence index for the crop-type maps computed (values closer to 1 show higher trust levels for the assessed parcels);
- (4) Maps with the mismatches between the type of crop declared by the farmer and the one observed by the application;
- (5) Lists of parcels with problems, in order of the surfaces affected by inconsistencies, according to the data in product 4 above;
- (6) RGB backgrounds with mediated aspect, uncontaminated by clouds and shadows, computed for a period of time, with national coverage. The computed synthetic images use the principle of weighted mediation, in a fuzzy logic, which guarantees a superior visual quality; they have a very natural look,

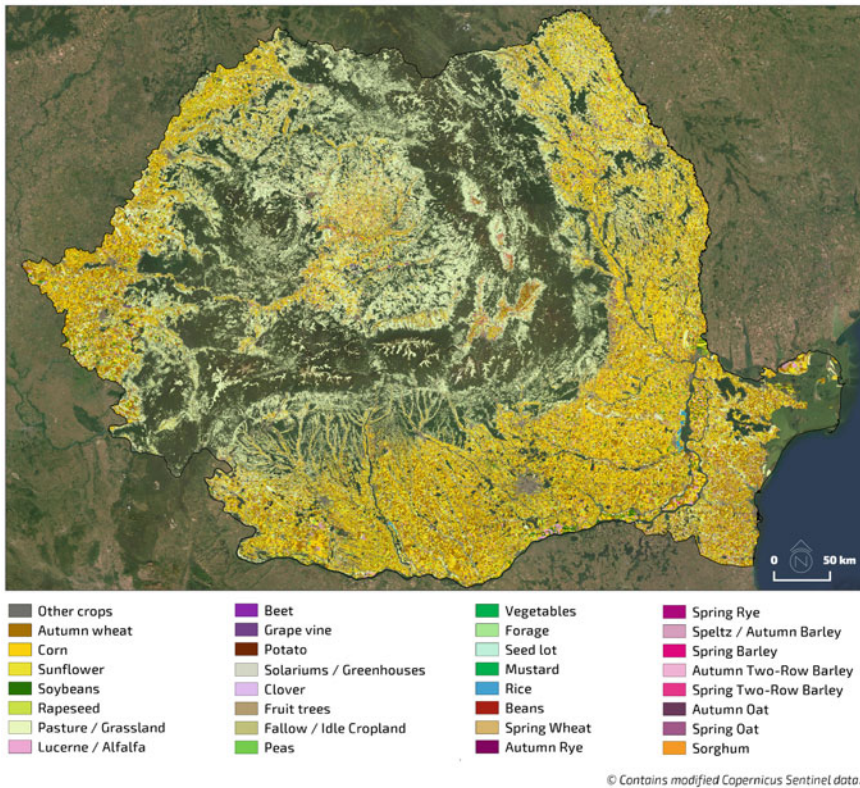


Fig. 20.6 Romania—observed crop-type map (2019) for the entire territory of the country, showing the 32 crop types that the algorithm is able to recognize, summing more than 97% of the total declared area in Romania

similar to a unique scene, however without the image being associated with a moment of time;

- (7) RGB mosaics uncontaminated by clouds and shadows, computed for a period of time, with national coverage;
- (8) NDVI maps uncontaminated by clouds and cloud shadows, computed for a period of time, with national coverage;
- (9) Early discrimination maps between winter and summer crops.

Technology Pipeline for the Trial Stage in Italy

For the trial stage in Italy, a set of markers have been computed in relation to predefined scenarios in terms of crop families and reference periods during which agricultural practices have been defined. The methodology is working at parcel level, therefore computing several markers for each parcel depending on the specific crop family.

The full list of tuned markers includes plowing, vegetation presence\growing, harvesting and mowing. However, considering the typical phenological cycle and the agricultural practices for each crop class, not all the markers have been computed for all crop classes. For example, the markers considered for wheat (autumn–winter crop family) are plowing, vegetation presence/growing, harvesting and mowing, while for permanent grassland, only the presence/growing and mowing markers have been computed [7].

For the definition of markers, it should be considered that each of them should be defined according to the geographic location and specific algorithms and related parameters should be identified, therefore requiring a proper tuning by leveraging on time series analysis. This operation is supported by the analysis, for each crop family, of the spectral behavior along time, in order to identify from a mathematical point of view, markers related to specific activities.

For example, Fig. 20.7 shows the NDVI temporal trend of a corn parcel in the center of Italy, from which it is possible to identify, together with the support of false-color images, the relevant stages in the phenological cycle:

- Plowing: between January and April (false-color image A);
- Presence\growth: between April and August (false-color image B);
- Harvesting: between July and September (false-color image C);
- False-color image D shows the parcel after the harvesting.

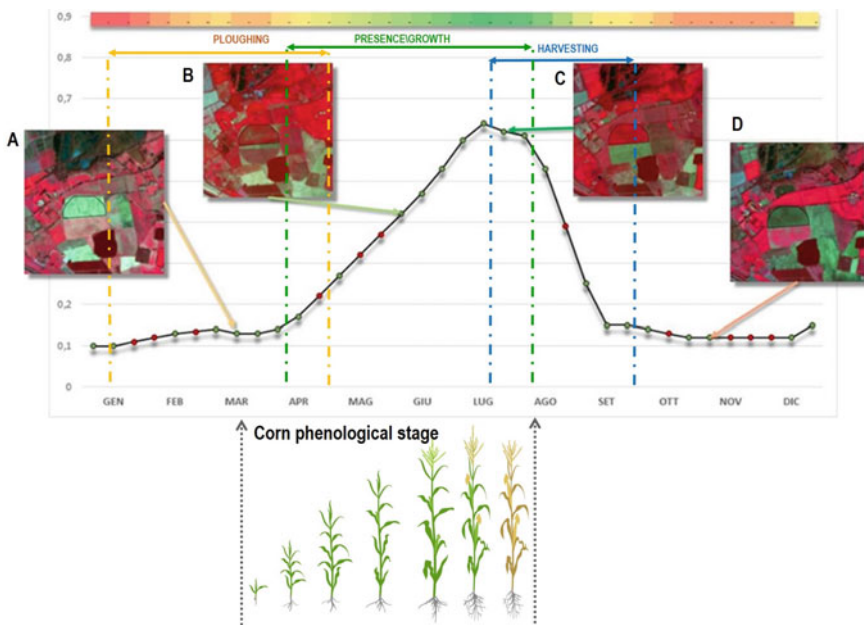


Fig. 20.7 NDVI temporal trend with identification of relevant stages in the phenological cycle

Once the markers are tuned according to type of interest, relevant periods, thresholds on NDVI values and geographic location for each crop type/families, they should be detected using a proper algorithm operating on time series. The results of the marker computation (positive/negative) can feed the internal workflow of the paying agencies, by:

- supporting the analysis and computation of parcel compliance versus administrative regulations of farmers' applications for subsidies;
- supporting the detection of LPIS anomalies (incorrect classification or update need) and then re-classification, testing the validity of markers of the other macro-classes.

Technology Pipeline for the Trial Stage in Greece

For the trial stage in Greece, a set of data collection, processing, and visualization components has been used to technically support the pilot activities [7]. More specifically, the following technological components should be acknowledged:

In terms of *data collection*:

- In-situ telemetric stations provided by NP, so-called gaiatrons, that collect ancillary weather data;
- Modules for the collection, preprocessing of Earth Observation products, the extraction of higher-level products, and assignment of vegetation indices at parcel level.

In terms of *data processing*:

- GAIABus DataSmart Machine Learning Subcomponent (NP), supporting EO data preparation and handling functionalities, multi-temporal object-based monitoring and modeling and crop-type identification;
- GAIABus DataSmart Real-time streaming Subcomponent (NP), supporting:
 - Real-time data stream monitoring for NP's gaiatrons installed in the pilot sites;
 - Real-time validation of data;
 - Real-time parsing and cross-checking.
- Neural Network Suite (CSEM), used as a machine learning crop identification system for the detection of crop discrepancies;
- Georocket, Geotoolbox and SmartVis3D (FRAUNHOFER), having a dual role: a back-end system for Big Data preparation, handling fast querying and spatial aggregations, as well as a front-end application for interactive data visualization and analytics.

In terms of *data visualization*:

- Neurocode (NP), the main component, allowing the creation of the main pilot UIs in order to be used by the end-users (FSCs of GAIA EPICHEIREIN);
- Georocket (FRAUNHOFER), an additional DataBio component providing information visualization functionalities.

20.3.2 Data Used in the Pilots

All the three trials (Romania, Italy, and Greece) aimed to demonstrate the advanced capabilities of Earth Observation data in monitoring agricultural areas [8].

Therefore, the input data consisted in:

- Sentinel-2 and Landsat-8 optical satellite data;
- The declarations of the farmers regarding cultivated crops and areas covered;
- The map of the parcels of interest or the map of the physical blocks of interest;
- List of crop codes used;
- List of crop classes to be followed (LCCF, i.e., very related groups of crops, which have similar aspect and phenological behavior);
- Ancillary sensor measurements from agro-climatic IoT sensor stations (used for the trial stage in Greece);
- A collection of a validation dataset, representative for the crop types/crop families distribution, derived from very high-resolution imagery (used for the trial stages in Romania and Italy).

20.3.3 Reflections on Technology Use

Reflections on Technology Use for the Trial Stage in Romania

The crop monitoring technology developed by TERRASIGNA is able to recognize a large number of crops families, of the order of tens. For Romania, it addressed the first most cultivated 32 crop families (according to the information provided by the National Paying Agency), which together cover more than 97% of the agricultural land [7]. The success rate in recognition was not equal between crops families, but an overall performance of 98.3% (Table 20.1) was obtained for the first most important 8 crops (winter wheat, sunflower, maize, green peas, winter barley, meadows and pastures, rapeseed, soybean). The countercheck data was obtained using a manual classification of a statistical sample in a test zone of the size of a Sentinel granule, supplemented with field-collected data regarding cultivated crop types and areas covered. The performance proved to be quite uniform reported to the size of the plots and remained high even for parcels smaller than 1 ha (Table 20.1).

At the moment, taking into account the agricultural specificity of Romania, defined by excessive land fragmentation, as a result of the existing legislation, the developed technology is using only optical data, consisting in both Copernicus Sentinel-2 and Landsat-8 imagery. According to the Romanian National Paying Agency, out of the total of 6 million plots for which payments have been granted, 2.7 million plots have an area smaller than 0.5 ha (44% of the total number), while 1.8 million plots consist of an area between 0.5 and 1 ha. Therefore, the small narrow plots are not suitable for SAR analysis for crop-type identification, taking into account the noise level, despite the good spatial resolution of Sentinel-1 images. Moreover, as stated before, in terms of overall accuracy (OA), the classification result using only Sentinel-2

Table 20.1 Romania—results of the validation based on independent data consisting of very high-resolution imagery and field-collected data

	<1 ha	1–1.5 ha	1.5–2.5 ha	2.5–5 ha	5–10 ha	10–20 ha	>20 ha	Overall performance
Winter wheat	<i>99.1%</i> N = 130 ha	<i>98.6%</i> N = 226 ha	<i>97.5%</i> N = 619 ha	<i>98.2%</i> N = 1,919 ha	<i>98.5%</i> N = 3,073 ha	<i>98.3%</i> N = 4,494 ha	<i>99.4%</i> N = 22,208 ha	98.7% Omissions: 0.73%
Maize	<i>99%</i> N = 22 ha	<i>94.4%</i> N = 30 ha	<i>90.7%</i> N = 81 ha	<i>88.1%</i> N = 216 ha	<i>90.1%</i> N = 396 ha	<i>93.1%</i> N = 679 ha	<i>99.1%</i> N = 4,877 ha	93.7% Omissions: 4.93%
Sunflower	<i>97.8%</i> N = 25 ha	<i>97.8%</i> N = 41 ha	<i>99.6%</i> N = 109 ha	<i>96.7%</i> N = 320 ha	<i>99.3%</i> N = 664 ha	<i>99.1%</i> N = 1,008 ha	<i>99.5%</i> N = 3,663 ha	98.8% Omissions: 3.99%
Soybean	<i>100%</i> N = 4 ha	<i>92.1%</i> N = 18 ha	<i>90.1%</i> N = 41 ha	<i>99.6%</i> N = 186 ha	<i>99.9%</i> N = 558 ha	<i>100%</i> N = 800 ha	<i>100%</i> N = 2,370 ha	99.3% Omissions: 2.12%
Rapeseed	<i>99.7%</i> N = 77 ha	<i>99.6%</i> N = 93 ha	<i>98.9%</i> N = 268 ha	<i>99.6%</i> N = 811 ha	<i>98.6%</i> N = 1,107 ha	<i>99.4%</i> N = 1,633 ha	<i>99.8%</i> N = 8,346 ha	99.5% Omissions: 1.16%
Pastures	<i>98.1%</i> N = 111 ha	<i>97.3%</i> N = 165 ha	<i>99.2%</i> N = 393 ha	<i>99.2%</i> N = 1,199 ha	<i>99.1%</i> N = 2,161 ha	<i>99.5%</i> N = 3,306 ha	<i>99.5%</i> N = 7,302 ha	99.3% Omissions: 2.64%
Peas	n.a.	<i>96.5%</i> N = 2.4 ha	<i>100%</i> N = 4 ha	<i>100%</i> N = 23 ha	<i>100%</i> N = 75 ha	<i>86.9%</i> N = 93 ha	<i>100%</i> N = 348 ha	97.4% Omissions: 1.44%
Winter barley	<i>100%</i> N = 10 ha	<i>100%</i> N = 3.5 ha	<i>91.5%</i> N = 24 ha	<i>95.3%</i> N = 58 ha	<i>100%</i> N = 1,477 ha	<i>100%</i> N = 442 ha	<i>100%</i> N = 1,682 ha	98.7% Omissions: 2.79%
All crops	97.8% N = 336 ha 434 plots	97.4% N = 543 ha 436 plots	96.6% N = 1446 ha 722 plots	97.5% N = 4454 ha 1,216 plots	98% N = 8075 ha 1,109 plots	98.2% N = 12,261 ha 865 plots	98.2% N = 50,106 ha 1,060 plots	98.3%

The italic values represent the percentage of accuracy for the specific size classes analyzed for each crop category

The bold values represent the overall percentage of accuracy for each crop category and also the overall percentage of accuracy for each size class

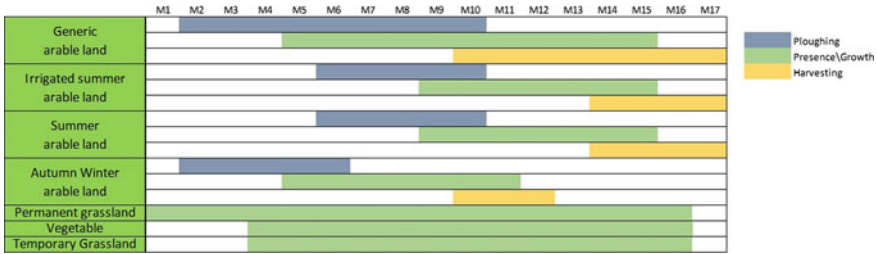


Fig. 20.8 Example of predefined scenarios regarding agricultural practices for the crop categories analysed

imagery reached 0.98. Thus, a major increase in overall accuracy using SAR data was not foreseen.

Reflections on Technology Use for the Trial Stage in Italy

The crop monitoring technology developed by e-GEOS for the trial stage in Italy was based on NDVI profile trends [7], which allowed the computation of a set of markers related to agricultural practices that should take place (e.g., plowing, vegetation presence/growth, and harvesting), in relation to predefined scenarios (Fig. 20.8), in terms of:

- selected macro-crop type;
- reference periods;
- NDVI thresholds.

At the beginning of the trial activities, the LPIS crop types have been aggregated in macro-classes (23 families) and the predefined scenarios have been tuned for the seven classes suitable for the automatic detection of anomalies and reclassification, based on the Sentinel-2 time series.

Analyzing their distribution and considering that the largest part (about 67%) of the agricultural crop families in the AOI belongs to 2 main groups, permanent grassland and arable land, only the crop families of these 2 groups have been considered in order to test the algorithm of anomalies detection and re-classification at macro-class level.

The markers computed in relation to predefined scenarios have been implemented in a decision model to verify their correct classification. The model has been run for each parcel of the macro-classes considered as suitable for the automatic detection of anomalies. Examples of parcels for which the original macro-class has been confirmed or detected as anomalous through the automatic analysis based on the related markers are displayed in Fig. 20.9.

Parcels detected as anomalous have been automatically re-classified testing the validity of the markers of the other macro-classes, thus updating the LPIS in terms of macro-classes (Fig. 20.10).

As expected in the arable land area, due to the usual crop rotation practice, the largest part of parcels changed their agricultural use between 2016 and 2018

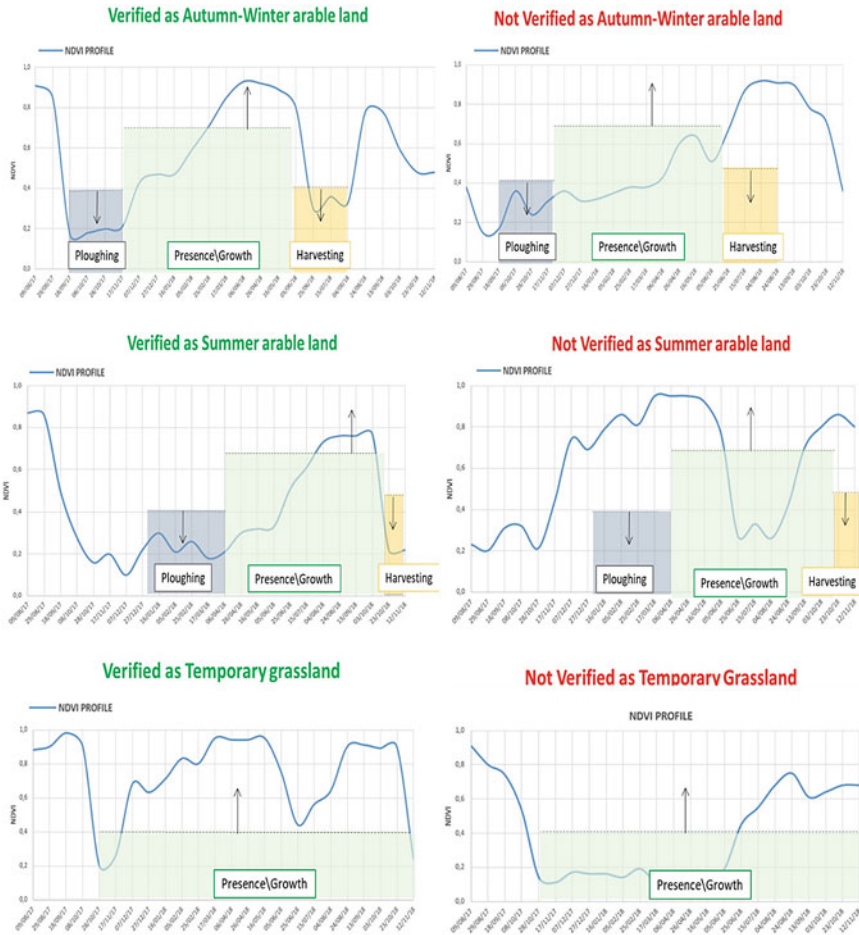


Fig. 20.9 Example of marker analysis based on predefined scenarios

(Fig. 20.11). In most cases, it is simply a change from winter–autumn to summer or temporary grassland and vice versa (Fig. 20.12).

The results are confirmed by the pie charts (Figs. 20.13 and 20.14) that describe, for different crop families (autumn–winter arable land, summer arable land and irrigated summer arable land) the percentage of parcels for which the crop family has been confirmed (in green) and the percentages of anomalous parcels, re-classified as other crop families.

Irrigated summer arable land parcels (e.g., rice paddies) are mostly confirmed (few anomalies) probably because these types of crop field, supported by irrigation systems, are not subject to crop rotations (Fig. 20.15).

In terms of permanent grassland areas, as expected, the percentage of anomalies is meaningful lower, considering the fact that usually the agricultural use of these

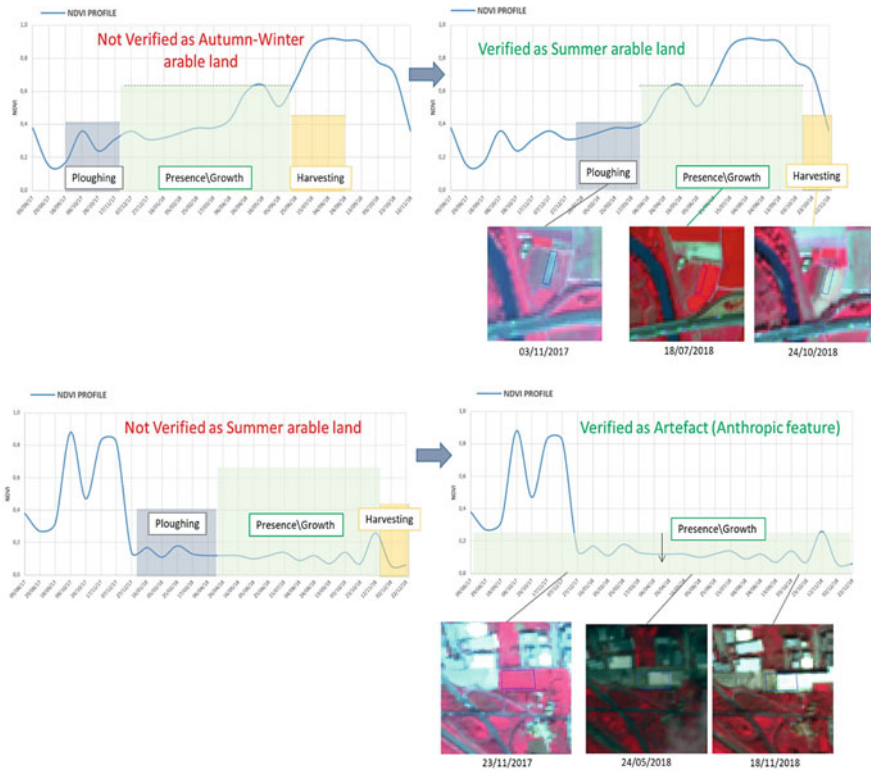


Fig. 20.10 Examples of non-compliant (left) and re-classified (right) parcels



Fig. 20.11 LPIS arable land parcels classified as verified (green), anomalous (red) and not analysed (gray)

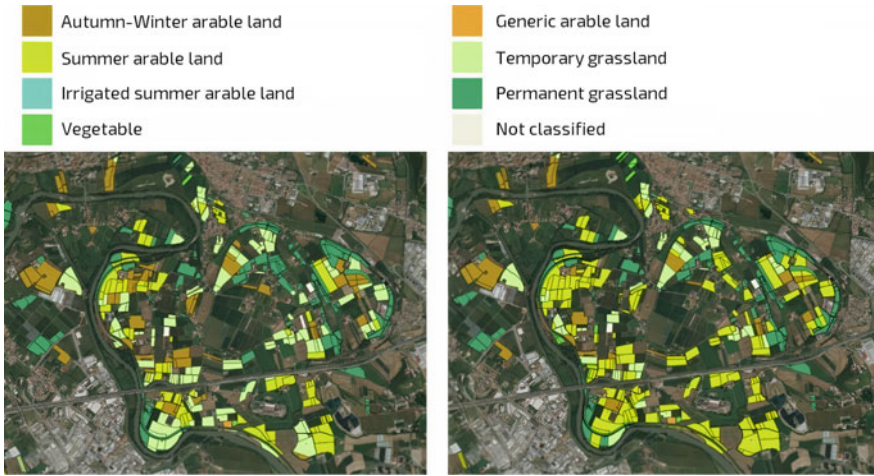


Fig. 20.12 LPIS arable land parcel classes in 2016 (left) versus 2018 (right), after re-classification of anomalous parcels

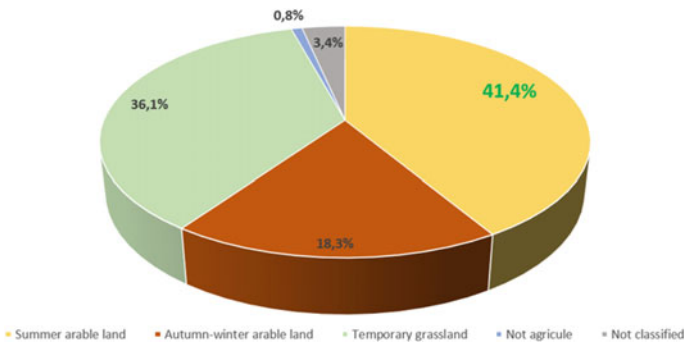


Fig. 20.13 2016 LPIS summer arable land parcels updated to 2018

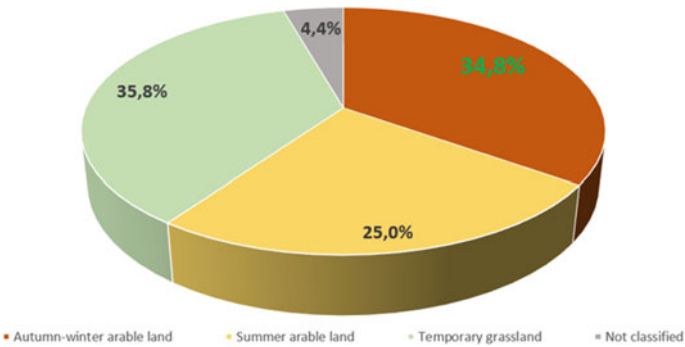


Fig. 20.14 2016 LPIS winter–autumn arable land parcels updated to 2018

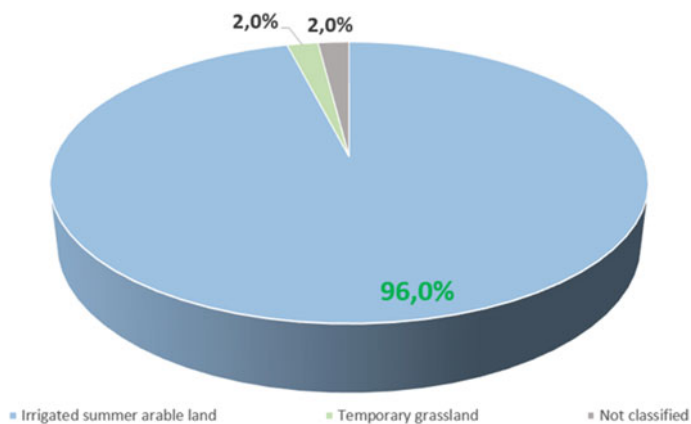


Fig. 20.15 2016 LPIS irrigated summer arable land parcels updated to 2018

parcels is stable for several years (a grassland field is defined as permanent if it is not plowed for 5 years, at least) (Figs. 20.16 and 20.17).

The accuracy of the methodology proposed for the LPIS anomalies detection and reclassification has been assessed through a validation activity based on data extracted from very high-resolution imagery. About 1000 parcels have been considered for

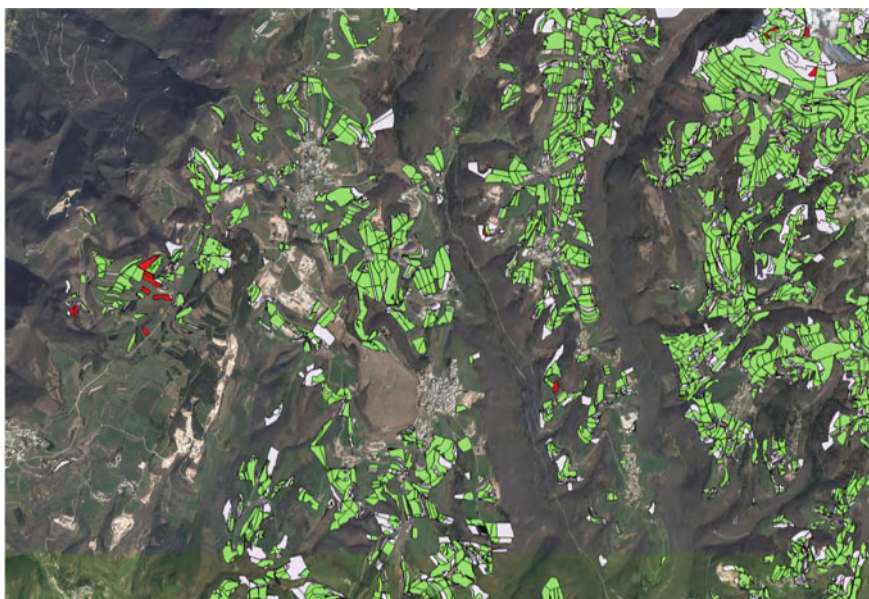


Fig. 20.16 LPIS permanent grassland parcels classified as verified (green), anomalous (red) and not analysed (gray)

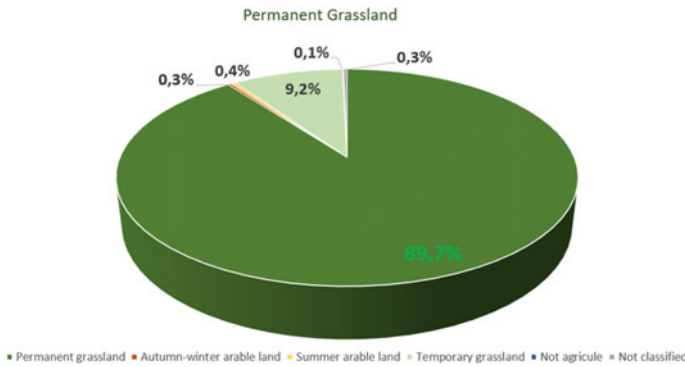


Fig. 20.17 2016 LPIS permanent grassland parcels updated to 2018

Table 20.2 Results of the validation based on reference data extracted from very high-resolution imagery

Crop family	Parcel number	Accuracy (%)
Autumn–winter arable land	26	84.6
Summer arable land	55	96.4
Permanent grassland	973	96.5
Temporary grassland	73	38.2

the accuracy assessment (Table 20.2). The resulting validation dataset consisted of four main crop families: autumn winter arable land, summer arable land, permanent grassland, and temporary grassland, reflecting the crop families’ distribution over the entire area. Other crop families, considered statistically insignificant in terms of number of parcels, have not been taken into account in the accuracy assessment.

The results reveal very high accuracy for permanent grassland and summer arable land (more than 95%), high for winter arable land (85%). However, the computed accuracy for the temporary grassland crop family with respect to the farmers’ declarations is just around 40%. The remaining 60% mis-classified parcels are distributed, according to farmers’ declarations, mainly as permanent grassland (33%) and they require an additional refinement of marker rules in order to improve the accuracy.

The performances will be further tested in wider areas in order to evaluate the potential to be used in operative scenarios.

Reflections on Technology Use for the Trial Stage in Greece

In Greece, “greening” compliance was assessed for the 2019 cultivation year and the respective aid applications [7]. The farmers that could benefit from the methodology were the ones holding parcels larger than 10 ha, eligible for checks for greening requirements related to crop diversification. The crop types that have been modeled by the GAIABus DataSmart Machine Learning Subcomponent were seven (7) in total and more specifically: wheat, cotton, maize, tobacco, rapeseed, rice, and sunflower.

Table 20.3 Normalized crop classification confusion matrix (horizontal axis corresponds to the true label, whereas the vertical one to the predicted label)

	Maize	Cotton	Rapessed	Sunflower	Tobacco	Rice	Wheat
Maize	0.932	0.004	0.022	0.021	0.004	0.003	0.005
Cotton	0.009	0.954	0.006	0.019	0.079	0.002	0.008
Rapeseed	0.000	0.000	0.713	0.000	0.000	0.000	0.000
Sunflower	0.023	0.007	0.019	0.824	0.061	0.000	0.017
Tobacco	0.000	0.001	0.000	0.005	0.712	0.000	0.001
Rice	0.002	0.001	0.000	0.000	0.000	0.994	0.000
Wheat	0.032	0.032	0.239	0.131	0.144	0.000	0.968

If seen as a multiclass classification problem, the performance of the trained crop models to the 2019 testing data are offered at the confusion matrix, in Table 20.3.

Using the trained models as the backbone of the CAP support methodology, the assessment of “greening” compliance was conducted over 2019s aid applications. A traffic light system was employed to inform the farmers that there could have been a problem within their declarations:

- if the confidence level of the classification result was $>85\%$ and the declared crop type of the farmer was confirmed by the classification, traffic light should be green;
- if the confidence level of the classification result was $<85\%$ and the declared crop type of the farmer was confirmed by the classification, traffic light should be yellow;
- if the declared crop type of the farmer was not confirmed by the classification, traffic light should be red.

According to this approach, the farmer is more protected in order to receive the payment as robust and reliable feedback is provided to him/her. The farmer is notified for issues (especially when the main crop seems to cover more than 75% of the cultivated land—mandatory condition for ensuring crop diversification) that put at risk his/her eligibility for greening compliance, thus contributing to raising awareness and allowing follow-up activities to be taken. An example regarding greening eligibility assessment is shown in Table 20.4.

Moreover, in order to support the Greek pilot activities, an integrated analytics platform has been finalized and deployed (Fig. 20.11). The use of machine learning

Table 20.4 Greening eligibility assessment using a traffic light system (table and map projection)

Parcel	Crop group		DataBio Assessment		Traffic Light	Area (ha)	Map projection
	Declared	Detected	Status	Result			
001	Wheat	Wheat	Assessed	Compliant		2.08	
002	Wheat	Wheat	Assessed	Compliant		1.67	
003	Maize	Wheat	Assessed	Not compliant		1.1	
004	Maize	Maize	Assessed	Insufficient evidence		1.46	
005	Wheat	Wheat	Assessed	Insufficient evidence		1.25	
006	Cotton	Cotton	Assessed	Compliant		0.82	
007	Cotton	Wheat	Assessed	Not compliant		0.73	
008	Wheat	Wheat	Assessed	Compliant		1.88	
					Total	10.99	

services provided a proof of concept for its use in CAP support scenarios. FRAUNHOFER was responsible for the development of the UI, integrating pixel heat maps from the different classifiers and information visualization capabilities. A CSEM developed system for the management of machine learning models was used to facilitate the simple and retraceable management of models. RESTful services, combined with security features in the form of JSON Web Tokens (JWT) and encryption with Hypertext Transfer Protocol Secure (HTTPS), were implemented and integrated into the service. The service has also been containerized to allow simple deployment. This service enables the communication with FRAUNHOFER's component GeoRocket and UI for on-demand crop-type classification, in both pixel and parcel levels (Fig. 20.18).

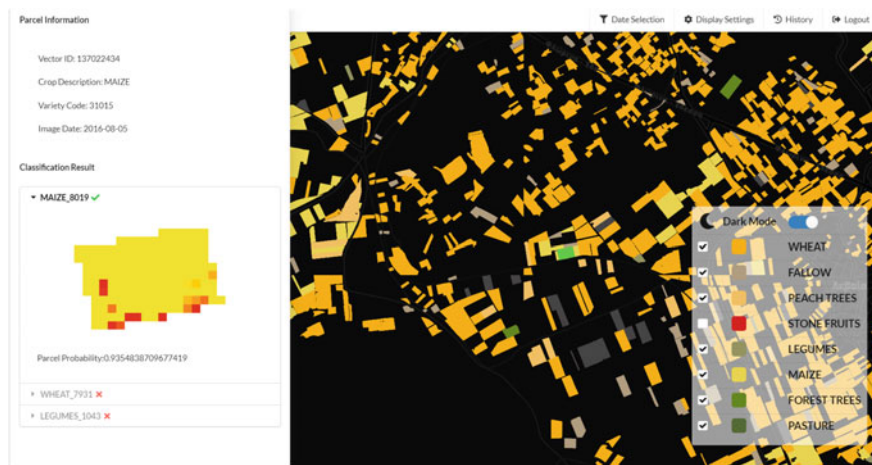


Fig. 20.18 User interface created by FRAUNHOFER for the Greek CAP support trial. The user interface integrates CSEM’s classification results into pixel heat maps

20.4 Business Value and Impact

20.4.1 Business Impact of the Pilot

All the three CAP support trials developed within the DataBio project were tailored according to the needs of specific end-users (national and regional paying agencies), and, therefore, the business impact of the pilots is closely related [9]. The added value of the three pilots effectively consists in the increase of efficiency that the payment authorities and other end-users experience in using satellite monitoring and big data technologies.

Business Impact for the Trial Stage in Romania

The possibilities for exploitation of the project’s result for TERRASIGNA focused on proving a concept and attracting a long-term collaboration with the National Agency for Payments and Investments in Agriculture (APIA), holding responsibility in Romania of the implementation of CAP mechanisms for direct payments. A cooperation agreement was signed with the agency, in order to offer and test the results of the pilot—crop compliance maps in support of APIA’s activity of monitoring the subsidies payments [9].

The CAP support pilot, through its EO crop monitoring component, offered the stakeholder the possibility to check, in a more efficient way, the compliance between the declarations made by the farmers in request of the subsidy payments and the real crop in the fields. While currently a minimum of 5% from the applications is cross-checked either by field sampling or by remote sensing, the developed methodology

allowed checking the compliance of the declarations submitted by the farmers for all agricultural parcels with individual areas exceeding 0.3 ha.

Moreover, as the service automatically detects and signals the parcels with the highest probability not to grow the crop declared by the farmer, it makes the regulator's decision more efficient in selecting the parcels for field control or control through very high-resolution imagery.

Business Impact for the Trial Stage in Italy

The marker-based approach proposed by e-GEOS for the trial stage in Italy has demonstrated its applicability not only for CAP monitoring, but also opening up the street for future innovation in the market. e-GEOS is active in the agro-insurance and CAP market segments, with a network of actual and potential customers and users, including the paying agencies operating in different regions of the country. Therefore, the CAP-related developed products have been strongly related to the Italian agricultural policy needs [9].

Business Impact for the Trial Stage in Greece

GAIA EPICHEIREIN, through its associated network of farmer service centers (FSCs), provides collection and advisory services to the Greek Farmers concerning the submission of the aid application for direct payments, including eligibility pre-check mechanisms for error reduction and proof provision. The total number of holdings in Greece for 2016 was 686.818. GAIA subsidy services are mainly oriented to aging small-sized farmers, which own 80% of the holdings in Greece. Over the last two annual periods, GAIA EPICHEIREIN provided collection services and cross-compliance checks to 76% of the holdings. Even if GAIA EPICHEIREIN has a market share of 76%, the ongoing CAP changes and trends, the differentiations in the internal market and the new business plans for smart farming (driven by the evolution in sensor and space technology) indicate that GAIA EPICHEIREIN needs to evolve its services in order to keep its competitive advantage and sustain its market share [9].

For the Greek scenario, the offered DataBio solutions allow the farmer (beneficiary) to deal effectively with the greening requirements. More specifically, DataBio solutions will be a valuable tool within the suite of digital CAP support services offered by GAIA EPICHEIREIN's and its FSCs that support the crop declaration process. During the process and usually after the declaration period closes and error-checking tools are applied, the FSC would be able to check the farmer's claim for the greening requirements, examine the results, and inform the farmer for follow-up activities that better serve his/her interests.

Apart from the exploitation value for the partners involved, the pilot introduced concrete benefits for the farmers and the agri-food sector as well. The results of the pilot effectively showed that EO-based crop identification services, tailored for monitoring greening compliance, offered a layer of protection against errors in the declaration process which could lead to a significant financial impact for the farmer. Additionally, and from a higher level, agricultural monitoring approaches could contribute

to more efficient funding absorption, thus securing investments and progress in the agri-food sector.

20.4.2 Business Impact of the Technology on General Level

The added value of multi-temporal copernicus sentinel data and applied new technologies (automated detection and determination using machine learning) in the context of CAP support can be explained through two different, but complementary aspects:

Copernicus Sentinel Data Stream

For the first time in the history of Earth Observation, almost every single region in Europe benefits from repetitive (5 days in average) observations with solid spatial and spectral resolution. Generally speaking, this stack of information enables early usage of EO data in the agricultural season, which in turn allows the extraction of preliminary conclusions that can be used within control with remote sensing (CwRS) decision trees (e.g., detection of winter crops, post-winter water ponding).

The 10 m spatial resolution enables the survey of the smaller plots, that in many European countries (including Romania), represent a significant number of CAP applications.

The spectral resolution provides all the necessary information (visible, NIR, SWIR) for observing the crops phenology and for distinguishing additional features (e.g., water, burned area, built-up).

The “turning data into information” policy is fully exploited, by transforming the wealth of satellite and in-situ data into valued-added services based on processing and analyzing the data, monitoring changes and making the datasets comparable, integrating them with other sources and, finally, validating the results.

Finally, the no-cost policy of the EC and the unprecedented volume of data on a full, free and open basis foster new business opportunities and job creation and provide the necessary sustainability to invest in developing copernicus data-based workflows.

Application of New Technologies

The usage of time series of Sentinel-2 satellite images for crop detection can increase the results precision, as the 5-days revisit time almost triples the number of surveys compared to the Landsat feed (16-days revisit time). An automatic nation-wide technology will warrant constant quality of the results over large areas and time periods. On the short time scale, this will allow avoiding human subjectivity.

The usage of the new technology is significantly decreasing the time, money, and human power required currently by the control with remote sensing (CwRS) campaigns. Instead of one year of administrative actions (from very high-resolution data selection to the real photo interpretation), the decision maker will have the option

to solely focus on areas already marked as red lights or on areas known as prone to risk.

The usage of the early results will also enable the use of technology as a deterrent tool; e.g., the farmer's declaration lists a winter crop, and no winter crop was observed at the end of March.

20.5 How-to-Guideline for Practice When and How to Use the Technology

The three CAP support approaches use earth observation data time series, thus providing wide and repetitive homogeneous coverage, translated into an unprecedented amount of information. The technologies benefitting from these data volumes represent a solid solution for a continuous monitoring of CAP compliance. The EU Copernicus Sentinel-2 satellites hold an enhanced revisiting time, delivering regular coverage over large areas and allowing a uniform observation of the agricultural plots. The superior spectral resolution allows the identification of the phenological growth stages and the distinction between various crop types or classes.

However, the pilots also had to overcome some major drawbacks, mainly related to data fusion, georeferencing errors (deeply affecting the quality of the crop mapping results for narrow or small plots), cloud and shadow masking, or semantic confusions between crop classes.

The highly automated proposed approaches allow the implementation of big data analytics using various crop indicators, resulting in reliable, cost and time saving procedures, and allowing a more complete and efficient management of EU subsidies, strongly enhancing their procedure for combating non-compliant behaviors.

The developed techniques have undergone continuous development and improvements, are replicable at any scale level and can be implemented for any other area of interest. Any further developments of the CAP monitoring technologies will be able to provide products tuned in order to fulfill the requirements of the present and future EU common agricultural policy. This application of big data processing technologies based on copernicus sentinel data will also significantly improve the way which farmers are doing online aid applications and, for the paying agencies, will help to keep the LPIS up-to-date and to move to the new checks by monitoring.

Moreover, the Copernicus free and open data policy, together with the long-term availability certainty, are important factors that highly help the developed solutions enter the European market and trigger collaborations between government agencies (regional or national paying agencies) and private sector companies.

The DataBio European Lighthouse project offered new business opportunities and aimed to directly improve a series of CAP support activities for providing supporting tools and services, in line with the commands of the EU's new agricultural monitoring approach. The effort is expected to continue in the next years for all the three companies, setting strategic targets such as integration of information available in

several institutional registers, active use of technologically most relevant and cost-efficient remote sensing services and proactive cooperation with rural communities and farmers.

20.6 Summary and Conclusion

Common agricultural policies and activities from national and regional paying agencies can radically benefit from the use of continuous satellite monitoring instead of random and limited controls.

The DataBio European Lighthouse project, with its three different CAP Support approaches, in Romania, Italy, and Greece, demonstrates the potentiality for final users to exploit Copernicus data in the agriculture domain, a key economic sector for most of the European countries. The proposed methodologies have undergone continuous development and improvements over the last years, offering a wide range of opportunities in order to enhance the implementation of the CAP. The continuous agricultural monitoring services, based on the processing and analysis of Copernicus satellite imagery time series, are not just CAP compliance tools, but can also offer a great range of supplementary information for both public authorities and farmers and can support the set-up of more environmentally friendly and efficient agricultural practices.

This market is one of the more promising in terms of exploiting the full potential of earth observation deployment and represents a successful example of how policies and strategies drive advancement in big data processing technologies, encourage innovation within the public sector and trigger long-term private–public partnerships.

References

1. DG Agri—European Commission. (2007). *Europeans, agriculture and the common agricultural policy (CAP)*. Retrieved August 16, 2020, from https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_410_en.pdf
2. Ivanov, D. (2017). Future development of common agriculture policy after 2020. *Trakia Journal of Science*, 15(Suppl. 1), 158–162.
3. Commission Implementing Regulation (EU) 2018/746 of 18 May 2018 amending Implementing Regulation (EU) No 809/2014 as regards modification of single applications and payment claims and checks C/2018/2976 (OJ L 125, 22.5.2018, pp. 1–7).
4. Commission Implementing Regulation (EU) No 809/2014 of 17 July 2014 laying down rules for the application of Regulation (EU) No 1306/2013 of the European Parliament and of the Council with regard to the integrated administration and control system, rural development measures and cross compliance (OJ L 227, 31.7.2014, p. 69).
5. Bartolini, F., Vergamini, D., Longhitano, D., & Povellato, A. (2020). Do differential payments for agri-environment schemes affect the environmental benefits? A case study in the North-Eastern Italy. *Land Use Policy*. <https://doi.org/10.1016/j.landusepol.2020.104862>
6. Erjavec, K., & Erjavec, E. (2015). Greening the CAP—Just a fashionable justification? A discourse analysis of the 2014–2020 CAP reform documents. *Food Policy*, 51, 53–62.

7. Estrada, J., et al. (2020). DataBio deliverable 1.3—Agriculture pilot final report. Retrieved 3.11.2020 from https://www.databio.eu/wp-content/uploads/2017/05/DataBio_D1.3-Agriculture-Pilot-Final-Report_v1.2_2020-01-21_TRAGSA.pdf
8. Klien, E., et al. (2018). DataBio deliverable 5.3—EO services and tools. Retrieved 3.11.2020 from https://www.databio.eu/wp-content/uploads/2017/05/DataBio-D5.3-EO-Services-and-Tools_v1.0_2018-06-15_Fraunhofer.pdf
9. Stanoevska-Slabeva, K., et al. (2020). DataBio deliverable D7.2—Business plan v2. Retrieved 3.11.2020 from https://www.databio.eu/wp-content/uploads/2017/05/DataBio_D7.2.Business-Plan-v2_v1.2_2020-03-19_UStG.pdf

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

