Smart Pro CAP

Hercule III Programme

State-of-the-Art and future perspectives for the protection of EU financial interests through data analysis at AVEPA

Summary of a selection of experiences shared during the Study visit held in Padua on the 17th and 18th of November 2021

AVEPA – Paying Agency of Veneto Region

Table of content:

ABSTRACT **3**

PART I Mapping the areas, methods, software and resources used for the data-based risk analysis. 3

SUBSECTION 1: Sentinel data in AVEPA land monitoring 4

Background and starting point 4

Resources used. 6

Our solution 8

The idea. 8

Indexes. 10

Scope. 11

Approaches. 12

Data Model: The Classical way. 12

Data Model: The Modern way. 13

Lessons learnt 15

Idea for the future 17

SUBSECTION 2: Farm's data management in AVEPA 17

Background and starting point 18

Our solution 19

Lessons learnt 25
SUBSECTION 3: Reasonableness of machineries' costs – AVEPA's experience 28
Background and starting point 28
Our solution 30
Lessons learnt 35
SUBSECTION 4: Innovative IT approach for OTSC - the experience of ERDF Regional Operative Programme 2014-2020 38
Background and starting point 38
Our solution 39
Lessons learnt 39
PART II Experiences of using Arachne risk scoring tool **41**
The problem 41
Our starting point 41
Our solution 42
Lessons learnt 44
Conclusions **46**

ABSTRACT

A Study visit in the frame of the Smart Pro CAP Project took place in Padua (Italy) on the 19th and 20th of November 2021. Main aim of the visit was to share and to highlight good practices related to the protection of financial interests by acting at regional level by providing an in-depth overview of procedures and methods currently in place at AVEPA and to further deepen and discuss some key issues with the purpose of exploring new way forwards.

The Visit represented an opportunity to touch upon a broad variety of issues moving from the comprehensive Agency's integrated strategy bringing together obligations and initiatives related to anti-corruption, anti-money laundering and anti-fraud, its data-driven risk assessment tool for internal audits and a general approach to tackle the creation of artificial conditions.

Following these overarching introductory themes which are relevant for establishing and maintaining an effective internal control system as a first step towards the protection of financial interests and for the protection against fraud, three specific areas were deepened during the project meeting and form part of the covered topics selected for this report, notably:

1) Use of Sentinel data for control activities (Part I).

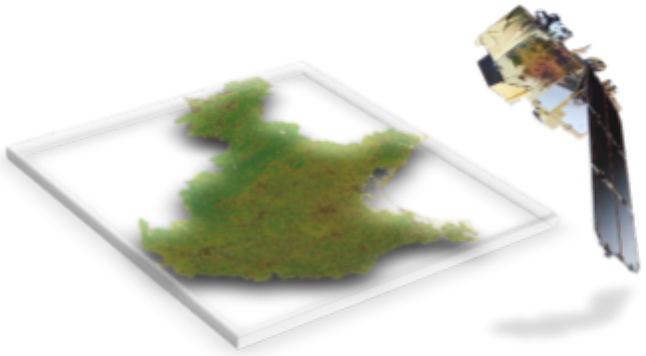
2) Management of Farm's data collected through the Farm Register (Part I).

3) Arachne – AVEPA's first experience in using this tool (Part II).

In addition to these, AVEPA's experience in dealing with the reasonableness of machineries' costs and a glimpse of an innovative method to perform OTSC remotely which has been used in the frame of ERDF Regional Operative Programme (ROP) in time of pandemics (i.e., through Video call checks) have been also presented and are deemed to be reported within this document (see Part I) as additional experiences to be taken into further considerations.

PART I Mapping the areas, methods, software and resources used for the data-based risk analysis.

SUBSECTION 1: Sentinel data in AVEPA land monitoring



This part:

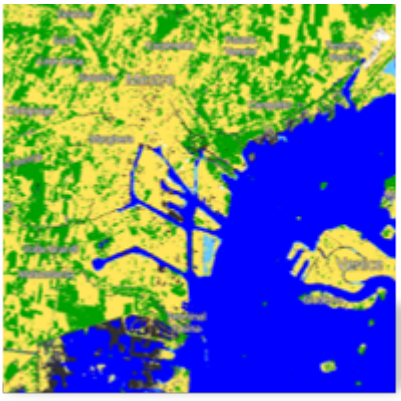
- ***Shares methodology, tools, software, resources used for the data-based geometry classification risk analysis.***
- ***Improves the awareness of the PA's staff on land monitoring and remote sensing available results, useful to reduce risks of errors and frauds.***
- ***Points out main learning points of Machine Learning classification algorithms used for Checks by Monitoring in investigations on parcel.***
- ***Shares an idea of possible/desirable future evolution of entire Veneto Region classified near-real time using this algorithm.***

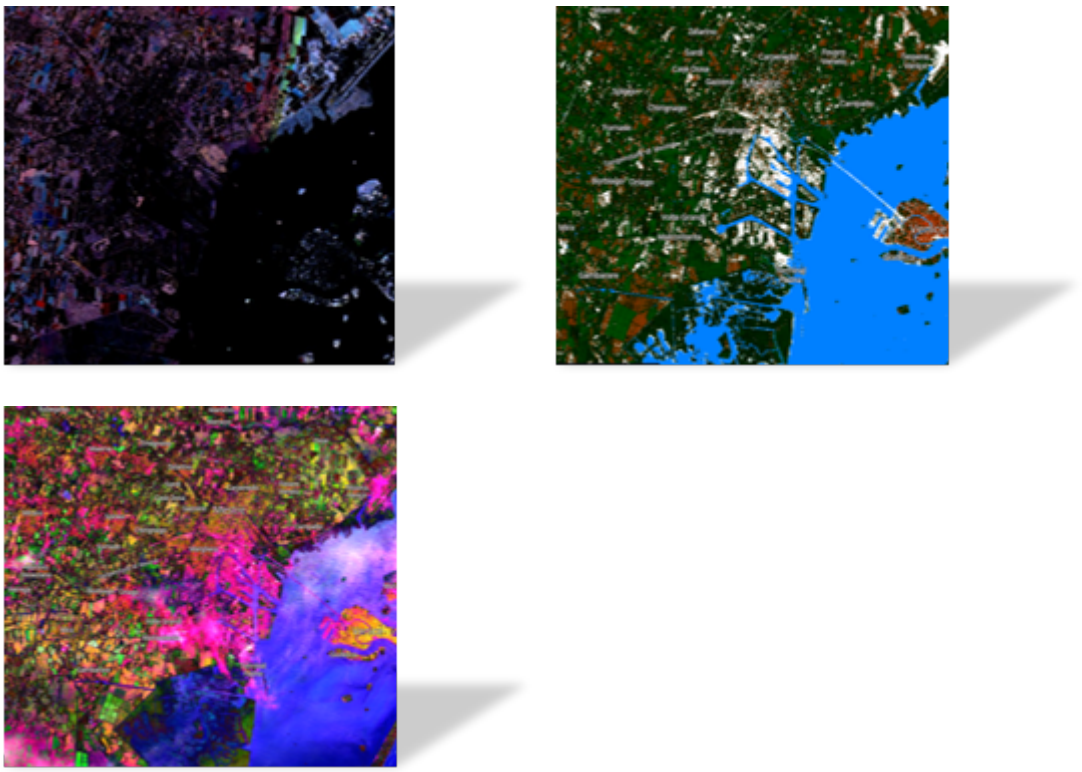
Background and starting point

Remote Sensing data gives Avepa the possibility to increase the knowledge of the regional territory, enriching in a relevant, pervasive, and semi-automatic way the amount of information associated with the declared crop geometries.

In this sense, this counterpart wants to present the work of last years of research and refinement of a model (more than one, to be honest) to automatically classify grasslands and pasture meadows using Sentinel data. We initially built it through a

traditional approach - i.e., Searching for the rules that characterize a phenomenon - and recently testing Machine Learning algorithms - i.e., Using data to train an algorithm to produce a reliable model. Avepa carries out On-The-Spot-Checks in the territory with the aim of collecting information necessary for the investigation on payment of contributions. It is not difficult to understand that these checks can only be a sample of the universe of crops: the human resources available are limited and the number of parcels to be checked is high. This exposes the Agency to risks, among others, errors, frauds and lack of information that can cause wrong payments. This is where Avepa's research work, on the use of remote sensing for checks by monitoring, comes in: can Sentinel scenes be used to reduce this risk? And can Avepa increase knowledge of the territory by using this data? If a person simply sees the photos constructed from the bands that the Sentinel collects, he might instinctively think that there is no added value in devoting resources to obtaining images that have such low spatial resolutions (10m/pixel at most), “free tools can give us better results”, he can think.





Different pictures of the Venetian lagoon, these real-time images are generated from the bands collected by the Sentinel through ad-hoc scripts: row by row, True Colour, Scene classification, Barren soil, Land Use Visualization, Urban classification, Growth Stage.

Well, we have experienced that **spatial resolution** is only one of the quality indicators, alongside **spectral resolution** (for the Sentinel 2, or S2, we have 13 bands) and **time resolution** (we can have with Sentinel 2, one scene at least every 5 days), which together with native **radiometric and geometric correction** make the Copernicus project useful in constructing effective indices to capture a lot of phenomena on the ground. By using the 13 bands and frequent scenes available for a given geometry, Avepa could calculate the time trend of several significant indices, compare it with the time course of other similar geometries and estimate land use, drastically reducing the need for on-site checks. All this material can be added to actual LPIS, increasing significantly the information connected to a parcel and giving to Avepa the ability to better monitor the land and its change.

Resources used.

We used a lot of Sentinel-hub publicly available resources:

- Dashboard/Configuration: https://apps.sentinel-hub.com/dashboard/

- Eo-browser: https://apps.sentinel-hub.com/eo-browser/
- Playground: https://apps.sentinel-hub.com/sentinel-playground/
- API: https://docs.sentinel-hub.com/api/latest/
- WMTS/WMS/WCS: https://docs.sentinel-hub.com/api/latest/api/ogc/
- Custom-Scripts: https://docs.sentinel-hub.com/api/latest/evalscript/

We consult Database indexes:

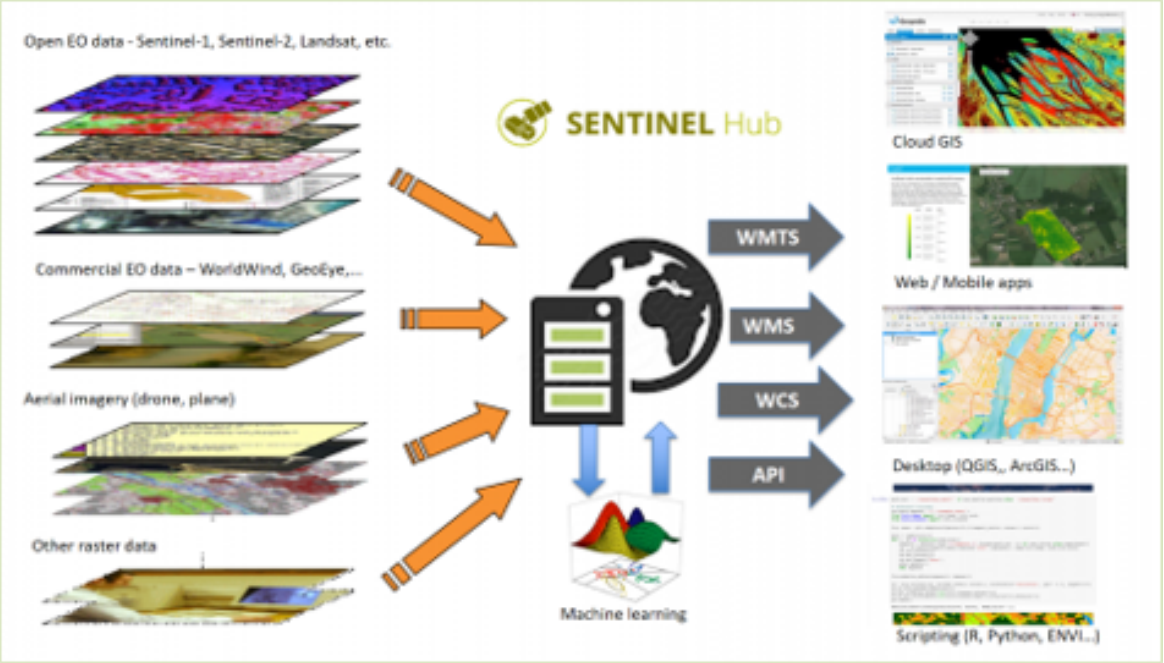
- https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/indexdb/
- https://www.indexdatabase.de/

Active community

- Forum: https://forum.sentinel-hub.com/

Other tools/languages

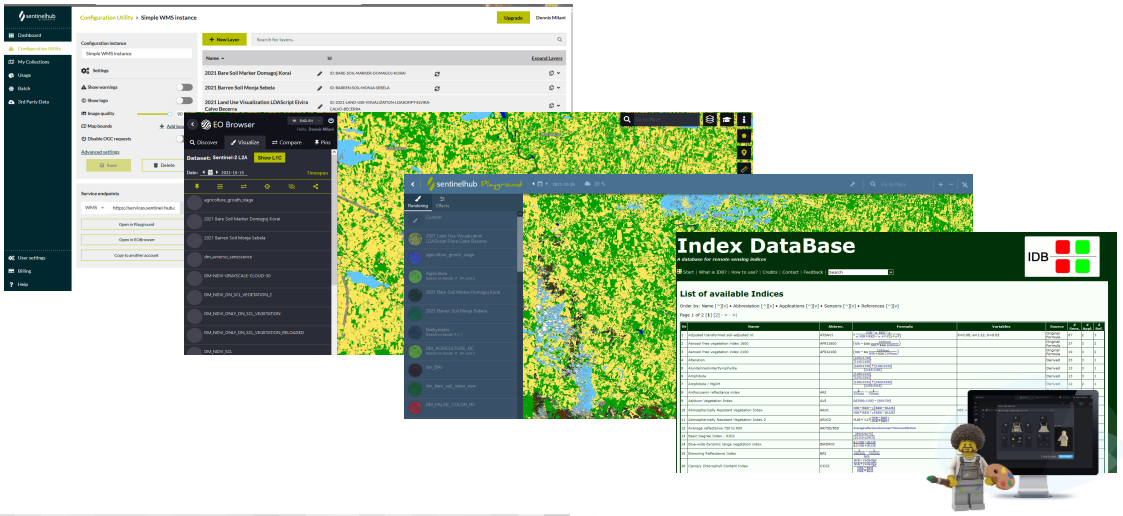
- First model was built in Python (with Sentinel and GDAL libs).
- Second model was accessed through API and a geo-package.
- Qgis.
- Avepa LPIS/Geoserver.



Our solution

Here is the *game the Remote Sensing Team played* in Avepa:

> "Give us the geometries with a particular declared crop practice to analyse and we will give you back the belonging of the geometry to the crop practice through simple *traffic light* system"



So, green-yellow–red flags to say ok–don't know–ko for accordance declaration and classification. The Copernicus project offers not only data from bands to study soil, but also provides many tools, training sessions and a very rich and active community, all elements that help and encourage the use of the entire ecosystem.

The idea.

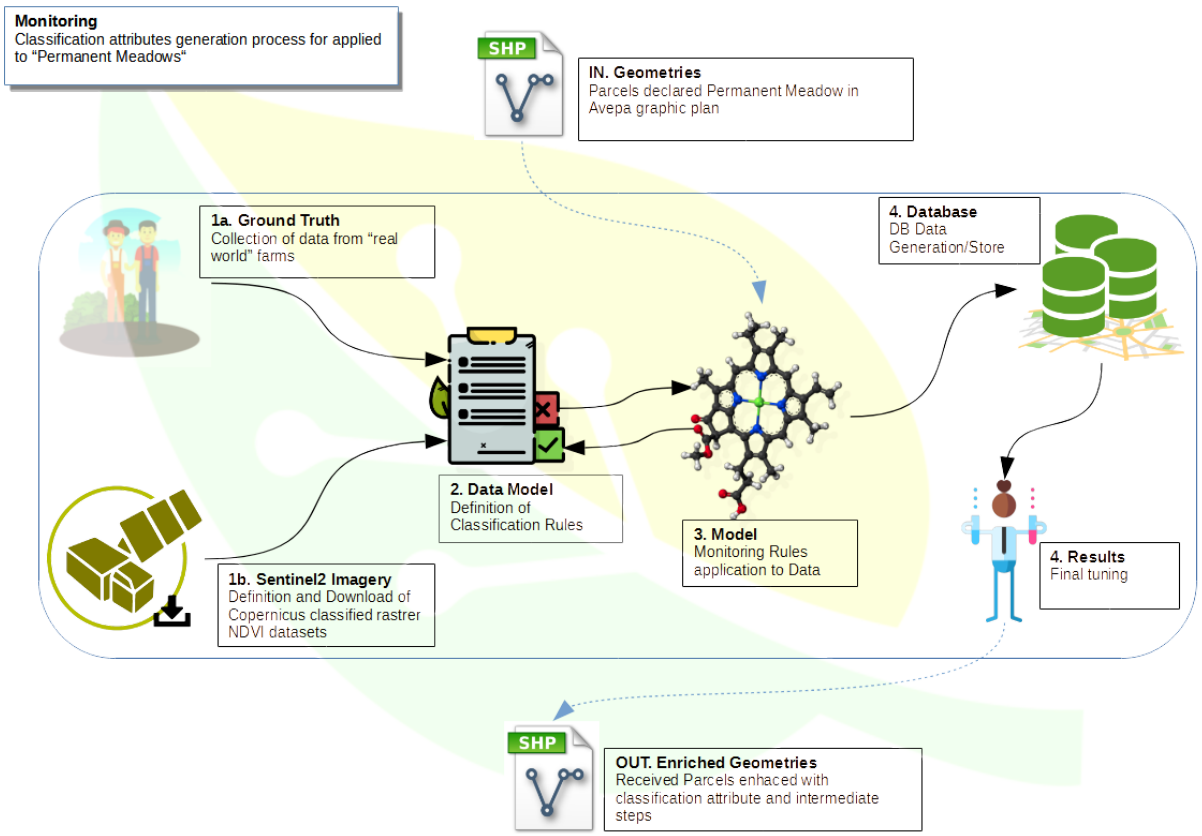
The starting idea was to apply S2 on this specific Avepa use-case: **Permanent Grassland specific farming practice**, so:

1. You (farmer) apply for maintaining permanent grasslands,
2. You are obliged to adopt a specific farming practice on your grasslands (no plough and mowing),
3. Avepa pays you if you keep the commitment to maintain grassland,
4. Avepa must verify this commitment (search for grass presence, there mustn't be ploughing, there could be 1 or 2 mowing events, …).

The *Remote Sensing Team* receive e specific engagement, to verify a classification. **We aim to build a specific model** to well capture grasslands and mowing events and use it to classify given geometries:

1. Input.
    a. A question to deepen.
    b. A shapefile from Avepa Geo Spatial Aid Application.
2. Elaboration.
    a. Collect enough ground truth.
    b. Download adequate S2 data.
    c. Eliminate *noise,* cleaning small and dirty geometries or *no data* pixels or cloud scenes.
    d. Analyse Indexes (Normalized Difference Vegetation Index, NDVI) trend.
3. Output: build a geodatabase and/or a shapefile with results.

To accomplish this task, we design the flow below.



To concretize this flow, we have followed two paths, let us call them **"the classical way"** and **"the modern way"**. Before start *walking*, however, let's talk about indexes, an essential concept to understand our study.

Indexes.

Sentinel 2 offers, for every pixel captured, 13 spectral bands every 5 days. These data can be used in many ways. After delving into the subject[1], the most interesting use for us was to construct the time course of the Normalized Difference Vegetation Index, or NDVI, to study vegetation: *"It normalizes green leaf scattering in Near Infra-red wavelengths with chlorophyll absorption in red wavelengths"*.

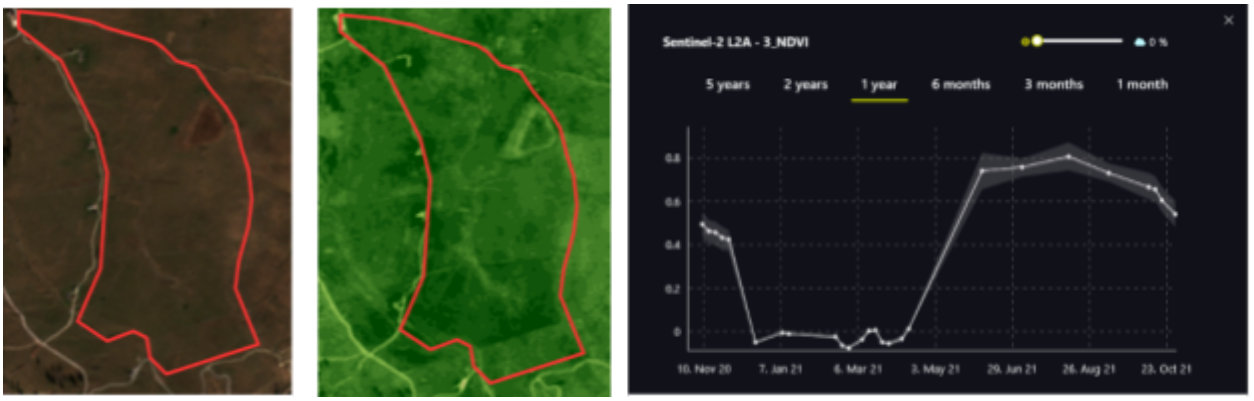
It is defined as:

$$NDVI = \frac{NIR-RED}{NIR+RED}$$

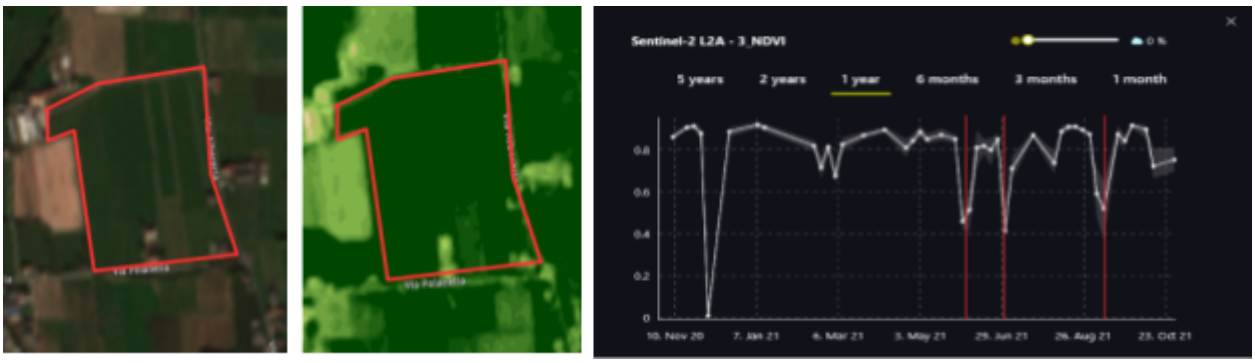
For Sentinel-2, the index looks like this:

$$NDVI = \frac{B8-B4}{B8+B4}$$

Let's see some examples.



[1] Ref. https://custom-scripts.sentinel-hub.com/; NDVI https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/ndvi/



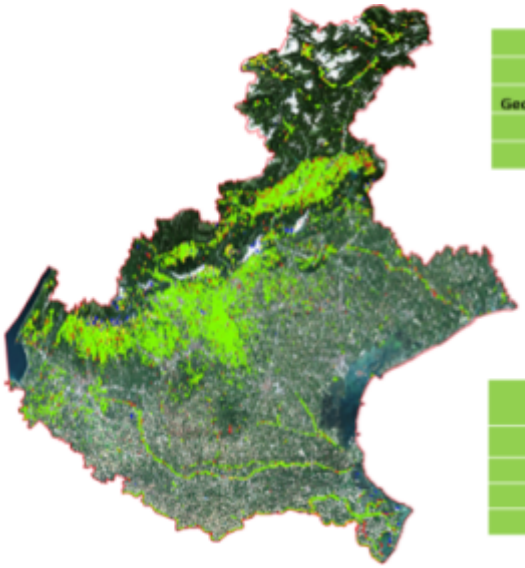
Examples.

First Geometry Row. NDVI is very low in spring season then grows. Could be a summer crop.

Second Geometry Row. NDVI is always high, could be e grassland with mowing events.

Scope.

Here are some measurements of the phenomena analysed to give an idea of the quantities we dealt with.



VENETO REGION DATA	mq	ha	kmq
Total Area	18.345.350.000,00	1.834.535,00	18.345,35
Geo Spatial Aid Application 2020	9.979.910.000,00	997.991,00	9.979,91
Agricultural 2020	8.034.260.000,00	803.426,00	8.034,26
Permanent Grassland	382.712.023,00	38.271,00	382,71

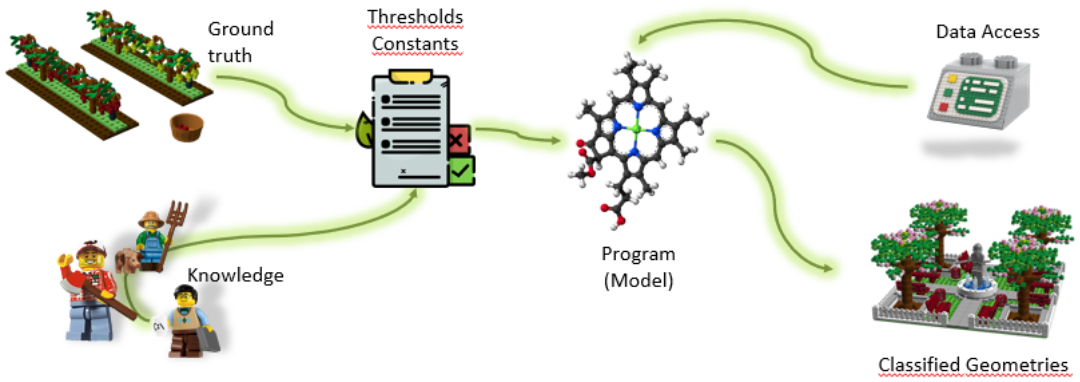
GEOSPATIAL AID DECLARED CROP	∑ ha	∑ ha (>1ha)	# Geoms	#Geoms (>1ha)
Perm Grassland 2020	48.564,00	30.144,77	76.464	13.304
Pastures 2020	46.585,00	41.358,56	25.013	6.074
Soya 2020	105.737,00	93.051,75	58.221	29.676
Beet 2020	8.775,00	8.616,96	2.372	1.785

Approaches.

Data Model: The Classical way.

In our first approach, named "**The Classical Way**", we find rules that regulate grasslands, using our and others knowledge and ground truth. We must know the grassland phenomenon we have to study: the phenological cycle, the incidence of moisture, the influence of altitude, the mowing and growth practices, …

Every information is translated in one or more rules that define *markers*, you can imagine it as a flag planted in the geometry, exposing some specific behaviour of the associated parcel. When all the markers are calculated, the process of classification is a decision tree, that assign a label based on values founded in markers.

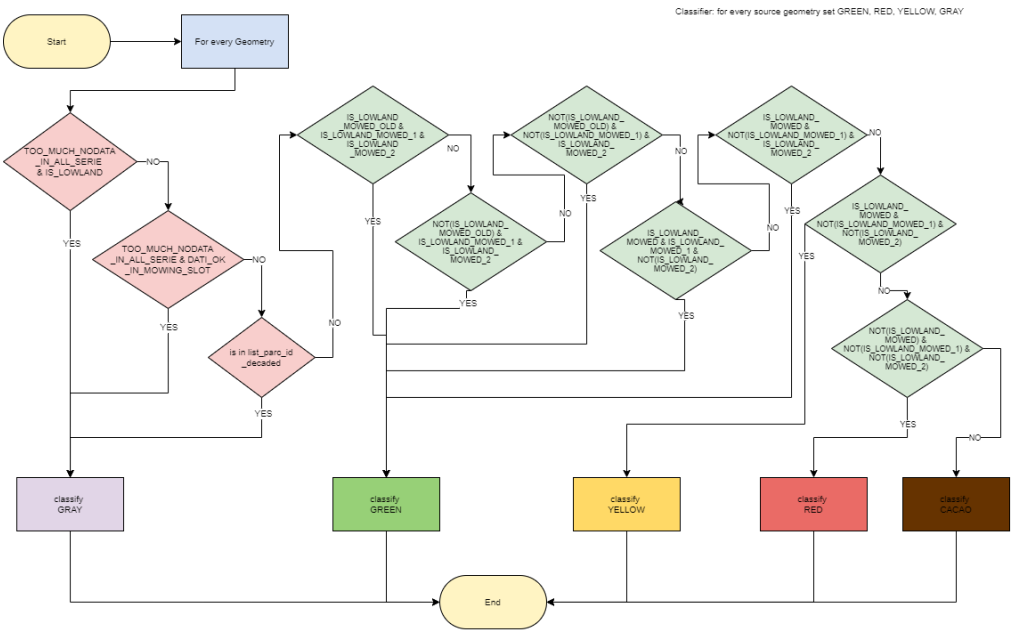


The Model: The Classical Way

Summarizing the process is:

1. Gather ground truth of the phenomenon to study.
2. Choose adequate index to capture grasslands (specific NDVI trend).
3. Prepare and download appropriate S2 images of the past agricultural year.
4. Set thresholds and constants.
5. Build a set of markers to describe grasslands inside the geometry.
6. Define a set of rules using markers to classify geometries.
7. Apply our rules to all input geometries.
8. Build a result database.
9. Verify the results.

10. Give back classified geometries and relative markers



Classification Flow

Results, “The Classical Way”.

Traffic Lights	Geometry Class	Automatic Classification	Supervised Reclassification
GREEN	Is mowed grassland	**26740** (75.4%)	**30007** (84.6%)
RED	Is not mowed grassland	**2610** (7.4%)	**2663** (7.5%)
YELLOW	Is not possible to define a class	**3108** (8.8%)	**1696** (4.8%)
GRAY	Insufficient data	**2704** (7.6%)	**796** (2.2%)
NULL	Poor geometry	**297** (0.8%)	**297** (0.8%)
TOTAL		**35459** (100%)	**35459** (100%)

Data Model: The Modern way.

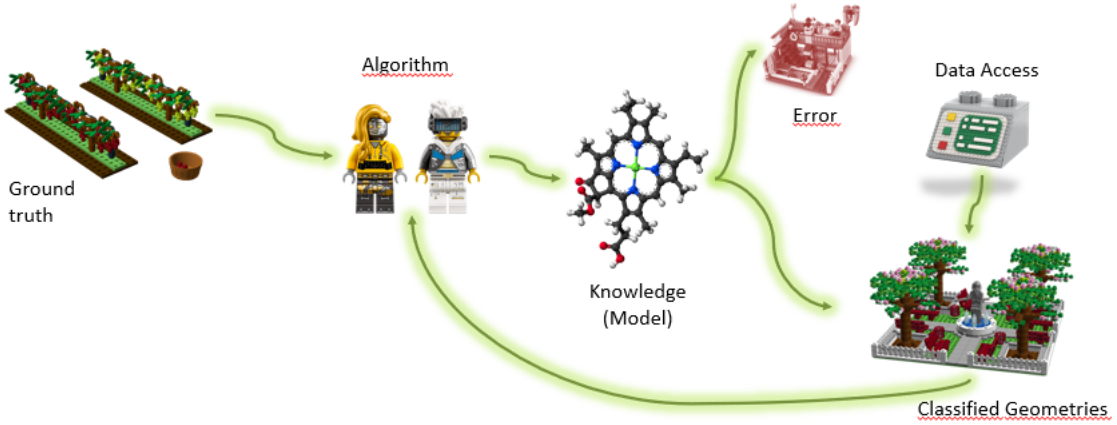
In our second approach – relying on an external company –, named “**The Modern Way**”, we use Machine Learning (ML) algorithms to study grasslands, soya, and beet on a smaller area (not the entire Veneto Region). Now a machine wants «the results needed» to define underlying implicit rules.

The process is:

1. Show to machine what a grassland is, it will learn;
2. Show it what a grassland is not, it will learn;
3. Tell it when it returned wrong output, it will learn;

We started from expected results, it built the desired model. As before, *markers* are defined, but now, using ML algorithms we:

a. Train an algorithm to define a model and then to improve results ("the more ground truth and valid result you give me, the more I will be precise");
b. Validate it ("tell me where I went wrong");
c. Improve the amount of analysed data.



The Model: The Modern Way

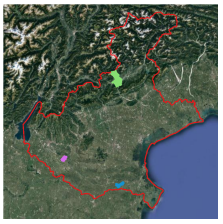
We used these markers:

1. *Mowing*: Machine detects mowing events;
2. *Mean*: Machine calculate mean NDVI value of all valid observations for geometries;
3. *Homogeneity*: Machine calculates probability of simultaneous crops inside same geometry.
4. *Bare soil*: Machine identifies a bare soil;
5. *Similarity and Euclidean distance*: Machine compares geometry to nearby geometries claimed or not.

We use, again, a decision tree to classify geometries using calculated markers.

Title	Results
Total Geoms in 3 test areas grassland, soya, beet:	**17496**
Total number of valid Geoms	**6057**
- Removed because of duplicated geometries	**26**
- Removed because topologically invalid	**1**
- Removed because they don't contain even one S2 pixel (10mx10m)	**11412**
Results provided for	**6057**
Grasslands declared	**2343**
- Valid	**2266 (96.6%)**
- Not Valid (found bare soil)	**77 (3.4%)**
- At least once mowed detected	**1982 (87,5%)**

Results "The Modern Way"



Lessons learnt

1. Avepa has experience in remote sensing through a drone (UAS, Unmanned Aircraft Systems). S2 allows us to have in the short time more multispectral

multitemporal images, even if resolution is lower than UAS, the advantages are interesting. Summarizing (UAS vs S2):

   1. Resources involved: 3 people vs 0 (automatic process in production environment);
   2. Time resolutions: once per mission vs 1 every 5 days;
   3. Space Resolution: 2 cm/px vs 10 m/pixel
   4. Spectral resolution: 4 bands vs 13 bands

2. A strong Ground Truth is essential to have robust results;
3. We will experiment extending crop types to analyze, now we have some experience only with 3 crops: Grassland, Soya, Beet;
4. It could be useful to extend markers, adding other information occurred in year (weather, drought, adverse events, …);
5. Publish results for expert judgement preliminary evaluation could be a useful result, now we give ready-to-use classification, without further explanations;
6. Approach “Near real time monitoring” can permit to build markers *as soon as possible*, not only at the end of the season;
7. Strong weather integration enforces good results: we saw that rains and droughts influences NDVI;
8. Results could be available and useful to farmers, giving them feedback and information;
9. Integration with other payment or monitoring contexts (i.e., Veneto Region) can improve data knowledge on land;
10. Further satellite constellations can lead to quality improvements, for example in no-data and clouds pixels;
11. We can obtain better results with contribution of knowledge from different fields (multidisciplinary);

Challenging issues

1. There are some topographical errors on source geometries;
2. Small geometries must be excluded (due to S2 resolution);
3. Clouds can interrupt time series (no-data problem);

4. NDVI variability connected to same crop at different altitudes;
5. A geometry can have different crops in the same season;
6. A geometry can have multiple crops at the same time;
7. Season water availability influences NDVI trend;
8. Need for processing capacity (disk, cpu);
9. Need for dedicated human resources;

Idea for the future

1. Can Avepa classify all the geometries contained in the LPIS, starting from a solid ground truth, near-real time, and independently from the farmer declaration? We can aim to have a database of remote sensing classified geometries, usable whenever needed, also in the application phase.

SUBSECTION 2: Farm's data management in AVEPA

This part:

- ***Improves the awareness of the PA's staff on the usefulness of having an organised set of updated informatic data representing the farm organization and people related to the farm, in order to prevent and tackles frauds and ease the payment procedures***
- ***Points out main learning points of which features must have a tool like Farm-register in collecting information and data from external sources insert here***
- ***Shares an idea of possible/desirable future state of approaching a preventive way to evaluate the eligibility of the requests for support applied by the farmers***

Background and starting point

When they want to apply for a request for support, farmers must supply to the Payment Agency several data concerning their activity (i.e., crop cultivated, animal breeder, machinery owned, buildings available, people) and if they want to apply for several requests, they must replicate these data each time. This situation can lead to mistakes in filling the requests and worst makes possible for unreliable applicants to declare voluntary wrong data in a part of the requests, in order to show a farm situation different from the reality. Beside this, there is often a repeated activity in checking the same data for different requests and a resulting heavier burden for inspectors.

Another problem observed in the declaring farm related data each time in several applications is that IT area takes more time to adjust the software when a new claim must be managed, because all the data checks must be fixed each time as new for the claim, and in the same way the admissibility checklists are longer.

Finally, another critical point which can be met when farmers apply several requests renewing some each time is to establish “who” is applying, e.g., some farms are managed as single ownership firms other as company and some farmers are involved in different farms, so when a farmer is filling an application, it can happen that it is not always clear “which” farm he is referring.

It can be considered that most of farm data don’t change frequently (not every year or every month) and there is data set that represents the “stable structure” of the farm, varying very slow through the years (fiscal identification, type of production, people working in the farm, machinery, buildings, land owned in property or with long time contracts).

If these data were managed in an IT tool (an “IT box”) separated from the applications, data checking can be concentrated in this tool and control can be more efficient, quick and engaging less human labour, especially when an automatic check is possible (e.g., verify with other public database, tax agency).

The several applications a farmer wants to apply could use this “stable and already controlled data set” to fill automatically a part of the request and so it is possible focusing the IT resources only on the management of the additional data required for that specific aid.

If the “stable data set” representing a farm is identified by a univocal identification code it will also be always clear “which” farm is involved in the application.

In AVEPA this IT tool is represented by the Farm Register, or better by the informatic data contained in the Farm Register.

Our solution

What is Farm Register?

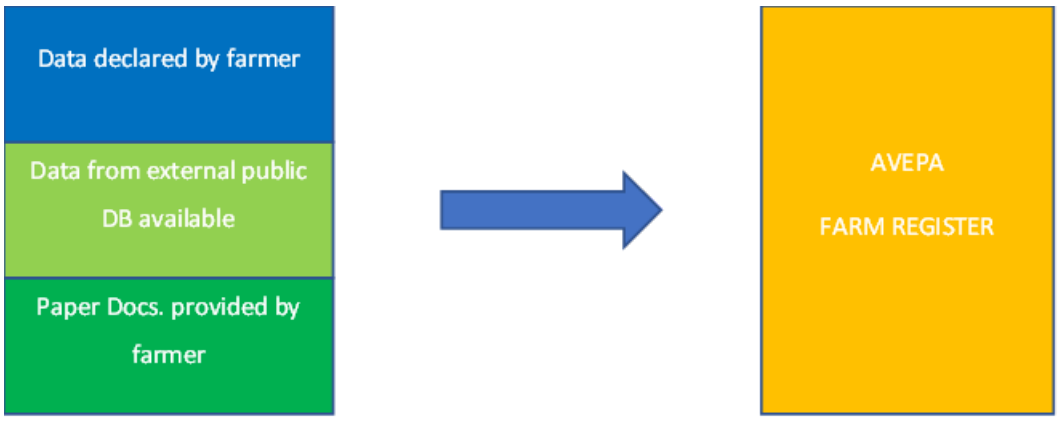
Farm Register in Italy first is a law requirement, since 2000.

Farm Register definition in the law was “the collection of information related to each farm”, no matter if private or public, “that has relationship (it means that wants to have relationship) with Public Administration”.

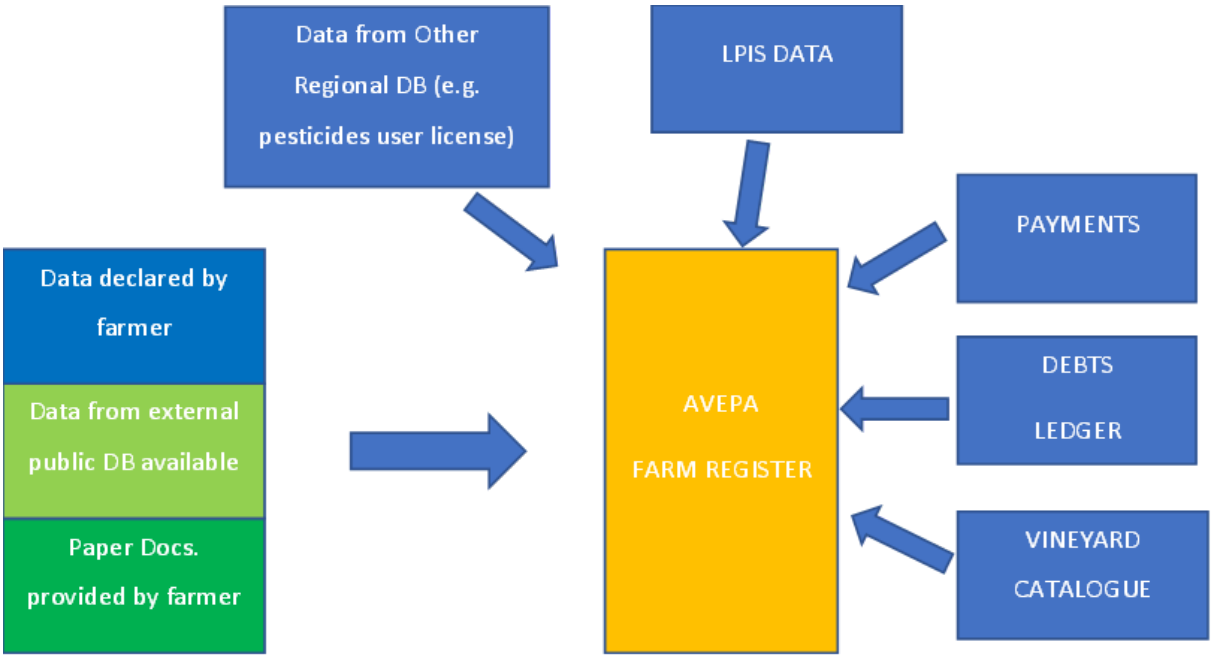
So, each farm that wants to receive a support, or that wants to require an authorization must have his own Farm Register formed. Simplifying, Farm Register represents the photograph of the farm. Each farm is identified by a unique code, that is called CUAA, translatable with “Farmer unique identification Code or Number”. CUAA depends on the type of farm organisation. For single owner farms it is the Fiscal Code, while for companies it is the Fiscal Code or the VAT number. In most of cases for companies has the same number as Fiscal Code and VAT number.

What Farm Register essentially is?

It is the sum of paper documents provided by the farmer, with various information he can declares and the data that may be available on several database managed by public authorities. So, we can imagine some little separate boxes that are joined together, and the result is the Farm Register.



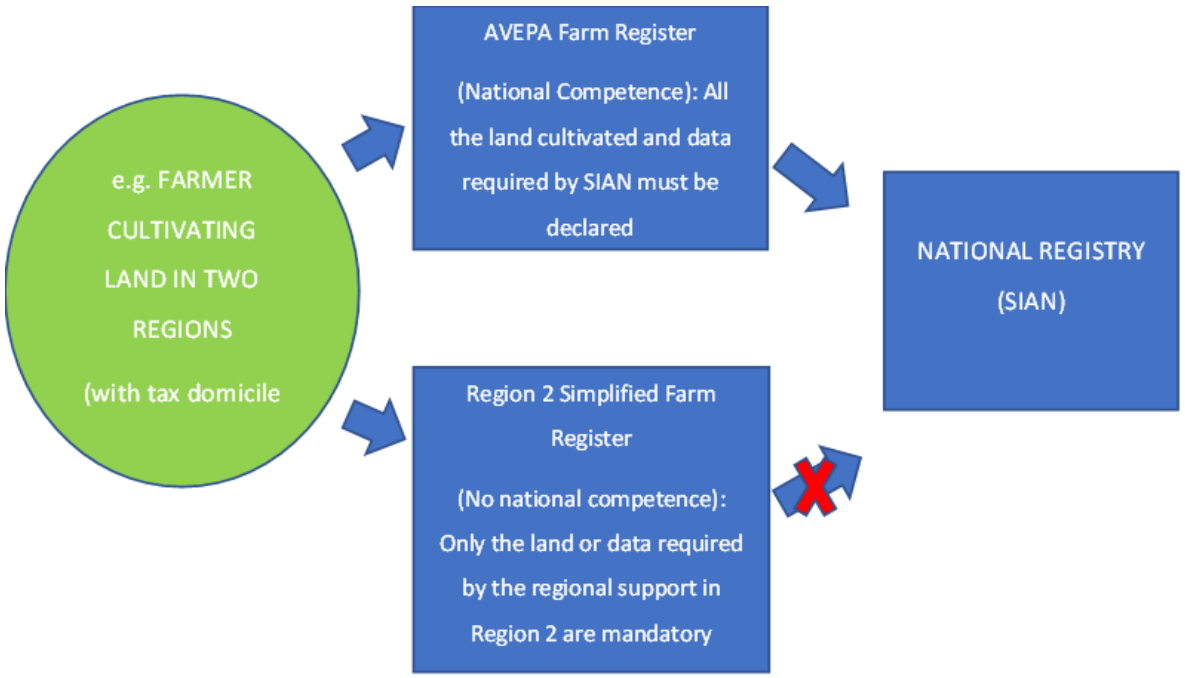
Farm Register is also a platform that is strictly linked to internal and external database that are available in the PA organisation.



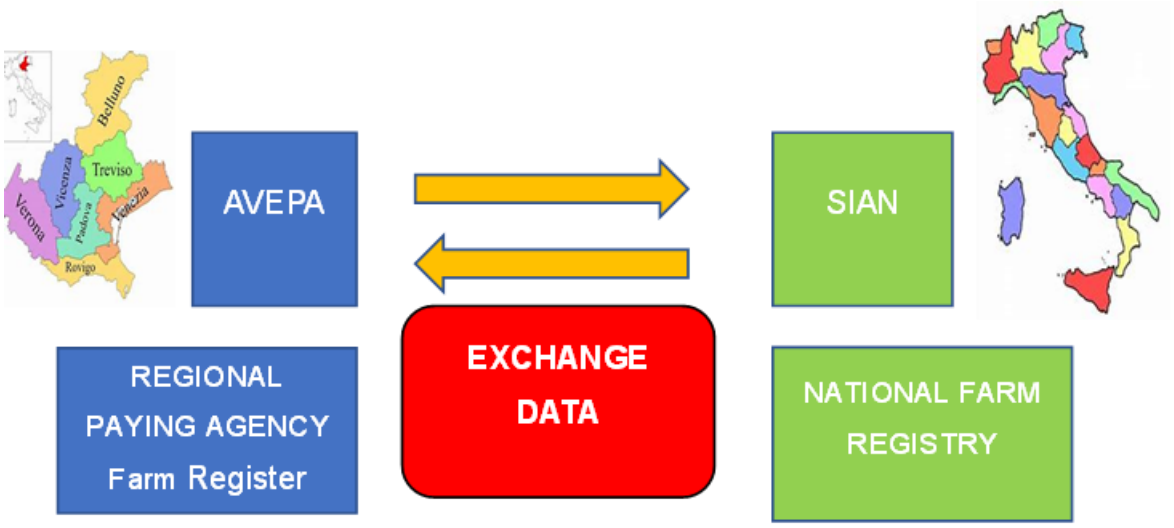
Since the national law require the farmer to have a register, Avepa Farm Registry is also part of the “national farm registry”, that is managed by a national informative system named SIAN. In Italy agriculture is an activity that is delegated, for its administrative organisation, to the Regions. And each Region has an own organisation, included the way to constitute farm register. Beside this, some requests are however managed only in the national level, through the SIAN platform. In most

of cases farmer operates only in one Region, he has relationship only with a Region and has one register, but it can happen that a farmer cultivating land in two regions needs to have two registers, one for each region, because he must apply, for instance, for different RDP regional measures or for vineyard activity. In any case, the national system will recognize only one of these registers, on the base of a competence rule.

Competence rule is the way used by SIAN to assign a farm to only one PA. The standard way is considering the Tax Agency Data, so the tax domicile (“where the farmer lives e pays taxes” or “where the company is located for taxes”) indicates the PA that is entrusted with that farm. A farmer who wants to change his pre-assigned PA (for instance the farmer lives in Veneto, but he has land only in Lombardy region, which is a bordering Region) he can do it, but previously he needs to require to the two involved AP a special authorization. In the example of a farmer cultivating land in two regions, and having tax domicile in our Region, he will form his register declaring all the data of the farm (with the land in both regions) and this register will be sent to the national system. In the other region, if he must apply for some specific support, he will form a simplified register with the information required by the support he wants to request. Anyway, we must say that in some cases, regions can have an informative system that use directly data from the national registry, so the data mandatory for their application must be present in the national register.



For the reasons we have just described, Avepa system needs to exchange data with the national system, that is SIAN system. The synchronization, for the farm register, follows two ways. When a register is validated, within 5 minutes the synchronization process starts. This process concerns the single register. Sian will check the incoming data and if everything is ok, in few hours the data will be viewable on that system. If there is something wrong (for instance the register is not under our competence) a detail of the problems encountered will be show in a special farm register section. The second type of synchronization is scheduled every night and it concerns a massive updating of some data for all the registers.



What kind of data we can find in the farm register?

The list below shows the most common data that ca be found in the farm register.

FR - DATA SET

a) personal data;

b) location of the farm;

c) legal representative and registered office of the farm;

d) references for registration for the Chamber of Commerce;

e) all the lands cultivated by the farm;

f) methods of land use (crops);

g) breeding;

h) machinery and equipment (if needed for applications);

i) buildings (if needed for applications);

l) other useful information (e.g. bank details).

Clearly, a register can have data for these categories, for instance a farmer can breed animals or not, and so the register will be complex. We must mention two important sections that are managed inside the Farm Register: LAND and CROP. LAND contains data about the documents used to prove the possession of the parcels and their cadastral data. CROP section receives data from the Graphical Crop Plan, that is a special application used by the farmers to declare the “yearly

crop plan" and by the AP to manage LPIS and OTSC activities. The two sections are strictly linked, because the removal of a parcel in the land section cancels the same parcel even in the crop section. Farm Register makes some automatic checks when data loading is considered completed, regarding with the applications the farmer wants to apply and the available data on external database. Main external databases used in the informatic controlling process are:

TAX AGENCY: supplies the CUAA, registered office, date of birth, place of birth, VAT number

CHAMBER OF COMMERCE NATIONAL SYSTEM (UNIONCAMERE-PARIX SYSTEM): for the farms which are registered in, supplies registration number, starting activity date, primary activity fiscal code (Ateco Code), certified e-mail address.

LAND REGISTRY (periodically update): supplies the basic information on the cadastral parcels (number, surface)

NATIONAL VETERINARY SERVICE (Health Ministry): supplies the farm breeding code and the kind of breeding animals. For some categories it is also possible to import the number of animals owned.

Considering that uploading data process is in theory enduring through all year, each Farm Register can be in one the following state conditions.

EDITING: the register is not complete, or it is under updating by the register's manager. This state can last until the updating is finished and the register is ready to launch the informatic checks.

VALID: when the register overcomes all the internal informatic checks (e.g., all mandatory data entry is completed and required documents are filled, external databases are verified) can become VALID. This is the unique state condition that allows the farmer to apply.

NOT VALID: when the informatic checks have been launched but not all overcome, the system marks the register with the state NOT VALID and shows an information

box with the irregularities found (e.g., lack of surface in the land use, Chamber of Commerce not verified, …).

When the irregularities are fixed, it is possible to launch a new control.

CLOSED/CEASED: when a register is no more used for the submission of applications (e.g., the farmer is dead, or he closed his activity) the register's manager must bring it to this state. The CEASED state is stronger than CLOSED state. The difference between the two conditions is on the reversibility. Only Avepa Central Unit can allow the re-opening of a ceased register (with justification).

Lessons learnt

The use of a tool like Farm Register allows some specific advantages in the organisation of AP activity.

DATA ARE DECLARED BY FARMER AND UPLOADED ONLY ONE TIME

Data are uploaded by FR's manager only first time and then, if they don't need to be updated or modified, they are available for all the other applications linked to FR. Some data need to be changed each year (e.g., crop plan) but many others represent the «permanent farm structure» and change less frequently.

DIFFERENT APPLICATIONS MAY USE THE SAME DATA

The use of FR as reference point for different applications the farmer wants to submit allows to prevent (or reduce the risk) to find different "versions" of the farm declared in different requests for support, with reference to the same time or period.

EASE APPLICATIONS CHECKS

Since some of the checks are previously verified by FR application, the related applications' checks may focus only on specific data they require.

EASE THE CHECKER OFFICER ACTIVITY

Checker officer can use FR as tool to value the global farm situation, considering further data there were not mandatory in the application submitted.

EASE ACCOUNTING UNIT ACTIVITY

Farmer bank code declared in FR is the only one used for PA payments, then it is easier for AU to manage changes after the application was submitted. At the same time, a more efficient control on historical data bank declared by the farmer is possible.

FLEXIBILITY AND IT DEVELOPMENT

FR Sections can be added or modified in order to manage new data/documents which become mandatory.

When necessary, new specific IT checks can be customized for each section and/or for FR Validation process (cross relationships between sections).

Beside this some difficulties and advantages must be considered. We can list some of them:

1) Data declared by the farmer that are not automatically verifiable in official public databases should be always validated some way, after they have been uploaded in the Farm Register. Otherwise, their use in different applications increase the number of times they can cause a damage on EU financial interests. This previous control is not always possible and so it remains part of the admissibility activity after the single request presentation.

2) The automatic checks need a good compatibility between IT tools involved. This situation can be a big problem when the databases have different structure and require a specific adjustment activity, bringing increasing costs in dedicated operators. In some cases, other public authorities are not interested in investing resources to evaluate a better integration with our systems or to supply data they manage.

3) Regionalised managed of Farm Register on one side permits a better customization of PA's administrative activities and software with the specific reality of Veneto Region, but on the other side requires the system is integrated with the national level and in some cases, there are problems in the exchange data activities.

What could be done for the future?

IMPROVE DATA EXCHANGE WITH THE OTHER PUBLIC AUTORITHIES

Improving the use of data available on other database and/or further developing of the ones reached (animal register, social security) is an important challenge in order to verify the truth of beneficiary declarations or to reduce/avoid the need to require them, by directly proposing him the official data.

Irregularities are often linked with such a condition, when declaration is false or contains wrong data the possibility to discover it is much more likely in such a way.

BETTER USE OF DATA ALREADY PRESENT IN THE REGISTER TO FIND NEW INDICATORS

The data's historicization recorded in FR can help to prevent irregularities profiling the more "critical" situations or farmers. Examples of this approach are e.g., beneficiaries linked by the same IBAN code, companies with the same persons as members and previous heavy irregularities or debts, farm registers with rapid or suspicious land movements, different farms that are owned by the same persons and managed as one but divided in order to avoid being under maximum size restrictions. This "data mining" activity in perspective is considered being the most interesting and effective in order to use the data analysis as way to prevent and tackle frauds. Indeed, this activity could lead to a better investigation and recognition of same situations.

We may list some of these potential indicators and issues.

Same Parcels used for the different investments (double founding).

Location and proper use of the investment (commitments maintenance).

Machinery ownership and machinery financed (double financing, commitments maintenance, use in non-agricultural activities).

Recording of previous "bad behaviours" of applicants (as element for risk in OTSC)

Cross-checking for ceased farms and commitments (maintenance).

Numbers of farms with same address of investment (link between people).

Cross-checking for specific data, considered to be likely linked to potential critical situations: age of farm (if very low), age of farmer (very high or very low), land owned (contract type and duration), amount of aid requested (if very closed to minimum or maximum level), applicants from abroad (more difficult controls), same bank code for different applicants (links between people), bank code from abroad.

SUBSECTION 3: Reasonableness of machineries' costs – AVEPA's experience

Background and starting point

Administrative check on applications for supporting purchase of machinery and equipment by farmers must include an assessment of the reasonableness of the costs, a task that should be carried out according to appropriate procedures. Until 2018 AVEPA, as recommended by the RD Programme Managing Authority established in the Veneto Region, carries out such an assessment by comparing three different quotes relating to the same item (i.e., the "three-offer-rule"). These offers should come from independent suppliers, be comparable and competitive with respect not merely to "official" pricelists but to actual market prices, i.e., those prices reflecting the real market situation for the specific item to be financed and with reference to the time at which the quotes were issued. Although AVEPA has been able over the years to develop adequate tools for tracking the assessments carried out on the reasonableness of costs - thanks to the constantly updated checklists and template minutes attached to each RD measure-specific manual - assessing compliance with the above-mentioned requirements of independence, comparability and relevance to the real market situation is still complicated. The assessment of these aspects is increasingly carried out by requiring the beneficiary to considerably integrate the documentation to be submitted. On the one hand, this practice constitutes a bureaucratic burden on the beneficiary, who could decide to forego access to EAFRD funding; on the other hand, this verification is also a slowing-down burden on AVEPA's own investigation process, given that the Agency must commit a large part of its staff to the assessment of the reasonableness of costs. Even so, in

some cases, not all the doubts raised can be fully cleared. Even if an in-depth investigation were to be led to the suspicion of "ad hoc" manipulations of the quotes aimed at specifically selecting one of the three offers submitted, if the winning offer is adequate and consistent with the real market situation, it would not be possible to prove that an irregular behaviour (or even a fraud) was being perpetrated to the detriment of the EU budget. Such a control system, besides being neither particularly efficient nor transparent, does not allow to satisfactorily detect irregularities or fraud to the EU budget, so it is desirable to base costs reasonableness assessments on other evaluation systems. As early as 2000, the Veneto Region separated the functions of RD programming and RD management and assigned the downstream programming phases to AVEPA, which manages payment authorisations and aid applications for the regional territory according to the regional guidelines and under the coordination of the Regional Government. Over the years, several weaknesses have been identified in relation to the specific subject of the three-offer-rule, the main ones being the following:

- Verification of the effective independence of the three offers: cases were found where allegedly different undertakings had the same address and/or website and/or telephone or fax number, or even the same owner/legal representative.

- Manipulated prices: presence of identical or very similar offers giving rise to suspicion of non-independent suppliers.

- Misapplication of derogations: in case of purchase of highly specialised goods or in case of completion of pre-existing supplies, it is possible to derogate from the duty of getting three quotes. Cases of misapplication of this derogation were found, whereby, although the goods were neither highly specialised nor did represent a completion of pre-existing supplies (with unique characteristics), the beneficiary had not submitted the three quotes required and the assessor had accepted the justification provided without carrying out verifications of any kind.

- False offers: in some cases, suppliers contacted directly by the inspectors denied having made offers addressed to the applicants checked

- Bankrupt or ceased suppliers later the offers: some doubts have rising about the validity of offers in terms of their effective representativeness in the real market situation

- Comparability of quotes: it was found that the three bids concerned goods which were not comparable from a technical point of view, due to the fact the endowment of optional features was so “inflated” that quotes could not be compared even if the base model was the same.

- Inconsistency with the real market situation: quotes issued at times too far from the one of submission of the application and therefore not useful to capture the real market price.

The Paying Agency considered difficult to dismiss the three-offer-rule as a tool for assessing the reasonableness of costs, but the problem was raised to the Managing Authority anyway, explaining that it was advisable to develop an evaluation tool that could support or, in the best scenario, even replace that method. The goal was to overcome the well-known implementation issues, thus ensuring a more uniform evaluation by all the assessors involved. Other goals to be simultaneously pursued concerned the downsizing and streamlining of the controls to be carried out, and, finally, a substantial reduction of the bureaucratic burden on beneficiaries, so as not to discourage them from applying for aid, thus ensuring a numerically adequate participation in the RD calls.

Our solution

As a result of the fruitful collaboration established over the years with AVEPA, the Veneto Region has tackled the problem and suggested the development of a machine and equipment price list, an idea inspired by the positive experience gained from the agro-forestry price list used for the itemised estimates required within the structural intervention procedures. The working hypothesis was that the database would set the maximum cost of a given item on which both the eligible amount and the amount to be financed could be based. Therefore, the database had to possess certain characteristics in order to be useful for the purpose, namely:

Accuracy: the prices of the price-book had to reflect market values and, therefore, derive from a survey that was not limited to a mere collection of pricelists or data retrieved from suppliers.

Up to date: the price list had to be regularly updated, in order to reflect market trends which, in recent years, have shown a considerable volatility due to the underlying economic crisis.

Comparability: the system had to arrange prices collected from suppliers into homogeneous groups, classified by types of machinery and equipment, and compare them based on precise quantifiable technical characteristics that are specific to the intended use (e.g., engine power for the tractor).

The Agro-food Directorate of the Veneto Region deemed to outsource the development of the database and turned to the Electronic Market of the Public Administration. The only registered supplier suitable for the development of a reference cost calculator for agricultural machinery was the company Edizioni L'Informatore Agrario s.r.l., a publishing house operating in the agricultural sector with over 40 years of activity. Edizioni L'Informatore Agrario, besides having a deep knowledge of the agricultural and forestry sector in Veneto, also manages a database collecting technical characteristics of new and used agricultural machines on the Italian market, including the corresponding list prices, which are published in its own thematic magazines as well as on its own website (http://www.macchineagricoledomani.it/). This led to the adoption of Director of the Agri-food Directorate Decree No 111 of 7 November 2017, which directly awarded to Edizioni L'Informatore Agrario the contract for developing a price-book of maximum unit costs for agricultural and forestry machinery and equipment. In Veneto Region there are several kinds of environments, many and various soils and grounds with different slopes, a large variety of cultivated crops, of livestock breeding, and different farm management systems (intensive, organic, integrated, etc). This heterogeneous context implies that the Agricultural machines, in their variety and heterogeneity, can feature multiple parameters and functional characteristics influencing not only their performances but also affect their price. The aim was identifying, for each category of traction and operating machines, those

homogeneous and univocal parameters most affecting their purchasing price, as well as in quantifying their impact. The goal was to elaborate an algorithm calculating the reference cost for a given piece of machinery.

The categories of agricultural machinery considered in the survey are:

tractors (conventional, tracked, specialised, isodiametric and telehandler);

large harvesters;

machinery for soil tillage, crop protection (i.e. atomizers), mineral and organic fertilization (fertilizer spreaders), haymaking;

agricultural trailers;

mixer wagons;

machinery for viticulture, olive growing and forestation.

The activities carried out by Edizioni L'Informatore Agrario s.r.l. to achieve the goal were carried out as follows:

updated pricelists of the main manufacturers and retailers operating in Italy and in the Veneto Region were retrieved. The database used by Edizioni L'Informatore Agrario includes technical characteristics and list-prices updated to 2017 for tractors and self-propelled vehicles and to 2016 for other agricultural machines marketed in Italy. Before delivery, the database underwent a further check and a filtering process. Data were also filtered and processed before being used.

A subdivision of the categories of machines into homogeneous sub-categories was carried out, and the mechanical characteristics that specifically and unambiguously differentiate the categories and sub-categories were subsequently identified; the technical parameters most affecting the price were then investigated by using traditional descriptive statistics tools (indexes of central tendency and data variability, box plots, histograms, etc.).

A statistical analysis was carried out to identify the most significant correlations between the (numerical and non-numerical) parameters considered and the selling

price for each sub-category of machinery (dependent variable) using one of the following methods:

simple or non-linear linear, polynomial, multiple regressions.

multivariate methods.

The appropriateness of the regression models designed, and the statistical significance of the estimated parameters were then evaluated. When relevant, checks on statistical goodness of fit and significance levels include:

estimation of correlation indexes r and/or determination index R^2, testing correlation hypotheses by means of Student or Pearson tests

analysis of residuals

Estimates the standard error and confidence interval.

Verification of hypotheses by means of Fisher tests on the slope of the regression lines.

Algorithms were then defined, and user-friendly models were developed for estimating the reference price for each sub-category of machine.

The statistical analysis was carried out by Edizioni L'Informatore Agrario using Excel built-in functions and Adalta's statistical package Statgraphics Centurion XVI.

With reference to Article 62, paragraph 2 of EU Regulation no. 1305/2013, which provides, even in cases of adoption of a price-book, that the Managing Authority shall ensure the accuracy and adequacy of data through a fair, equitable and verifiable calculation, the Department of Land and Agro-forestry Systems of the University of Padua (TESAF) has been identified as a functionally independent institution possessing the necessary expertise to validate the accuracy and adequacy of the calculations. On September 25, 2017, TESAF issued a statement certifying the accuracy and adequacy of the calculation methodology implemented for the maximum cost reference price-book.

Once the algorithm was developed, its performance was verified - for each type of machine considered - based on 1.034 quotes (which included discounts applied by dealers on list prices) provided by AVEPA for the machinery and equipment

positively appraised in 2016-2017 for the applications submitted under the 2014-2020 RDP calls for proposals. These estimates have been compared with the simulated values to identify the goodness of the model and the average discounts to be applied.

In particular, the analysis of the estimates showed that the model overestimated the actual values; in order to achieve a better accuracy, a correction coefficient was calculated and applied to the model so that most of the data would not exceed a ±20% deviation from the actual figures. The coefficient is 0.775, which represents a 22.5% reduction in the price estimated by the model. This value can be considered reasonable based on two seemingly opposite factors, both of which may explain this variability. The first one concerns the list price, on which the simulations are based, which is never the final purchase price, as it can be modified because of discounts applied by the seller, payment methods, existence of tax benefits, second-hand goods return, etc. The second one regards the estimates provided, which often include the provision of add-ons or accessories that are difficult to estimate with the available data. At the end of the project, by decree of the director of the Directorate EAFRD Managing Authority, Parks and Forests n. 111 of 18/12/2017 the price-book of maximum reference costs for agricultural and forestry machinery and equipment for the Veneto Region was approved to replace the “three-offer-rule” for the purpose of submitting applications for support under the Rural Development Programme of Veneto 2014-2020. The pricelist is also made available to operators as reference price calculation app; it is downloadable form the Internet and works with both Windows and Apple operating systems. Edizioni L'Informatore Agrario s.r.l. has agreed with the Region to implement three updates by 2021 but is available right from the introduction of the price list to integrate or update it in a timely manner whenever the relevant departments of the Region deem it appropriate. The machinery and equipment price list were updated twice in August 2018 and in December 2021 including new items and revising the prices of those already included based on the optional equipment concerned. With reference to the documents to be attached to the submission of the application for support, the introduction the new system allowed a simplification of the calls relating to those

types of intervention of the RD involving the purchase of machinery included in the price-book among the eligible expenses, since applicants shall attach only the report produced by the app. Since private applicants are no longer required to get three quotes from three independent and competing suppliers, a significant reduction of the bureaucratic burden has been achieved. Moreover, AVEPA has been able to simplify its own procedures, not having to check and fill-in the checklist relating to the verification of the three quotes on a sizable share of applications regarding machinery and equipment. As far as the analysis of the add-ons to the basic models is concerned, the decision was made to take into consideration - as a starting point - only the ones deemed necessary to put the equipment in use. For each basic model, the pricelists the most frequently purchased add-ons in the reference market have been included, thus allowing the applicants to choose her preferred set-up in the subsequent application for support. Moreover, add-ons that are deemed not in line with the RDP grant are not included in the price-book.

Lessons learnt

The of ongoing evolutionary process pursued by AVEPA and the Veneto Region in terms both of simplifying administrative procedures and adopting innovative IT tools has triggered the development of this new procedure to assess the reasonableness of costs. The above-described path should be followed by any PA but, irrespective of the subject entrusted to the development of such a “reference cost” database, the availability of a significant amount of data it is needed. In the case covered in this paper, the project started from a price collector organized by category of machinery and equipment and the collection of the quotes that AVEPA gathered during the RDP 2014-2020.

As first, in order to serve the highest possible number of applications for support, it is necessary to focus on those pieces of machinery and equipment that are mostly widespread in the pertinent territory and then, at a later moment, also goods characterised by a more restricted market, but for which a sizeable demand exist, can be included. At the same time, the add-ons to be combined with the basic versions should constitute a rather restricted set and include only the most

purchased ones in the reference market and the ones deemed necessary to put the equipment in use. It should be noted that the database should be organised by categories and corresponding sub-categories to ensure comparability.

The regular update of the database is crucially important, and it should not be carried out following a rigid calendar but, as far as possible, on the basis the evolution of relevant market conditions, so that real values can be effectively monitored. Each revision should be followed by informative notes alerting users and specifying which part of the database has been modified and/or integrated.

The pricelist should provide the maximum eligible cost for the purchase of a particular asset. The introduction of maximum eligible costs should allow the simplification the application submission. Costs should not merely derive from producers' official price lists but should be calibrated on the discounts usually applied in the territory, in order to ensure the reasonableness of costs principle.

Anyway, the "three-offer-rule" - where the offers are reported also in a technical paper explaining the final choice - cannot be totally replaced as not all categories of agricultural and forestry machinery and equipment fall into the "price-book of reference unit maximum costs" approved by the Veneto Region's RD programme Managing Authority. In fact, there exist in the market eligible items which cannot be included in the price-book because of peculiarities such as a high technology content or rare and not comparable characteristics of the very place in which these goods are put in use (Outliers).

The passage, even though still partial, from the evaluation of the reasonableness of costs based on the three quotes to the one based on the reference maximum unit costs list of agricultural and forestry machinery and equipment, has been meeting since the beginning a significant appreciation by AVEPA's assessors, beneficiaries and consultants.

Although the system has been only recently adopted, all actors involved agree that the price list allows a better identification of real market prices while providing a smart solution to the issue of the add-ons that often made the three-offer hard to comparable.

Current appraisal procedures have shown a significant streamlining of AVEPA's investigation process, thus facilitating managers in charge of the administrative procedure during making decision phase on the admission to funding. Moreover, a more uniform applications assessment has been achieved, thus ensuring that the same funding is given to beneficiaries requiring the same items.

However, potential areas for improvement are already evident:

1. The model is reliable for ordinary machines while it seems to underestimate high tech machines (so in few cases we came back to the three offers)

2. Very complex to update: this method has worked well until 2020 but now it seems to be not able to intercept the current market dynamics both in case of price decreases and increases (in December 2021 came the first update since 2018)

3. price updates publications should be synchronized with calls publications (avoid adopting price updates during the submission of applications phase) otherwise it could be the risk of different prices for the same good among applications.

4. Technical assistance provided by the suppliers should be considered in the eligibility cost of machinery

5. The downloadable «local» copy of the software programme once downloaded should be updated; anyway, it has often happened that applicants have not the latest version available at the time of submission. Therefore, it could be interesting switching to a web-based application and/or a mobile app to better manage the updates.

SUBSECTION 4: Innovative IT approach for OTSC - the experience of ERDF Regional Operative Programme 2014-2020

Background and starting point

The aim of this intervention was to explain the methodology used and the lesson learnt in doing OTSC during the pandemic crises.

Starting from the 2017 AVEPA, acting as intermediate body, deals also with the management of the calls financed with the measures defined by the ERDF ROP. In this field, AVEPA performs:

1. Checks on the eligibility phase which means verify subjective requirements, undertaking difficulty, compliance with state aids, SME's dimensions and organize the technical evaluation projects with experts;

2. Checks on the payment's requests related to expenses eligibility and correspondence between documents provided to the administration and the original ones, social security and contributory regularity with regularity certification (DURC), anti-mafia and Deggendorf clause documents, realization of the project inside given periods, realization of adequate advertisement and information as requested by the Regulation 1303/2013.

In order to complete all these checks required by the procedures, we have three means, conducted to the beneficiary:

3. In situ checks (carried out for all the payments requests)

4. On the spot checks (realized only for sampled beneficiaries)

5. Ex post checks (also in this case realized only for the sampled beneficiaries alter three/five years the ending of the project).

Aim of these visits is verify the real realization of the project, acquiring also photos of the intervention/prototypes/machineries. All checks are included in a report and a minute is edited.

Our solution

This procedure faced difficulties during the pandemic crisis depending on the impossibility to go to beneficiary, so in order to overcome this issue, in 2020 the Territorial Cohesion Authority allowed to use telematic modality in doing checks instead of in presence checks.

Consequently, AVEPA introduced a new procedure based on video checks which was implemented with the following steps:

1. To organize the meeting with the beneficiary, which means communicate to the beneficiary all the main aspect related to video call (time, modality, privacy regulation), prepare the meeting, filling out the necessary documents as checklist and minutes with the available information, identify the referent for the SME;
2. To find some geographical references (using street view);
3. To open OBS instrument to register (a free download program) the entire check in one or more video;
4. To start checks on: beneficiary identification, application stamp tax, n. of check extraction, description of some information related to the check, attendance list;
5. Following the descriptive relations already transmitted to AVEPA, to ask to examine invoices, plans and all other documents/objects/prototypes photos in order to demonstrate the realization of the project. In case of OTSC we acquire, before the meeting, all the accounting registration (VAT, Separate accountability...) and then we transmit the minutes to sign manually or digitally.
6. To check invoices and related expenditures.
7. To check advertisement obligations.

Lessons learnt

Using this new method, we faced positive and critical aspects. Such as:

- Reduction of administration costs, with particular reference to the elimination of travel costs.

- Greater efficiency in terms of time. If before to use video calls, to implement a check was necessary at least one day, now is possible conduct more checks in one day.

- Prevention of corruptive risks. Checks are entirely registered with video.

Critical aspects are:

- Human aspects such as difficulties in using IT systems for the beneficiaries or connection problems.

- Borderline situations hard to monitoring with video calls.

For the future we expect to continue to use both methodologies, using video calls as ordinary method with the possibility to conduct visits in presence for the problematic situations or where there is not the possibility to verify the realization of the project using IT systems.

PART II Experiences of using Arachne risk scoring tool

This part:

- ***Maps the challenges of using Arachne in AVEPA***
- ***Shares Arachne tested solutions (tips) for specific EARDF measures***
- ***Gives considerations and ideas of using Arachne as the tool in Agricultural funds***

The problem

AVEPA intended to test a new computerised procedure for prior checking of situations of conflict of interest and potential fraud between funded applications. ARACHNE is the tool that the European Commission has developed for ERDF Managing Authorities to help them putting in place effective and proportionate anti-fraud measures. This tool is intended to be used also in the frame of agricultural funds especially during the next programming period (2023-2027) therefore AVEPA started to test ARACHNE during 2021 and here are briefly reported the main steps followed and the general results of this first trial.

Our starting point

Testing activities were carried out by a team of employees belonging to different areas of the Agency and the testing project was developed in close contact with the services of the European Commission in charge of the development of ARACHNE.

The following steps were followed:

1. Identification of the sample

In order to verify the functioning of ARACHNE, the consistency and the nature of the return data, it was chosen to send a set of information present in the AVEPA databases related to the intervention type M0411AVI - "Investments to improve the overall performance and sustainability of the farm". This type of intervention was chosen for the following reasons:

(a) The high number of applications submitted provided a lot of data with great heterogeneity of the data;

b) The applications were in a "CLOSED" status, therefore it was possible to select and send information concerning all the project implementation phases.

2. Identification of valuable risk indicators and outputs

Based on the document "ARACHNE - Risk calculations", the preparatory activity for the sending of data was as follows:

a) Identification of the individual risk indicators, for which they could provide information and data, with delimitation of the type of intervention analysed and nature of the data available;

b) Identification of the related inputs that could be used (detailed data that could be extrapolated from individual projects). It emerged that the availability of data mainly concerns monetary data, on project expenses, while organisational, economic, financial and statutory data of the applicants/companies are scarce or incomplete).

c) Processing of the data held by the agency, coding of the same and compilation of the .xlm files to be sent to ARACHNE.

d) Sending data.

Our solution

A sample analysis carried out on the return data from ARACHNE revealed the following observations that can be grouped into two macro-categories:

GENERAL observations:

- ARACHNE assesses every single application as a 'project'; this affects the meaning to be given to some return data because changing the size and scope of a 'project' also changes the interpretation to be given to a data. Consider, for example, the individual risk indicator "overall eligibility score" and one of its risk factors "high percentage of costs awarded at the end of the project". If the factor were analysed on the totality of applications submitted, it is likely that the incidence of "risk" would

be proportionally reduced. On the other hand, if the risk factor were analysed at the level of the individual project application, the beneficiary who does not ask for interim payments but goes directly to the final payment would obtain a "high risk" indication. What for ARACHNE represents a serious alert, for the agency is not a problem due to the fact that each beneficiary can be allowed to report the whole initiative on the balance.

- The phase in which the project-application is sending the data to ARACHNE, the great variability of the types of interventions financed by the project-applications and the consequent heterogeneity of the types of expenditure, could lead to a different reading of the individual risk factors that ARACHNE highlights. The cross-referencing of the data provided by the agency with data of a different nature and origin present in other European databases could also alter the meaning of the results. The standardisation of risk factors at European level makes the system rigid in carrying out the assessment despite the interpretability of the results in relation to the context. Perhaps we should have a clearer idea of the 'basic context' used as a model and on which ARACHNE was built so that we can compare it with our own.

- Each piece of data sent to ARACHNE represents an input for multiple risk factors; the combination of these inputs leads to the processing of the risk factor regardless of whether all the inputs it requires are present.

Given the peculiarities of the agricultural, forestry, training, partnership, etc. realities where there are companies exempted VAT regimes, without the obligation to deposit the balance sheet in the Chamber of Commerce and exempted from depositing the information required by the algorithm, the few inputs sent create a high number of risk reports, due mainly to the very lack of inputs in the reference banks.

DETAILED observations:

- ARACHNE returns two types of assessment:

a) A global risk assessment on the application - project (OVERALL application score)

b) A risk assessment on the individual factors determining the global risk assessment (warnings/alerts within the 7 risk categories)

- The fact that the risk assessment on the single factors determining the global risk assessment is referred to the single application implies that, in order to ascertain the actual existence of a reported risk, it is necessary, application by application, to refer the corresponding factor (risk report) to the data codified in the sending table and therefore to the corresponding data in the application. This procedure becomes, in fact, a sort of ex-post investigation. Consider, for example, the sample under analysis made up of more than 980 project applications, the majority of which had a detailed risk factor highlighted (yellow to red dot).

- If each "high" risk factor contained in a single project-application had to be followed up or justified, this would become extremely burdensome in terms of resources and time and would require the constant supervision of a working group dedicated exclusively to this task.

Lessons learnt

From the analysis of the critical points, it emerges that the findings of ARACHNE are based on an extremely large variability of factors. It is therefore mainly the context to which ARACHNE is applied that makes it a more useful or less useful tool. The future introduction of ARACHNE in the procedures adopted by the Agency could be very complex. Selecting the most appropriate stage of the project applications, extracting the useful data, processing them, sending them, interpreting the results and identifying what corresponds to a risk and then decoding it in a specific control procedure, as ARACHNE is now, is a process that implies a substantial workload and the involvement of different sectors of the agency. It is certainly desirable to further explore the ARACHNE tool from several points of view:

- How it was developed (basic analysis and context);

- Specific training on its use for all the actors involved;

- Impact of the findings on the administrative process;

- Positive and negative implications of its use;

It is also appropriate to clarify with the services of the Commission how flexible and/or modifiable the tool can be in order to adapt it to the needs of a more restricted area of use (e.g., RDP measures of the Veneto Region), or if there is the possibility to downgrade ad hoc some risk factors (which we will demonstrate do not represent indicators of problems). It would also be desirable, before implementing it in the Agency's procedures, to start an additional testing of the tool in order to adapt it to the agricultural-forestry reality. We believe that this sector can broadly differ from that of the structural funds for which ARACHNE was conceived.

Conclusions

During the study visit held in Padua on the 17th and 18th of November a different variety of issues were touched upon. In particular, the following topics were deepened during specific brainstorming sessions:

1. Data-driven land monitoring.
2. Farms data analysis for the prevention of potential frauds.
3. IT tools for the prevention of fraud (Arachne).

In this report the experiences of the Agency in dealing with these topics, have been outlined and brainstorming outcomes particularly related to ARACHNE are listed as follows:

- As a general remark Arachne can be a useful tool that generates a huge amount of information and provides many search options but there is still room for improvement.
- Additional guidance from EC on how to use Arachne in a proper way would be needed. Clarifications on financial implications (financial corrections) in cases of limited / partial / non-check of indexes by PA would be necessary.
- As a general recommendation this tool should ease controls and not impose an excessive additional burden to PA - so there is a clear need to strike a balance between additional administrative burden imposed to PA and effectiveness of this tool.
- More flexibility on the investigation of indicators' results should be left to PA.
- A revision of aggregation of scores fitting into a unique aggregated index should be possibly taken into consideration.