grapevine

IBERGRID

Cecilia Grela Llerena, GRAPEVINE's Project Technician at CESGA

IBERGRID22 - 13th October 2022

GRANT AGREEMENT UNDER THE CONNECTING EUROPE FACILITY (CEF) - TELECOMMUNICATIONS SECTOR
AGREEMENT No INEA/CEF/ICT/A2018/1837816

CESGA

grapevine

hi**G**h pe**R**form**A**nce com**P**uting s**E**rvices for pre**V**ent**I**on and co**N**trol of p**E**sts in fruit crops

GRAPEVINE project

grapevine

European Commission

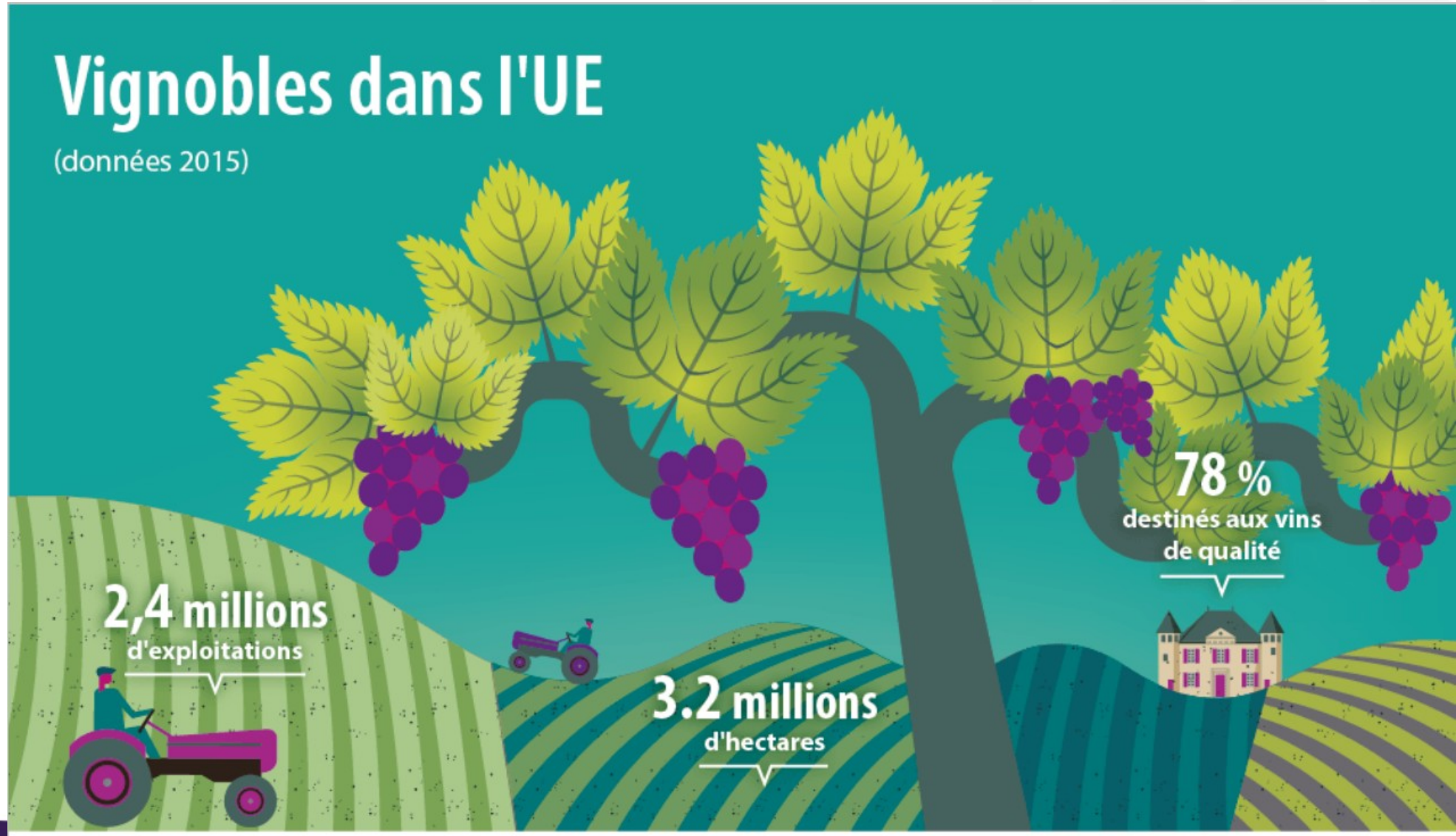
INEA
The Innovation and Networks Executive Agency

Co-financed by the Connecting Europe Facility of the European Union

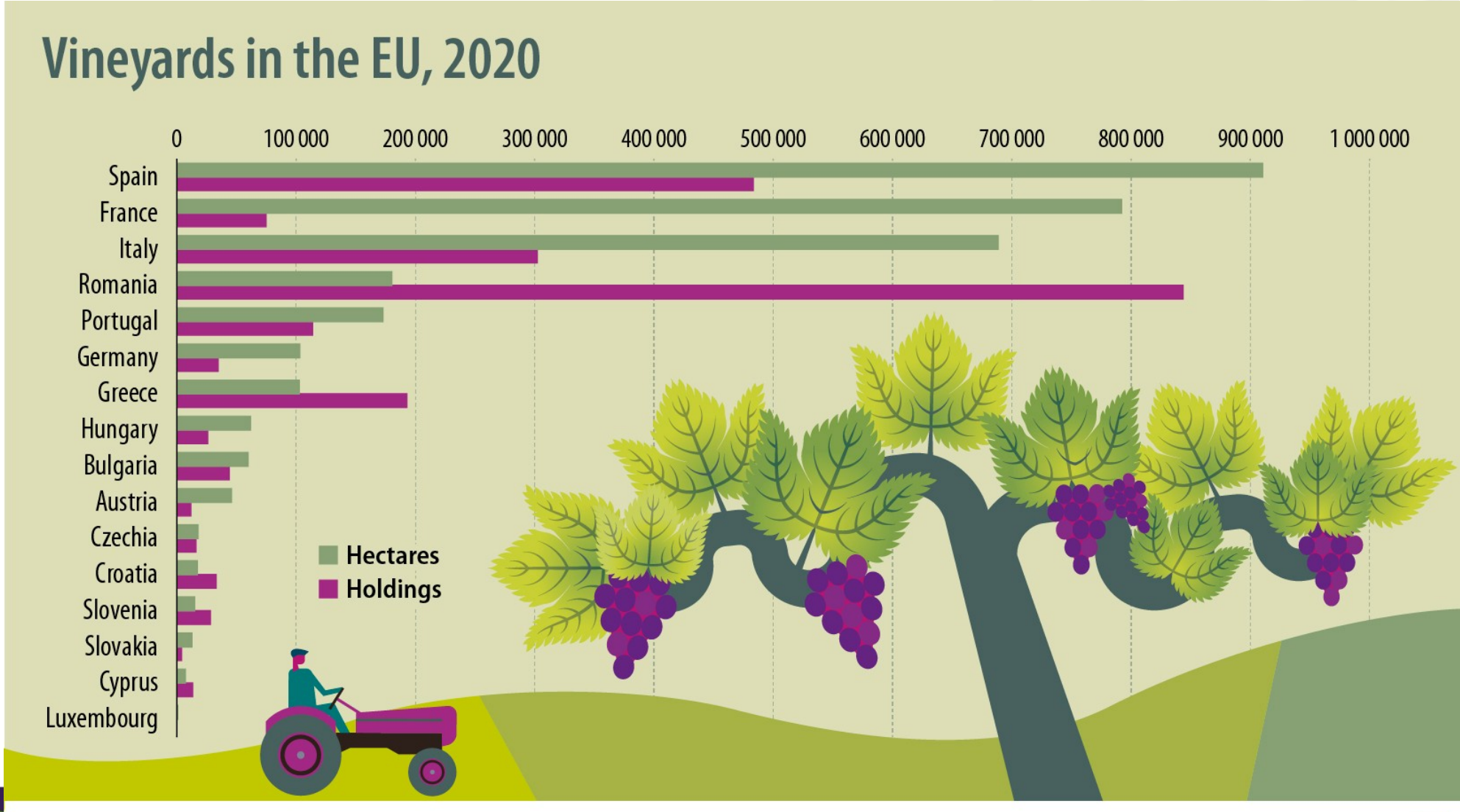
GRANT AGREEMENT UNDER THE CONNECTING EUROPE FACILITY (CEF) - TELECOMMUNICATIONS SECTOR AGREEMENT No INEA/CEF/ICT/A2018/1837816

Action No: 2018-EU-IA-0091

2015 - The EU had 3.2 million hectares of vines, about 45% of the world's wine-growing areas with 2.4 million wine-grower holdings



2020 - The EU had 3.2 million hectares of vines, about 45 % of the world's wine-growing areas but with a sharp reduction of 257.000 in the number of vineyard holdings across the EU.



Only EU Member States having a minimum planted area of 500 hectares (ha) of vineyards are included in the data collection.

Project objectives

- Reducing environmental impact by optimizing the use of phytosanitary products and increasing biodiversity.
- Providing an intelligent decision support tool for grape growers adaptable to other regions.
- Improve farms sustainability.

How do we do it?

- Developing physical models of grape vine phenology and plagues.
- Applying Big Data and Artificial Intelligence technologies to field data, IOT data and remote sensing data (Copernicus Images) for creating logic models of grape vine phenology and plagues.
- Validating physical models and artificial intelligent (AI) models by contrasting their results and by comparing their predictions with field data contained by RedFara, Aragonese Phytosanitary Network of Aragón.

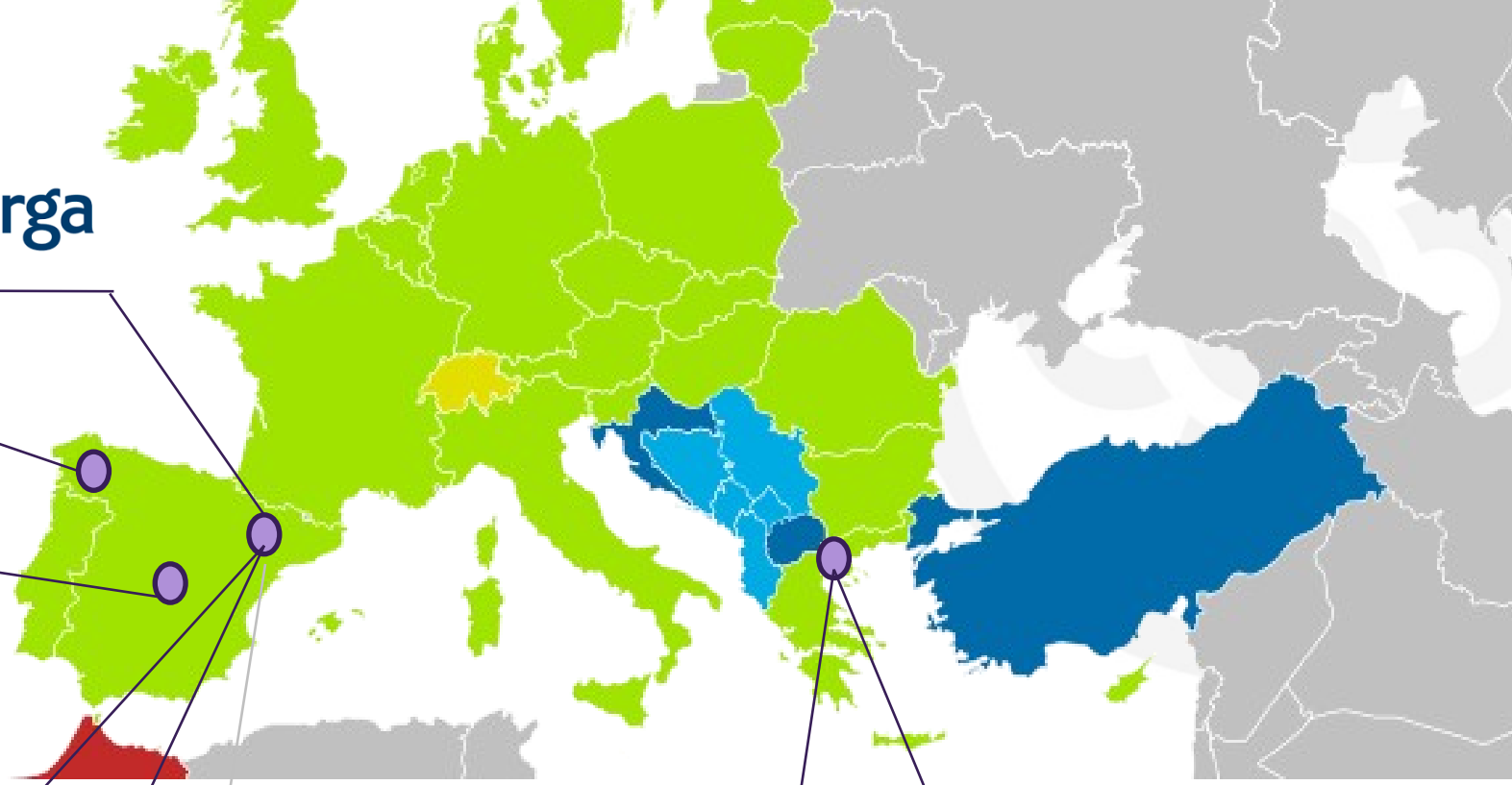
GRAPEVINE Project solution

In the GRAPEVINE the vineyard diseases selected to study are **heavily influenced by meteorological conditions** (temperature, relative humidity, and rainfall).

These diseases are:

- **Downy mildew** (in the photo)
- **Powdery mildew**
- **Black rot**
- **Botrytis**

grapevine



sarga

CESGA

Atos

Universidad Zaragoza

ITAINNOVA INSTITUTO TECNOLÓGICO DE ARAGÓN

ARISTOTLE UNIVERSITY OF THESSALONIKI

AgroApps

With the collaboration of:

CSCV

/ Know-how & contribution of every partner /

1 Atos

- **Project coordinator**.
- HPC support.
- Project management.

2 ITAINNOVA
INSTITUTO TECNOLÓGICO DE ARAGÓN

- Artificial Intelligence.
- Big Data.

3 sarga

- Sensorization and data capturing.

4 Universidad Zaragoza
1542

- Physical or biological models of pests and diseases.

5

- Hyperspectral and multispectral detection
- Integration from multiple sources

6 CESGA
Centro de Supercomputación de Galicia

- HPC infraestructure provision.
- Data visualization.

7 AgroApps

- Weather models.
- Market exploitation.

grapevine project: applying technology to **rural development** and **sustainability**

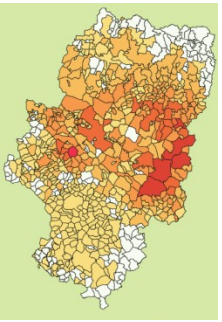
EC (External Collaborator)

CSCV

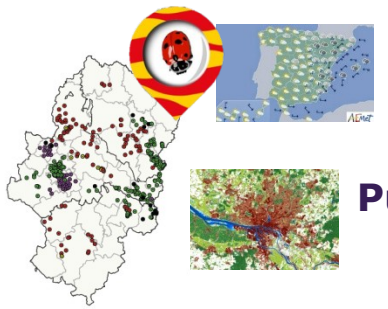
- Knowledge about crop pests and diseases.
- Collection of field data.

GRAPEVINE project

Vineyard field Data



Sensor Data



Public Open Data

/ Conceptual map of *grapevine* /

Many kinds of open data sources

Vineyard Data
- Grapeviners.
- Farmer Cooperatives.
- Wineries.
- Red Fara.

3 8

IoT & Sensors
- Metereological Data.
- Weather Forecast.
- Ground temperature.
- Leaf moisture.

Public & Open Data
- Galileo.
- Copernicus.

1 6

HPC infraestructure

1 6 7 — Visualization

2 — Open DSS Interface

4 8 — Physical Models

2 5 7 — AI Models

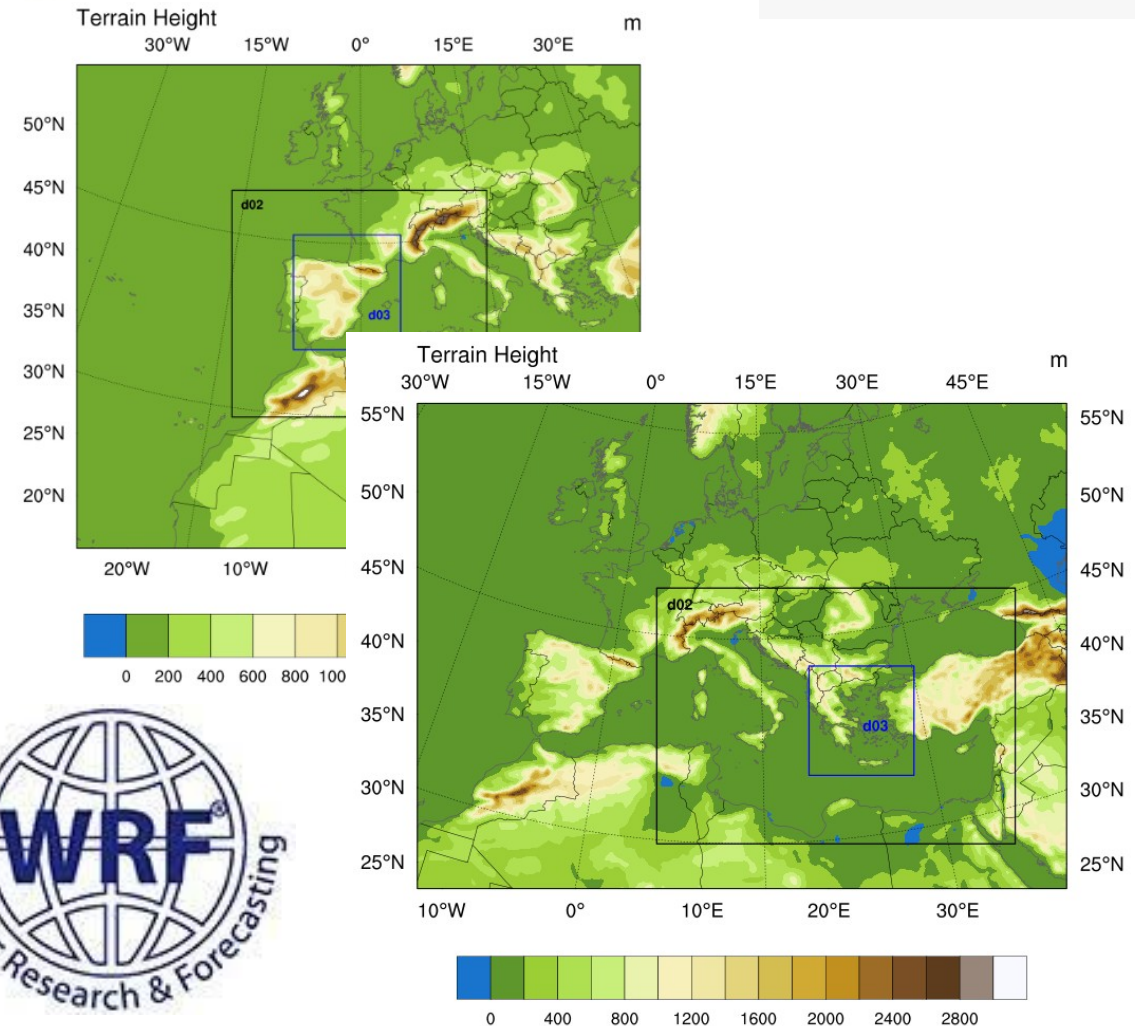
1 2 3 — Data Integration

(POWERED BY AI & MACHINE LEARNING)

grapevine

GRAPEVINE models: meteorological phenological diseases

Meteorological Model

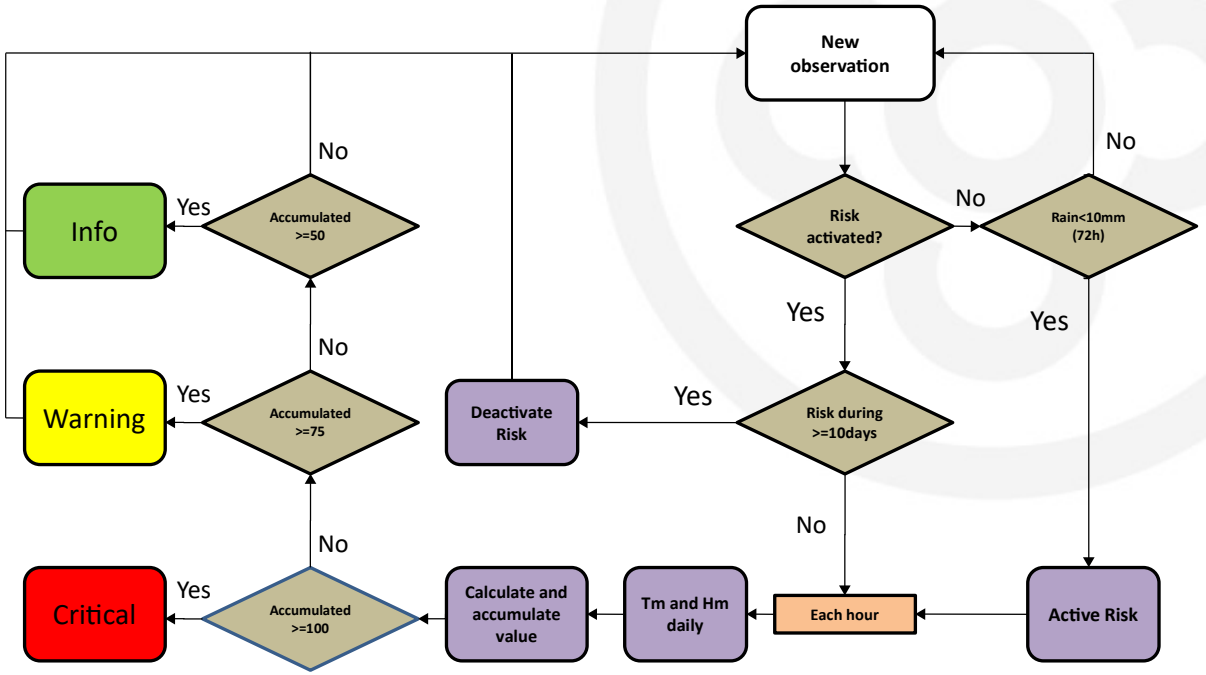


Model Integration Domains

- The WRF model produces **hourly weather forecasts** for up to 7 days time horizon, initated twice a day (00, 12 UTC) at 18km and 6km enclosing the pilots.
- **Sub-seasonal forecasts** are created, with hourly output at a 18km spatial resolution and for 2.5 months forecast horizon. Initiated from CFS once a month (multiple initial conditions to create an enslemble)
- A **THREDDS** server makes the data availble to the rest of the models.

Disease risk Models

- Targeted diseases: Downy mildew, Powdery mildew, Black rot, and Botrytis
- All of them are influenced by weather conditions, such as rainfall, relative humidity, and air temperature.
- The defined models warn for fungal diseases and are launching hourly alerts (Info, Warning and Critical) as the weather conditions indicate it is necessary to do so.
- They are integrated in AgroApps's Meteo Model framework to produce the warnings.



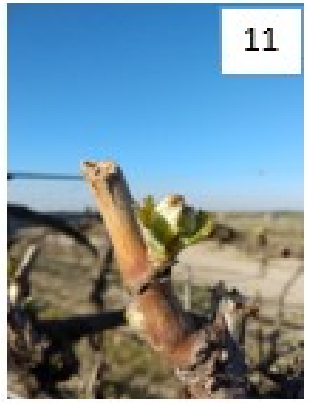
Flowchart of Downy mildew disease model

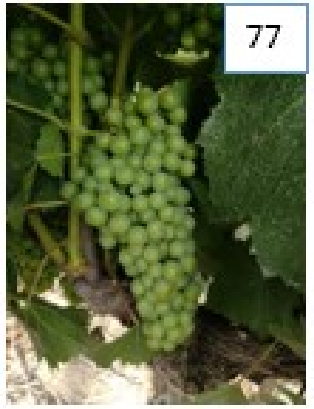
Disease risk Models

Disease / Alert level	No alert	Info	Warning	Critical
Downy Mildew	☑	☑	☑	☑
Powdery Mildew	☑			☑
Black Rot	☑			☑
Botrytis	☑			☑

Phenology Model

The 9 phenological stages modeled have been selected considering the main grapevine pests modeled in the project.





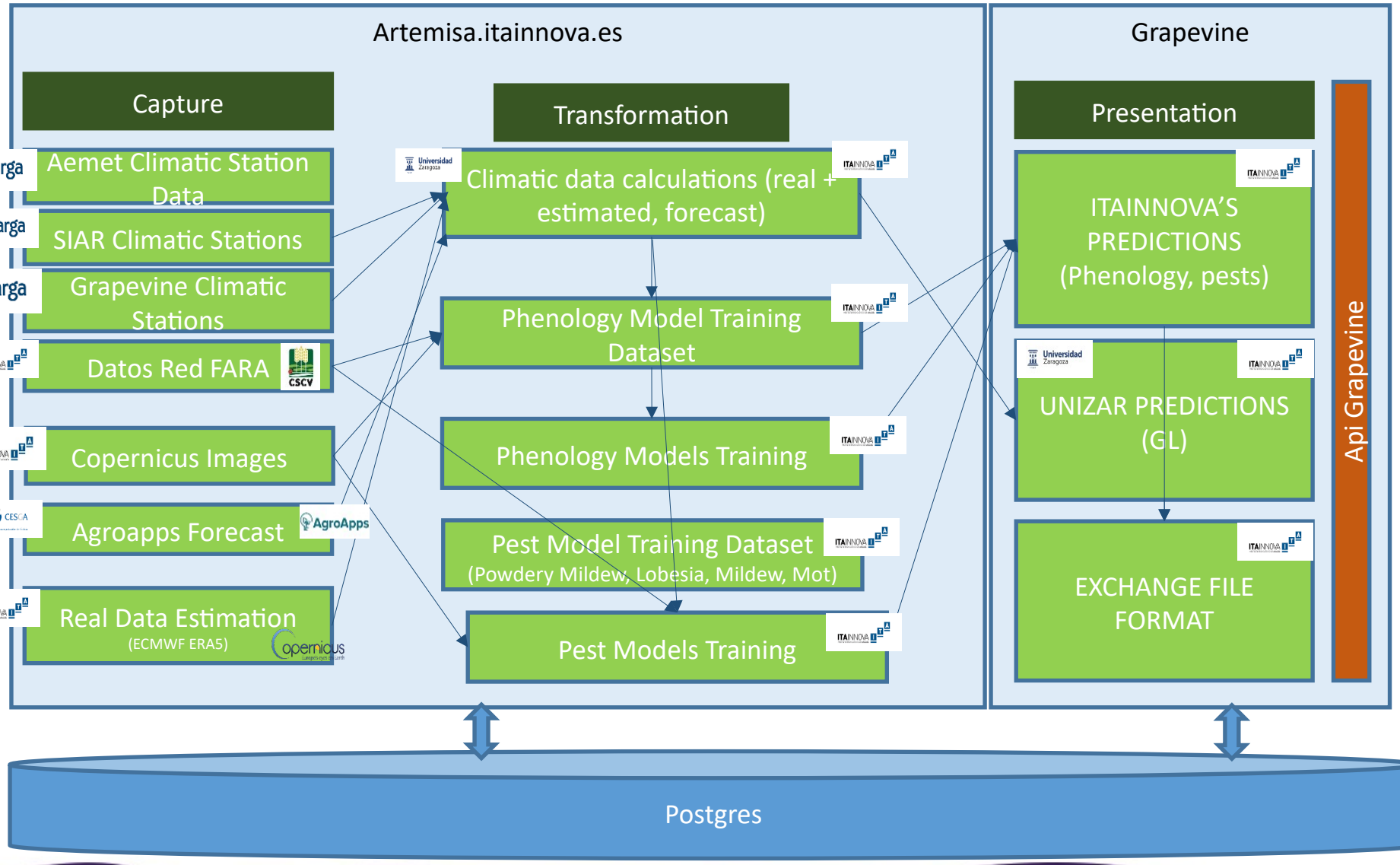
Phenology Model

Phenology Models: biological vs deep learning:

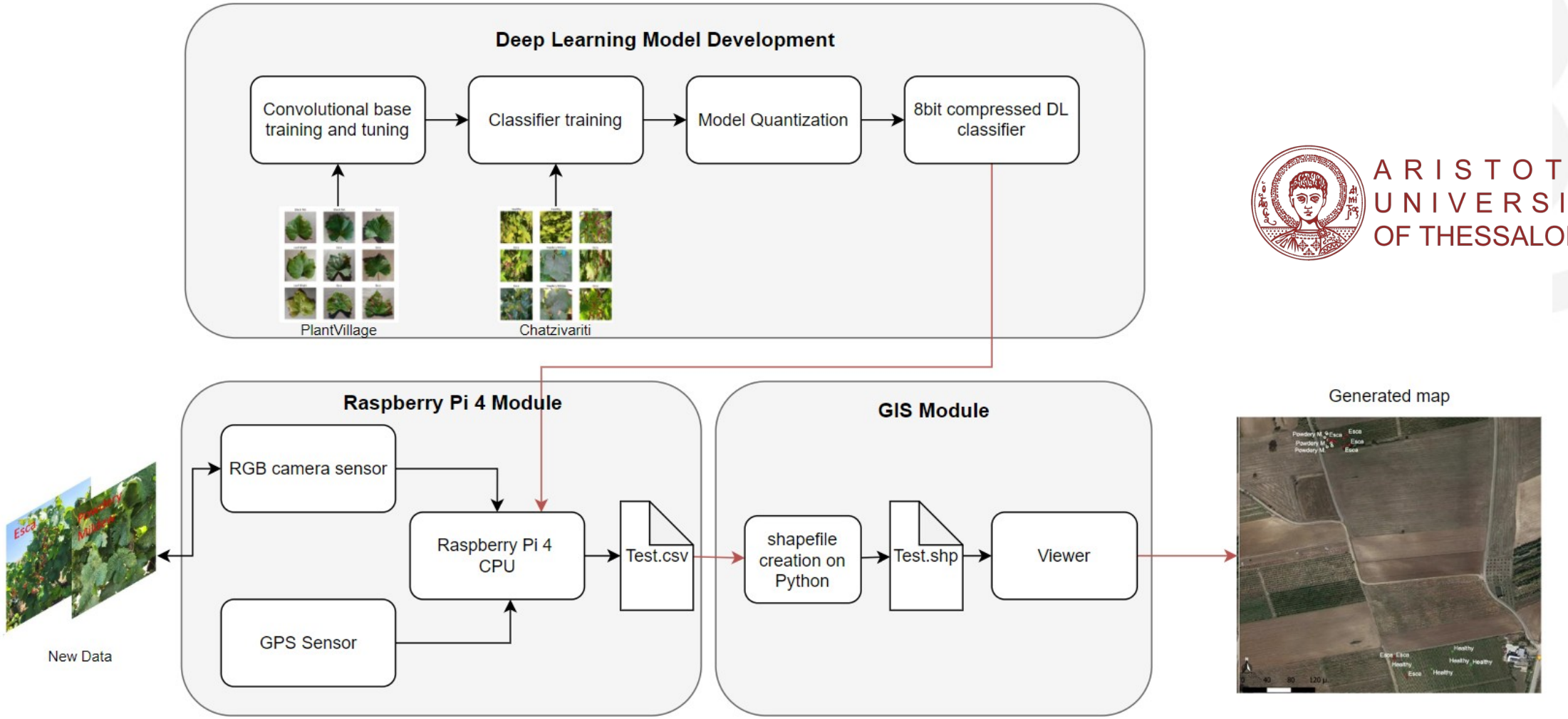
- A **biological model** has been developed by modifying the Winkler index, generating the WI_{10F1} with 10 ^{o}C base temperature accumulated since February 1st. This model only uses the daily temperature record.

- The **deep learning model** was able to consider all available data to train and validate its results. Finally, it ended up selecting the day of the year as the variable most related to phenology, ahead of WI_{10F1}. The model is able to predict the growth stage of the plant for a given day.

Deep Learning model - Modules Overview



Deep Learning model - disease recognition from RGB images



ARISTOTLE UNIVERSITY OF THESSALONIKI

grapevine

PROJECT'S ECOSYSTEM

Infrastructure
Data management and Storage
Orchestrator

GRAPEVINE INFRAESTRUCTURE

Standard APIs

Visualization

Integrated Model

Behavioural Model

Meteo Model

Phenology Model

DL Models

Data Manager

SLA Manager

Monitoring

Orchestrator

GRAPEVINE

HPC

CLOUD

CESGA

CESGA - Data Management

- **Shared storage** available for the model results:
  - Grows with the project.
  - Accessible from the FTII, FTIII, the THREDDS server and the Kubernetes cluster.
- Increased collection of project **datasets**
- PostgreSQL data base available for **visualization** purposes
- **THREDDS** server:
  - Provides access to the AGROAPPS meteorological model results through several methods: web interface, REST API.
  - Used by the models to obtain the data needed from the meteorological model.
- **CKAN** for data catalogs
- **Gitlab**: https://gitlab.com/grapevine-project/grapevine

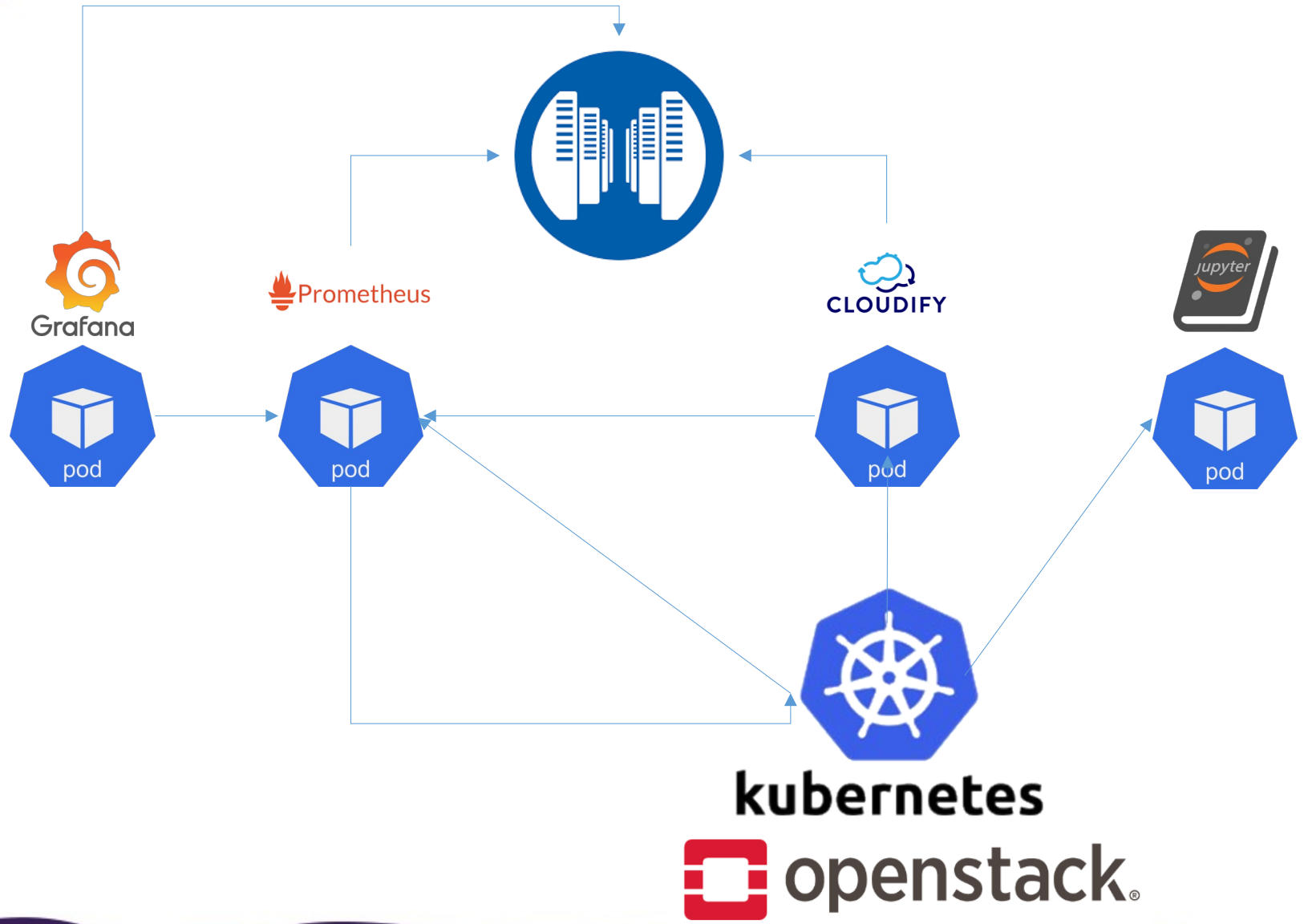
Data Management - THREDDS

Dataset

- One .00 file test
- One .00 file test all services
- One folder, all files and all services/
- One folder, all files and only ncss/
- 2021-04-22 - All files except .log, .out and .txt/
- 2021-04-23 - All files except .log, .out and .txt/
- 2021-04-24 - All files except .log, .out and .txt/
- AGROAPPS-CLIMATE-MODEL - All files except .sh, .log, .out and .txt/

THREDDS - AGROAPPS METEOROLOGICAL RESULTS - GRAPEVINE PROJECT at CESGA see Info Documentation

ORCHESTRATOR: Deployment Architecture



ATOS - Resources Management through the Orchestrator

- *Main Features*
  - Facilitate the execution of GRAPEVINE models in an optimized way, using HPC+Cloud resources
  - Enable an autonomous management, depending on applications behavior and current context → Link to Monitoring & SLA
  - Facilitate the usage of data management tools → Data movement tasks in workflows.
- *Improvements*
  - New infrastructures supported → CESGA FinisTerrae III + CYFRONET
  - Cloudify with Croupier plugin
    - Support for multiple HPC and Cloud solutions
    - Integration of data connectors and data management
    - Access to resource reservations and periodic executions
  - Blueprints available for Meteorological predictions
  - Analysis of data from queues / partitions usage:
    - Estimation of the best queue based on the HPC provider selected

grapevine

HPC RESOURCES

HPC simulations: AGROAPPS simulations, HPC resources, resource reservations

AGROAPPS METEOROLOGICAL MODEL

AgroApps	Responsible Team
WRF Weather Research & Forecasting	Numerical Weather Prediction Model
MET Model Evaluation Tools	Model Evaluation

- ✓ Operational execution on HPC
- ✓ Data Assimilation
- ✓ Model Evaluation

AGROAPPS METEOROLOGICAL MODEL

Agroapps is providing the meteorological data and forecast skill by running a weather forecasting model and a model evaluation tool. The **medium-range forecasts (**at 18, 6 and 2km spatial resolution) are provided twice a day for two domains, Spain and Greece.

The services run daily are:

- GV_R${DATE}_00_G: The forecasts initiated at 00 UTC for the Greek domain.
- GV_R${DATE}_00_S: The forecasts initiated at 00 for the Spanish domain.
- GV_R${DATE}_12_G: The forecasts initiated at 12 UTC for the Greek domain.
- GV_R${DATE}_12_S: The forecasts initiated at 12 UTC for the Spanish domain.

AGROAPPS METEOROLOGICAL MODEL

For the **evaluation** of the forecasts against reanalysis data, daily two services are run:

- MET_00: Evaluation of weather forecasts initiated at 00 UTC against re-analysis data.
- MET_12: Evaluation of weather forecasts initiated at 12 UTC against re-analysis data.

The aforementioned services were run daily, operationally.

AGROAPPS METEOROLOGICAL MODEL

To run the **sub-seasonal forecasting** and to produce weather forecasts for two and a half months ahead, this configuration will be followed:

- Each month for 10 subsequent days, 4 runs a day (SEAS_00, SEAS_06, SEAS_12, SEAS_18).

AGROAPPS METEOROLOGICAL MODEL

Once AGROAPPS started to run the model in the FTIII, the **sub-seasonal forecasting** configuration was:

- Each month for 11 days, 6 hours and 30 minutes:
  - 4 runs a day (SEAS_00, SEAS_06, SEAS_12, SEAS_18).

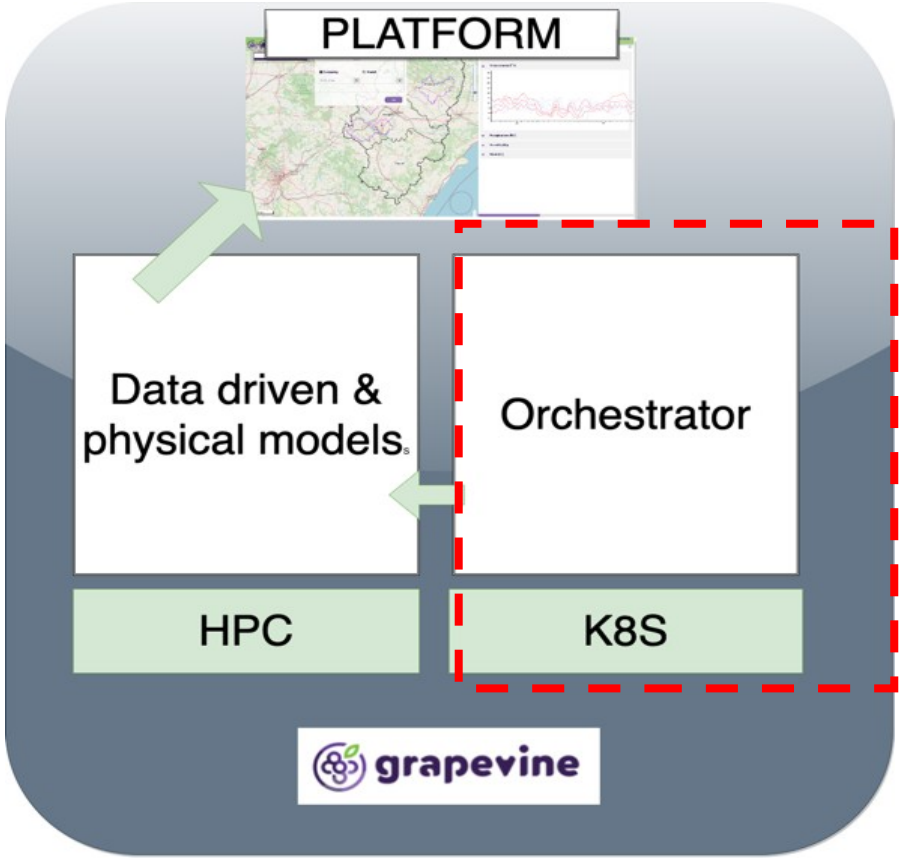
AGROAPPS METEOROLOGICAL MODEL

Reservation of resources were needed as some simulations (for example MET) need the results of previous simulations to be run.

CESGA created resources reservations for all parts of the model based on AGROAPPS request of resources.

The resources reserved were: full nodes and memory for an specific duration starting an specific date and hour.

Grapevine Tool: Conceptual map

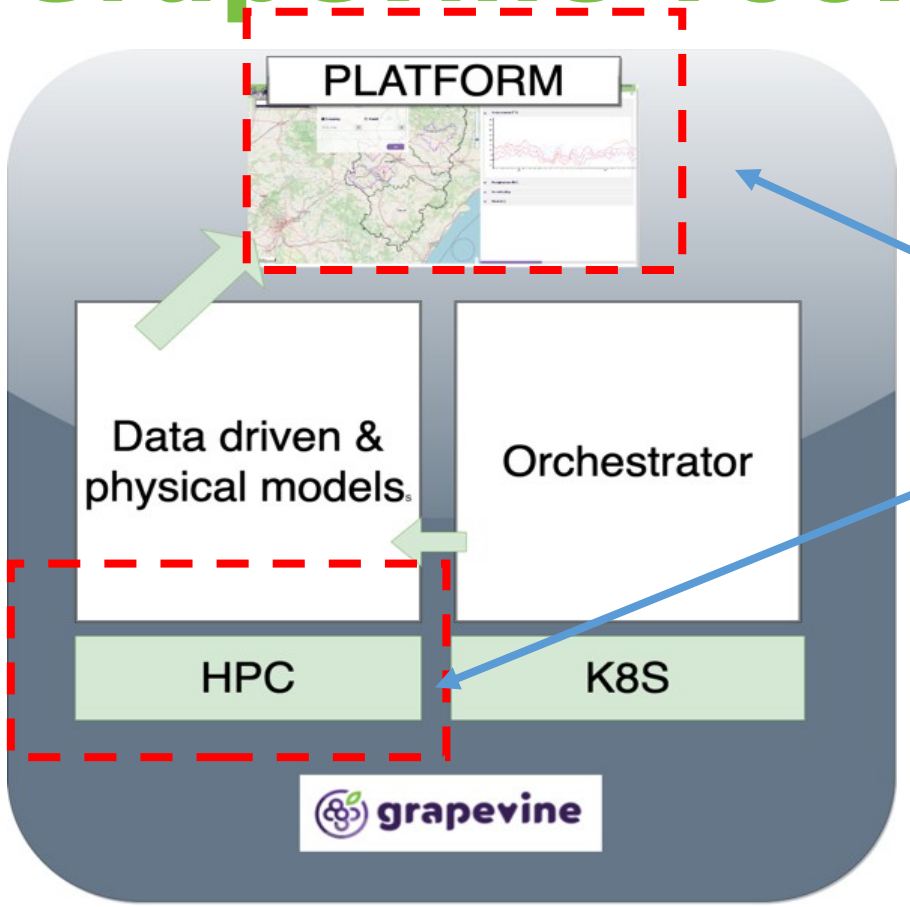


Atos

CESGA
Centro de Supercomputación de Galicia

AgroApps

Grapevine Tool: Conceptual map

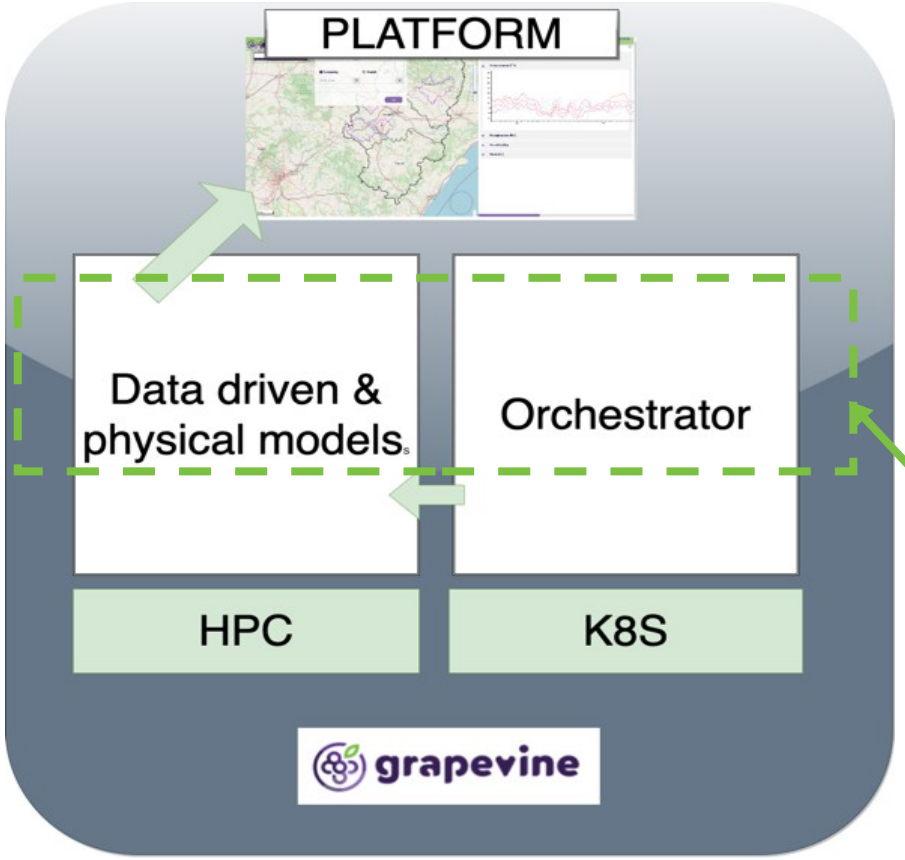


Atos

CESGA
Centro de Supercomputación de Galicia

AgroApps

Grapevine Tool: Conceptual map

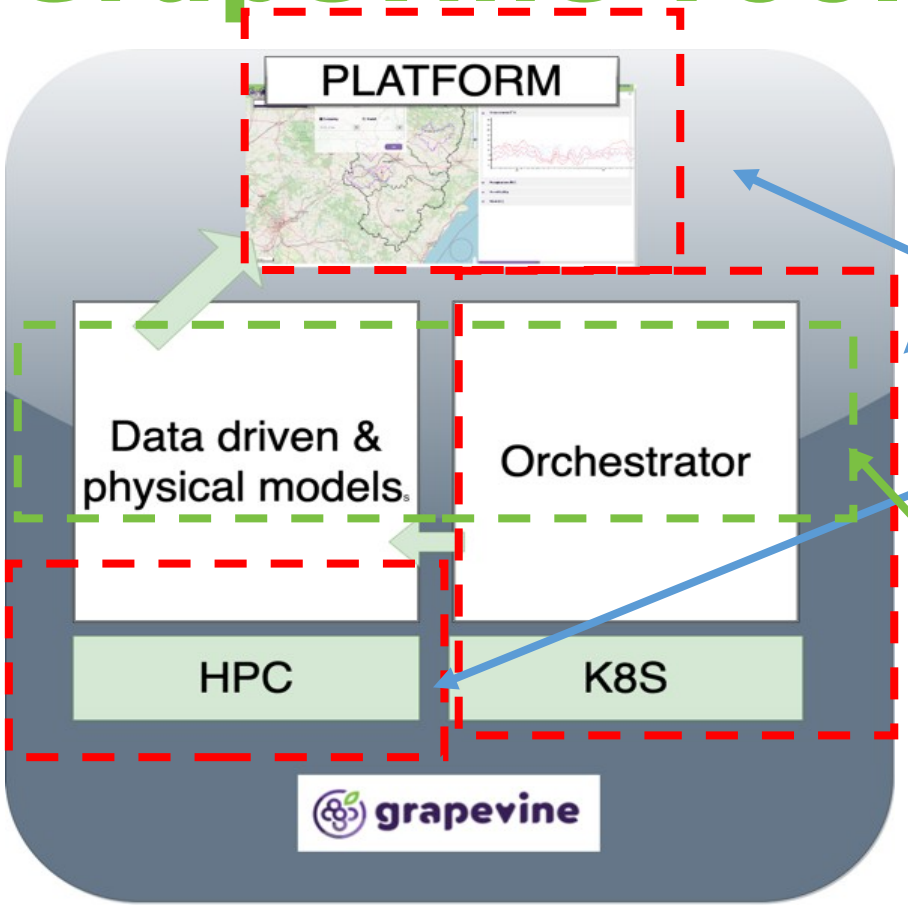


Atos

CESGA
Centro de Supercomputación de Galicia

AgroApps

Grapevine Tool: Conceptual map



Atos

CESGA
Centro de Supercomputación de Galicia

AgroApps

grapevine

INTEGRATION IN EOSC

HPC resources for the project
EGI-ACE project – project's integration in EOSC

By early October 2021 CESGA and AGROAPPS **estimated the HPC resources** needed for the simulations exceeded the amount available in the project.

HPC HOUR CONSUMPTION PROBLEM

Partners **tried to optimize the simulations**:

- Review the models' needs to simulate only what is really needed.
- Testing if the changes will affect the quality of the results.

The simulations in 2021 consumed all the available HPC resources for the whole project:

2.233.000 core-hours

HPC RESOURCES

The meteorological model consumes the highest amount of HPC resources.

An estimation for 12 month of resources with the previous configurations were:

4.226.400 core-hours

As a result of this issue in mid-November:

- Project **activities** that needed access to HPC resources were **stopped**.
- Partners decided that a new solution was needed to continue the simulations.

A significant delay in the project.

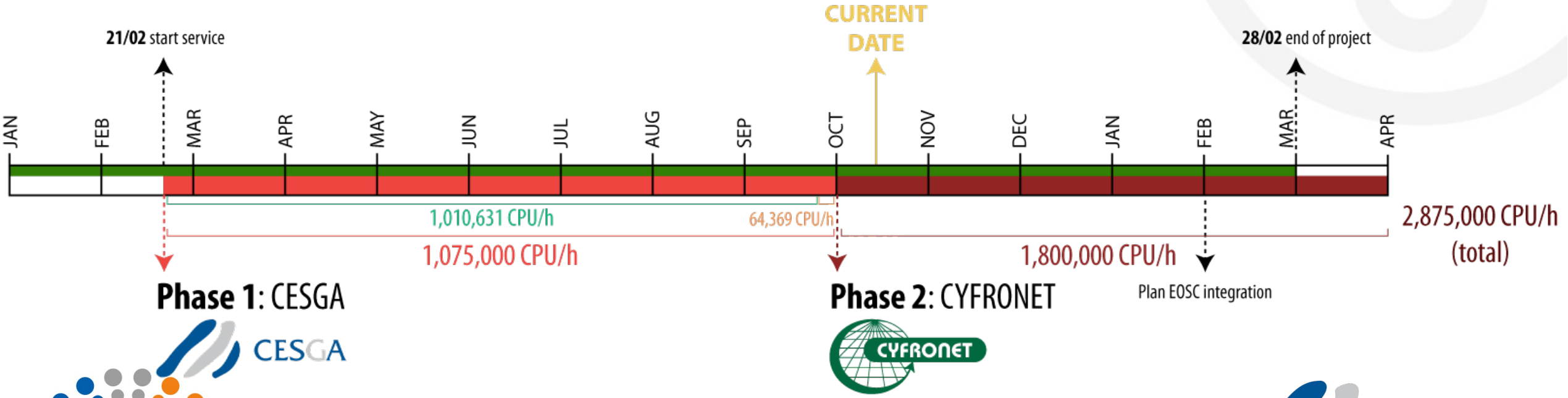
HPC RESOURCES

The delay could be fatal for the project so the consortium started searching for other **solutions**:

- Move resources in the internal budget of the project to obtain new HPC resources.
- Availability of partners infrastructures.
- Reduction of the model's HPC needs.
- Search for additional resources outside the project:
  - EGI-ACE open call for Use Cases.

EGI-ACE OPEN CALL FOR USE CASES – EOSC

An application to the EGI-ACE project was made in December 2021. By mid-January the consortium received a communication that the application was accepted.



2.875.000 core-hours assigned to the project in two providers:

- **CESGA**: 1.075.000 core-hours
- **CYFRONET**: 1.800.000 core-hours

Resource Centre CESGA — Elapsed time * Number of Processors (hours) by VO and Month (Custom VOs)

VO	Mar 2022	Apr 2022	May 2022	Jun 2022	Jul 2022	Aug 2022	Sep 2022	Total
vo.grapevine.eu	50,553	149,529	190,984	155,733	218,833	136,433	108,565	**1,010,631**

MIGRATION TO CYFRONET

Migration of AGROAPPS meteorological simulations from CESGA to CYFRONET (https://www.cyfronet.pl/en) involve:

- **Module** availability
- Adapt and test the **simulations.**
- **Correct** any errors found.
- **Data transfer**: results have to be moved from CYFRONET's storage to CESGA's storage to maintain the projects environment.

grapevine

Thank you for your attention!

https://grapevine-project.eu/

CESGA

https://twitter.com/grapevine_pro

https://www.linkedin.com/company/grapevine-project

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº INEA/CEF/ICT/A2018/1837816

Atos · Aristotle University of Thessaloniki · ITAINNOVA · Universidad Zaragoza · CESGA · sarga · AgroApps