ORGANISED BY

BDV BIG DATA VALUE ASSOCIATION

HOSTED BY

COSMOTE

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Odyssey of AI:

Navigating the Data Seas

27 - 28 May / Athens

IN COLLABORATION WITH

ATC ATHENS TECHNOLOGY CENTER

ΕΠΙΣΕΥ ICCS

core|IC

DEMOKRITOS NATIONAL CENTRE FOR SCIENTIFIC RESEARCH

FORTH INSTITUTE OF CHEMICAL ENGINEERING SCIENCES

NEUROPUBLIC Information Systems & Technologies

data-week.eu
#DataWeek25

GeoAI, Generative AI, Foundational Models and EO

Moderated by Karel Charvat (Plan4all

Main ideas

- Foundation models trained on petabyte-scale EO corpora minimise labelling effort and expedite transfer learning.
- Generative networks generate cloud-free Sentinel-2 mosaics, improving numerical weather prediction and surface-state monitoring.
- Multilingual LLM interfaces integrate EO, in-situ and textual data, removing language barriers and delivering agronomic knowledge directly to farmers.
- Federated fine-tuning propagates plot-level patterns to regional products while maintaining data sovereignty.
- Multi-agent GeoAI pipelines merge detection, forecasting and optimisation for resilient agro-forestry operations.

Large Language Models for Agriculture and Rural Development

by (Stein) Runar Bergheim (Asplan Viak Digital Services)

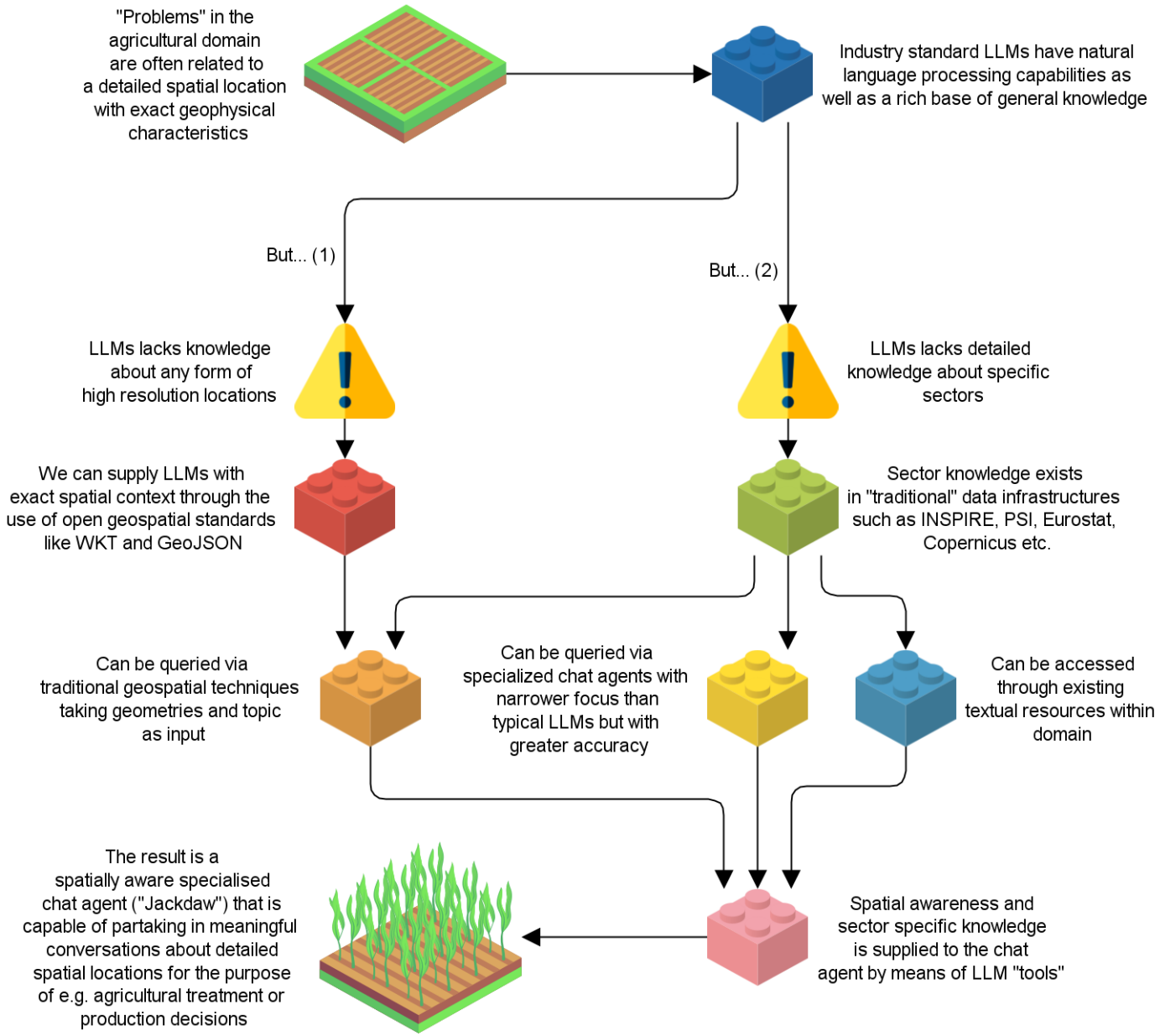
Hello, my name is (Stein) Runar (Bergheim)

I have been working with “data” for just about 25 years, first with tables, then with maps, now with IoT and AI

Have worked under the banner of my beloved Norwegian employer Asplan Viak for most of that time on 300+ projects in 25+ countries.

The Challenge

- We wish to talk to LLMs about specific locations
- In the context of agricultural decisions
- In the context of rural development decisions



The Preconditions

Agriculture

- High resolution in time and space
- Working on field level
- Working in «the present»
- Not specialised IT-users

Rural development

- High resolution in space
- Intermediate resolution in time
- Can expect some skills among the work force

JACKDAW Just-in-time AI Coordinator for Knowledge Discovery and Analytics on a World-scale

ME:
Could you characterize this area in terms of land use, industry and climate please?

JACKDAW:
Here's a summary of the characteristics of the area in terms of land use, industry, and climate:

Land Use

- **Forestry**: The predominant land use, covering approximately 73.4% of the area (57.31 km²).
- **Residential Use**: Accounts for about 11.3% (8.82 km²).
- **Transport Networks**: Comprises around 6.0% (4.71 km²).
- **Agriculture**: Represents about 3.1% (2.41 km²).
- **Transitional/Natural Areas**: Covers 2.1% (1.63 km²).
- **Cultural and Recreational Services**: Makes up 1.6% (1.27 km²).
- **Mining and Quarrying**: A small portion at 0.9% (0.69 km²).
- **Other Industries**: Includes logistics and utilities, making up less than 1%.

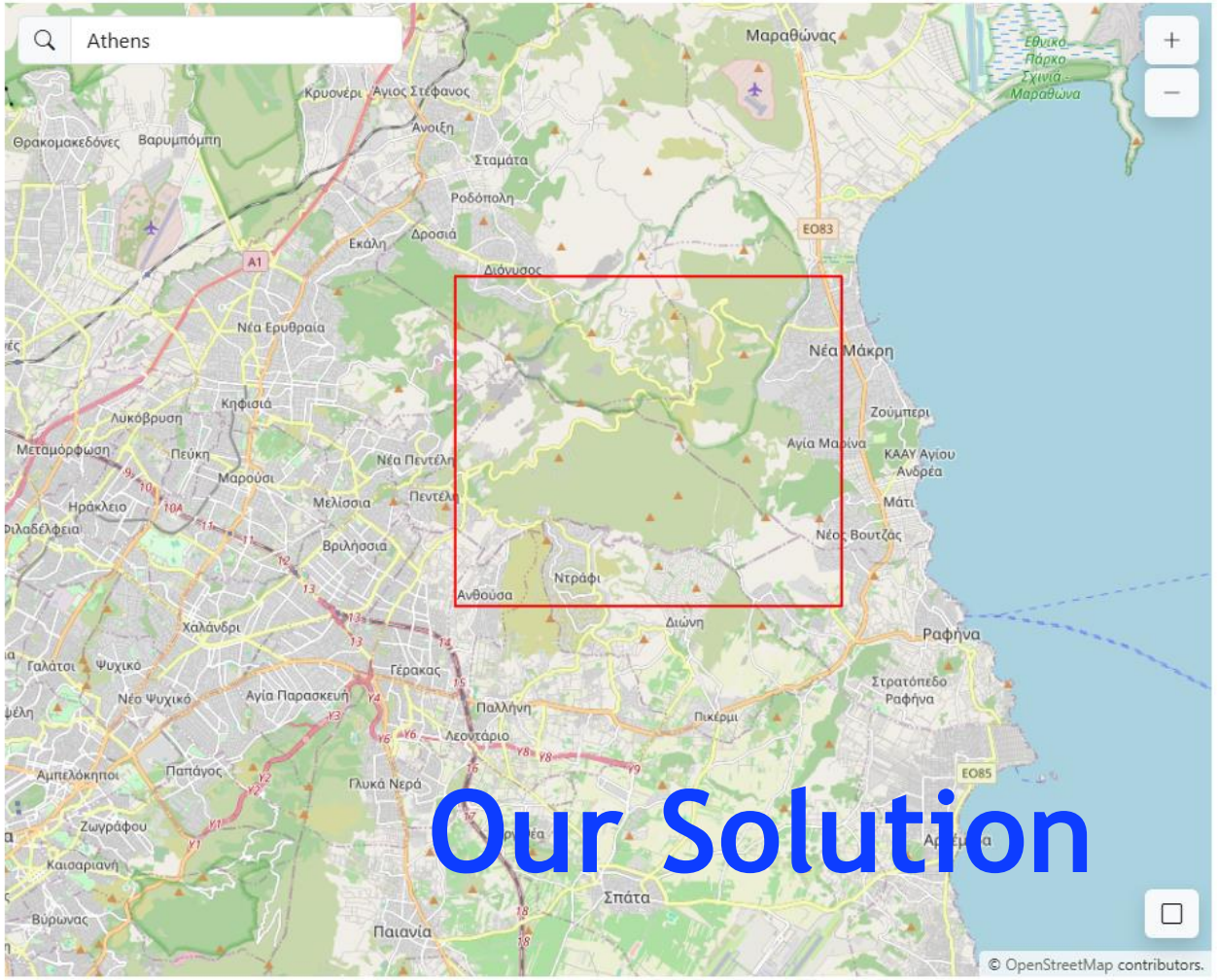
Industry

The area has a diverse industrial sector with significant employment across various sectors:

- **B-E Sector**: 40,401 employees, with a growth rate of 57.
- **F Sector**: 16,531 employees, with a growth rate of 31.
- **G Sector**: 84,493 employees, showing a growth rate of 122.
- **H Sector**: 13,108 employees, with a growth rate of 26.
- **I Sector**: 29,858 employees, with a growth rate of 95.
- **J Sector**: 34,063 employees, with a growth rate of 69.

Draw an area in the map and ask a question...

New chat



Our Solution

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

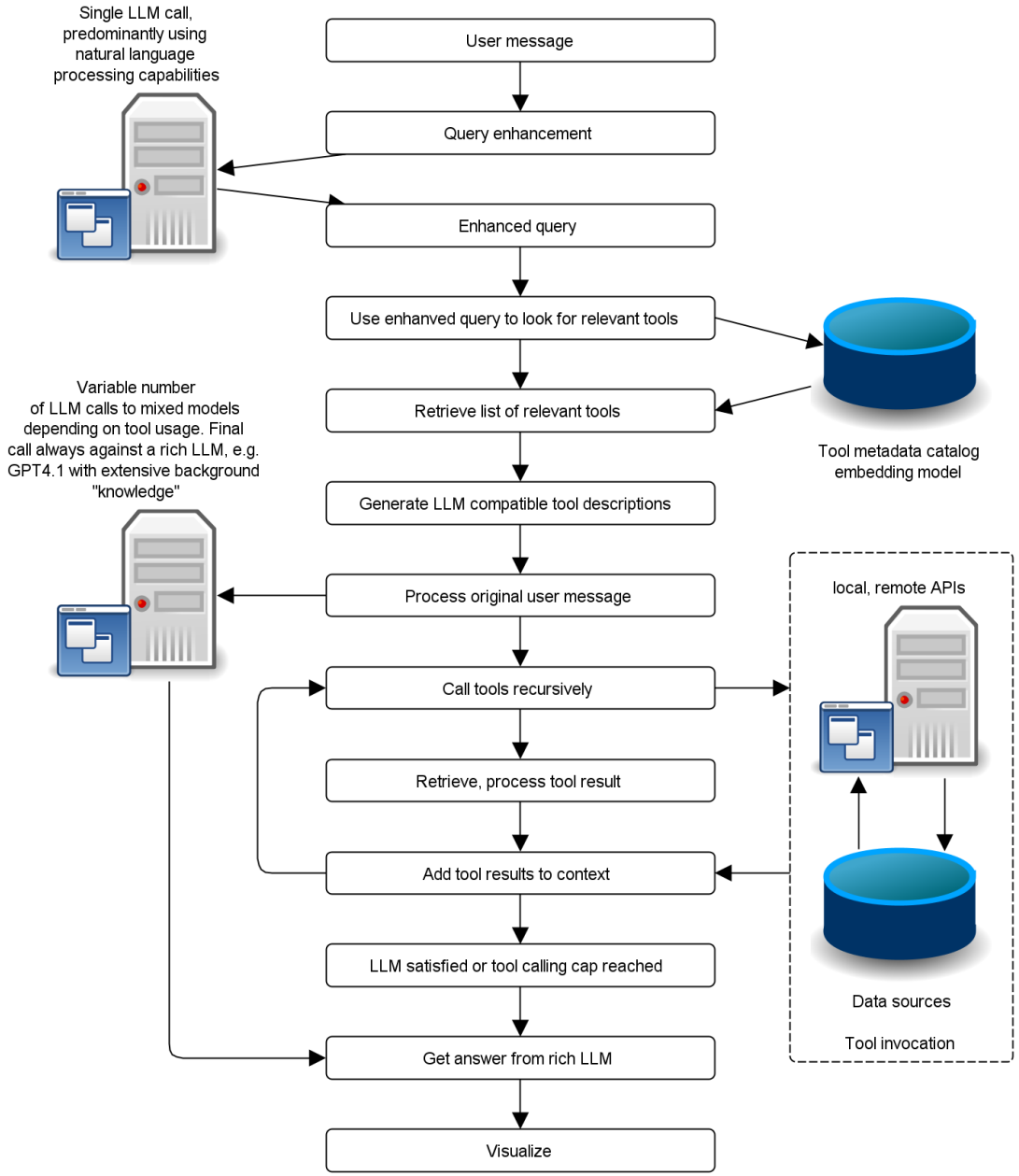
ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

HOSTED BY COSMOTE

IN COLLABORATION WITH ATC, CORE|IC, FORTH, ICCS, DEMOKRITOS, NEUROPUBLIC

How it works

- Orchestration of data sources, agents, tools and a generic LLM through a multi-step process
- Context augmentation prior to asking the original question
- Making «Jackdaw» an expert in spatial and socio-economic characteristics



The way forward

- Expanding our agent to handle hundreds, thousands of data sources
- Enabling users to add their own data sources
  - Spatial -> Geospatial analytics services -> Dynamic tools
  - Tabular -> Data analytics services -> Dynamic tools
  - Documents -> Rag generation -> Shared, generic tool

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

Many thanks

(to those who paid attention)

For further info, you can find me here…

Linkedin: @runarbe
E-mail: runar.bergheim@asplanviak.no

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

How to improve weather forecast and remove clouds from Sentinel 2 using AI (ALIANCE)

Karel Charvát jr. (LESPROJEKT-SLUŽBY s.r.o.)

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

HOSTED BY COSMOTE

IN COLLABORATION WITH ATC, CORE|IC, FORTH, DEMOKRITOS, NEUROPUBLIC

The project

- Czech national project co-financed by state support from the Technology Agency of the Czech Republic under the TREND Program.
- Czech partners developing innovative and technologically advanced solutions to support precision agriculture using artificial intelligence, satellite imagery, and sensor measurement

Aliance

HS - RS | ČVUT - FIT | WIRELESSINFO
ZS ČR | LESPROJEKT-SLUŽBY | BADKID

HS-RS

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

WIRELESSINFO

Zemědělský svaz České republiky

LESPROJEKT služby

badkid.

The Challenge

- Crop growth is influenced by many factors
- Cloud cover is a significant obstacle for satellite-based crop monitoring
- Weather is critical for agriculture
- The farmer needs forecasts tailored to his needs where, what, and when to forecast.
- Existing weather forecasts are too general, not local enough, do not create easily actionable predictions
- Integration of data sources is crucial

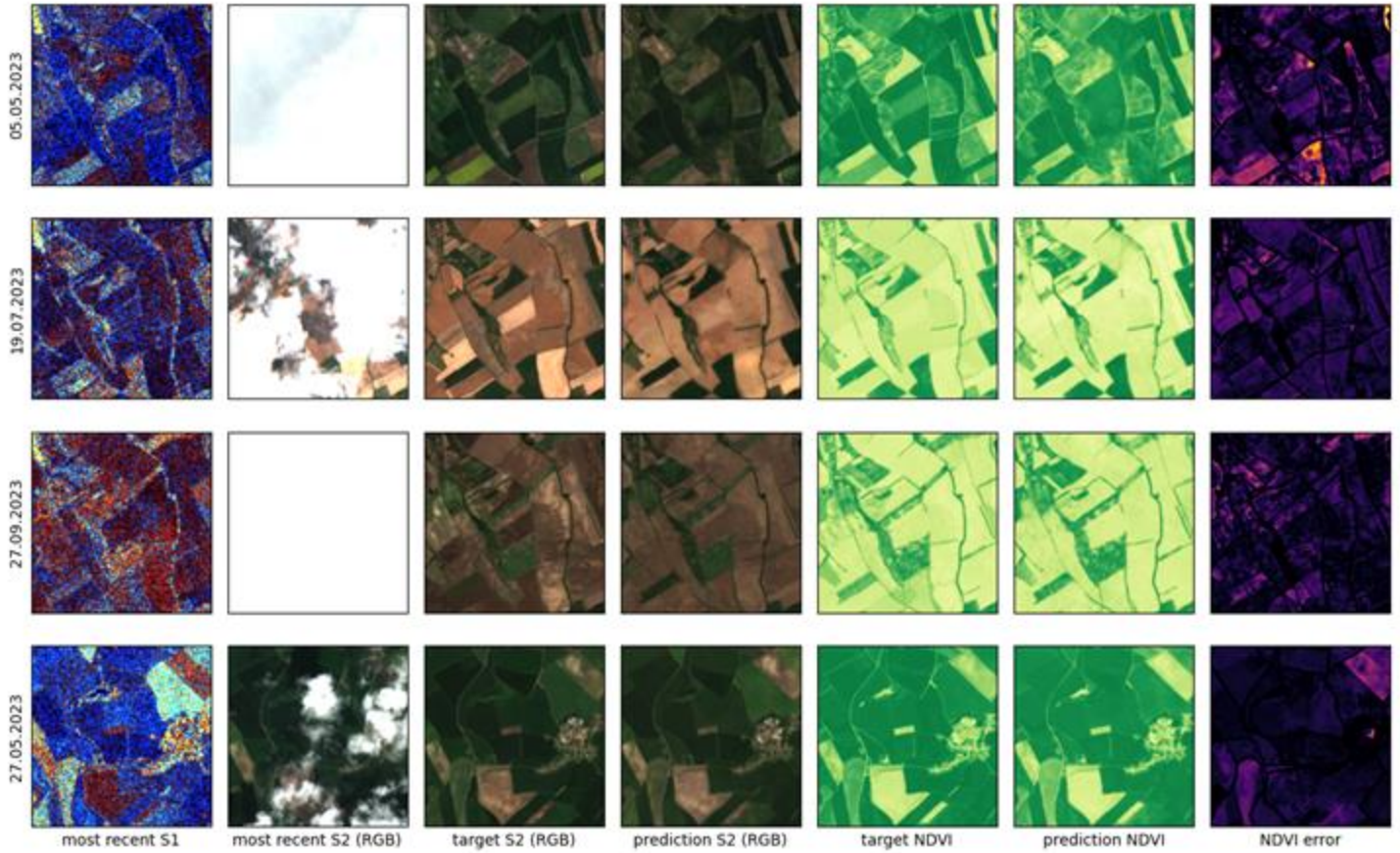
The Solution: cloud removal

- A combination of many available sources, Sentinel 2, Sentinel 1, weather, soil type, crop type, land use, elevation
- Developing an advanced deep-learning diffusion model that allows smooth assimilation of all the data
- The AI model must allow to quantify uncertainty in the prediction

The Solution: cloud removal

- We trained the AI model using thousands of patches overcentral Europe
- We choose only those times with a clean Sentinel 2 as a target
- We used the last five Sentinel 1 and Sentinel 2 (both cloudy and clean) images
- We minimalize mean squared error for RGB and also for NDVI
- The AI model allows uncertainty quantification

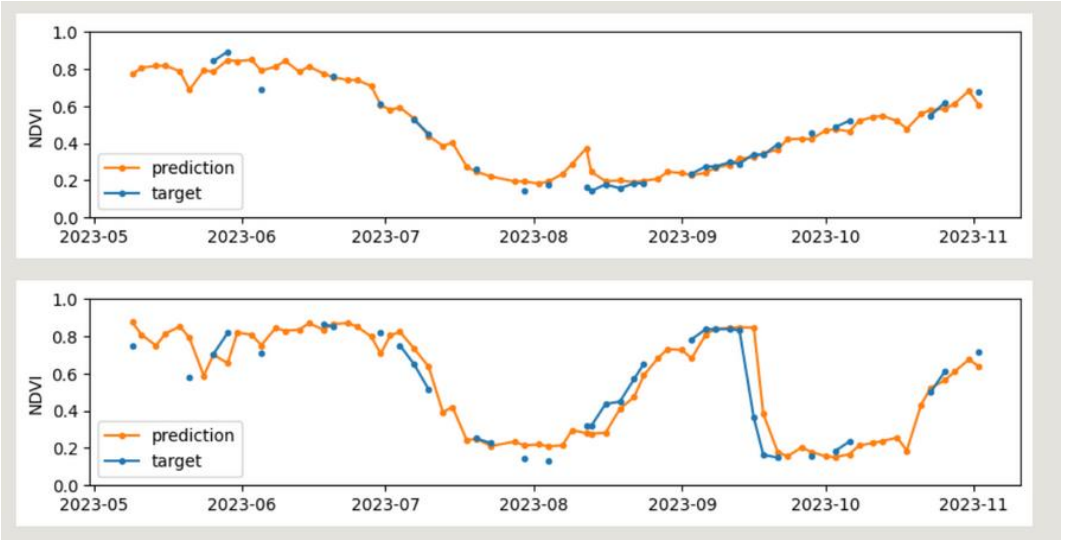
Tangible Results so far: cloud removal



Tangible Results so far: cloud removal

We created an AI model that allows:

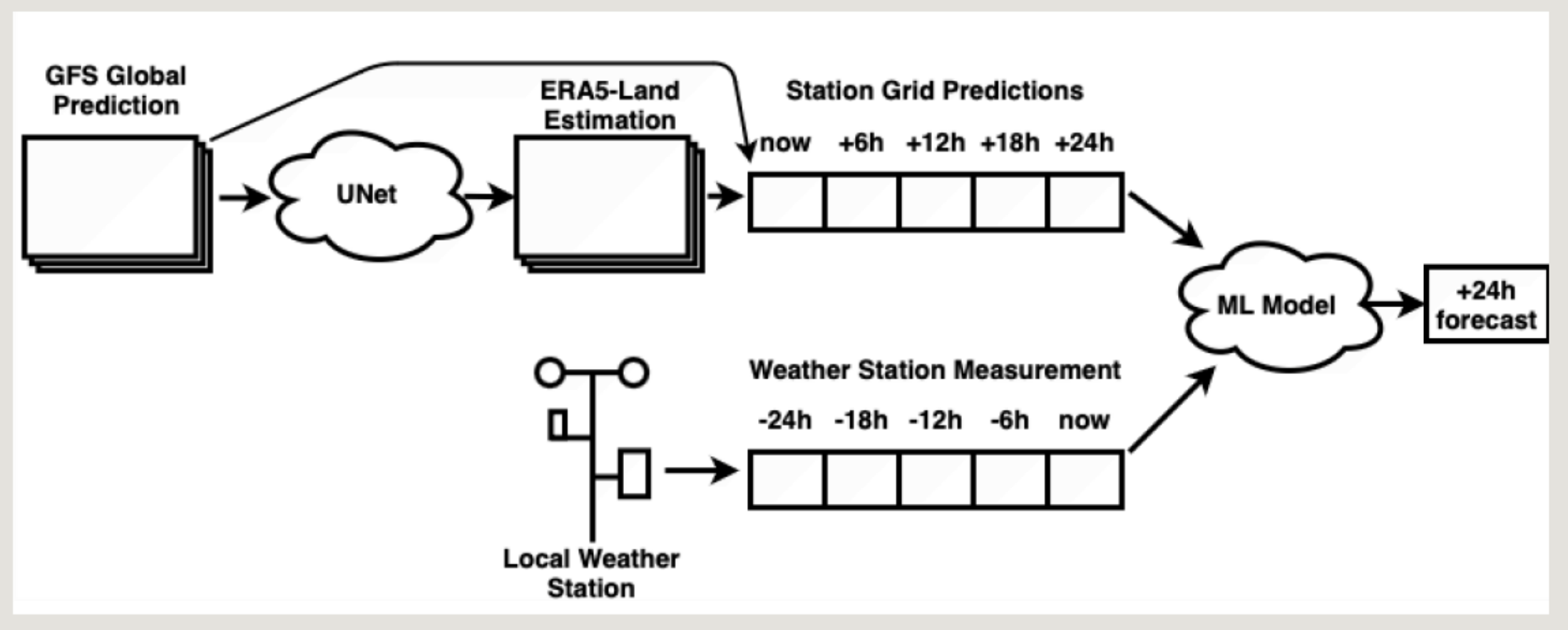
- Cloud removal
- Prediction of the next Sentinel 2 images (9 days ahead)
- Uncertainty quantification
- The model is state-of-the-art in cloud removal and shows good capabilities far from the training domain



The Solution: local weather forecast

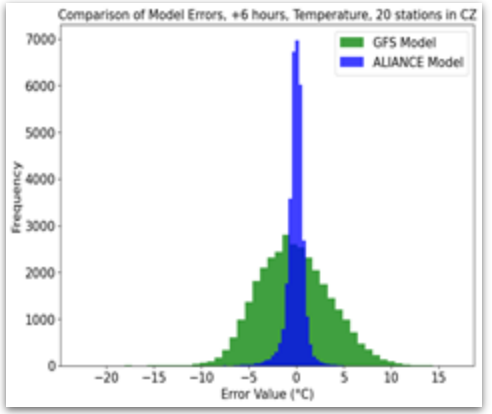
- Combining local weather station data with global forecasts
- We use historical data from farmers' stations together with the GFS global forecast
- We first improve the global model prediction using deep-learning model postprocessing
- We merge this prediction data with the last 24 hours of data from the weather station

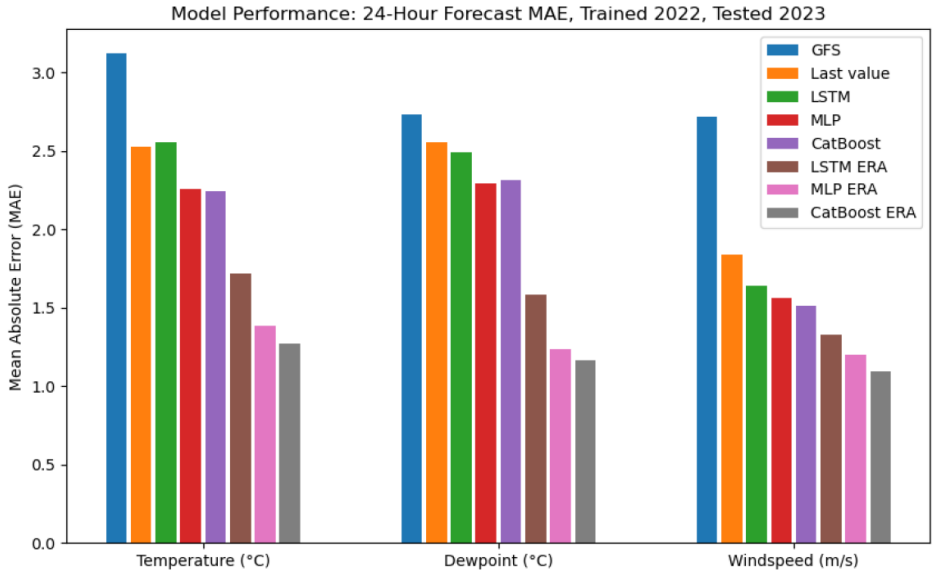
The Solution: local weather forecast



Tangible Results so far: local weather forecast

- We need only 4-6 months of local weather station data!
- We predict anything that the weather station can measure: soil moisture, ground temperature, leaf moisture





Radovan Hilbert

Sales development director

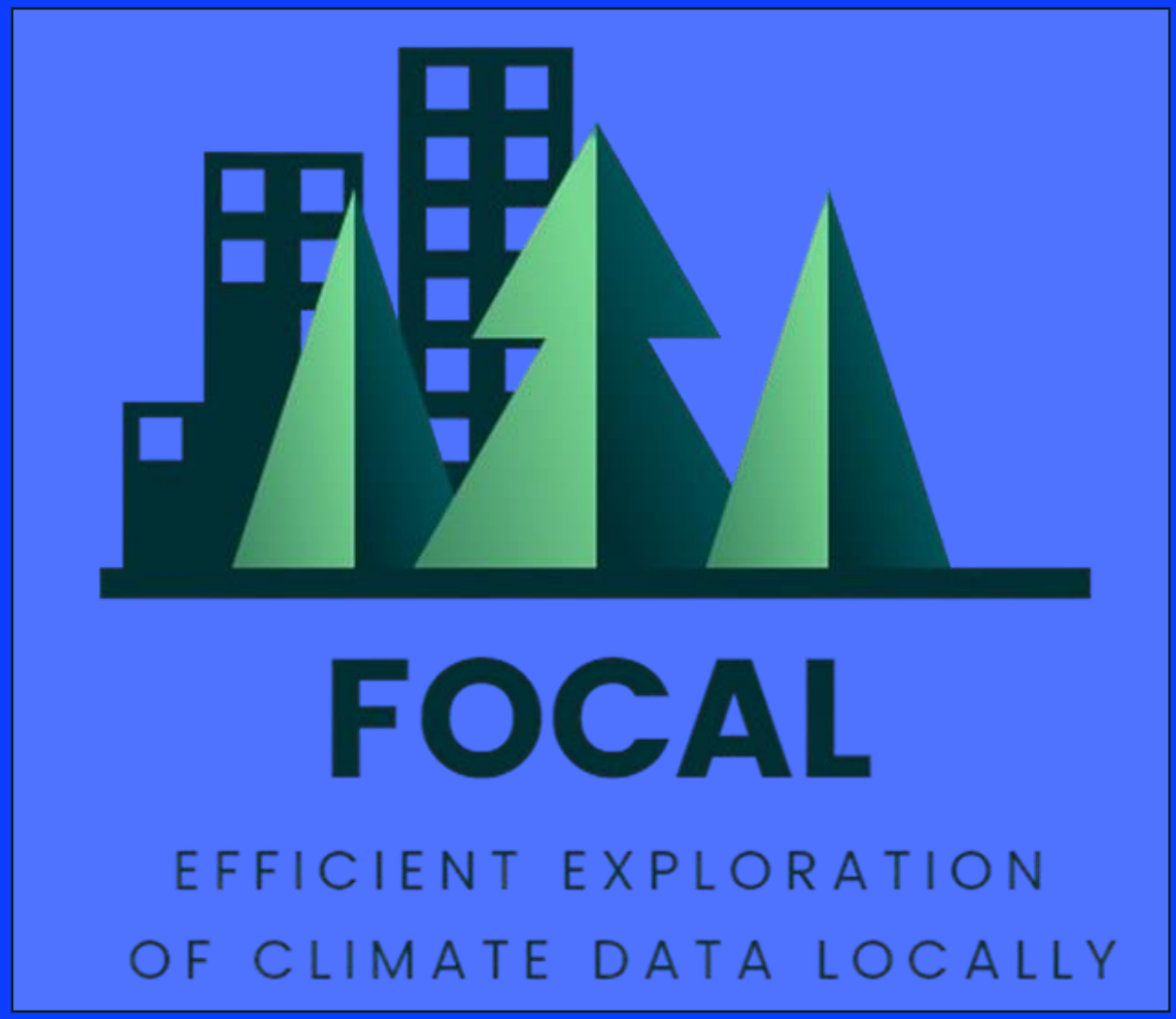
YMS
TURN ON TECHNOLOGY

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

FOCAL is dedicated to **enhancing our understanding of climate systems** and developing **innovative, AI-powered solutions** to address the challenges of climate change. This initiative brings together **top-tier experts** in AI, climate modeling, GIS, big data, HPC and cloud computing, while focusing on five core objectives:

https://www.focal-euproject.eu



ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

HOSTED BY COSMOTE

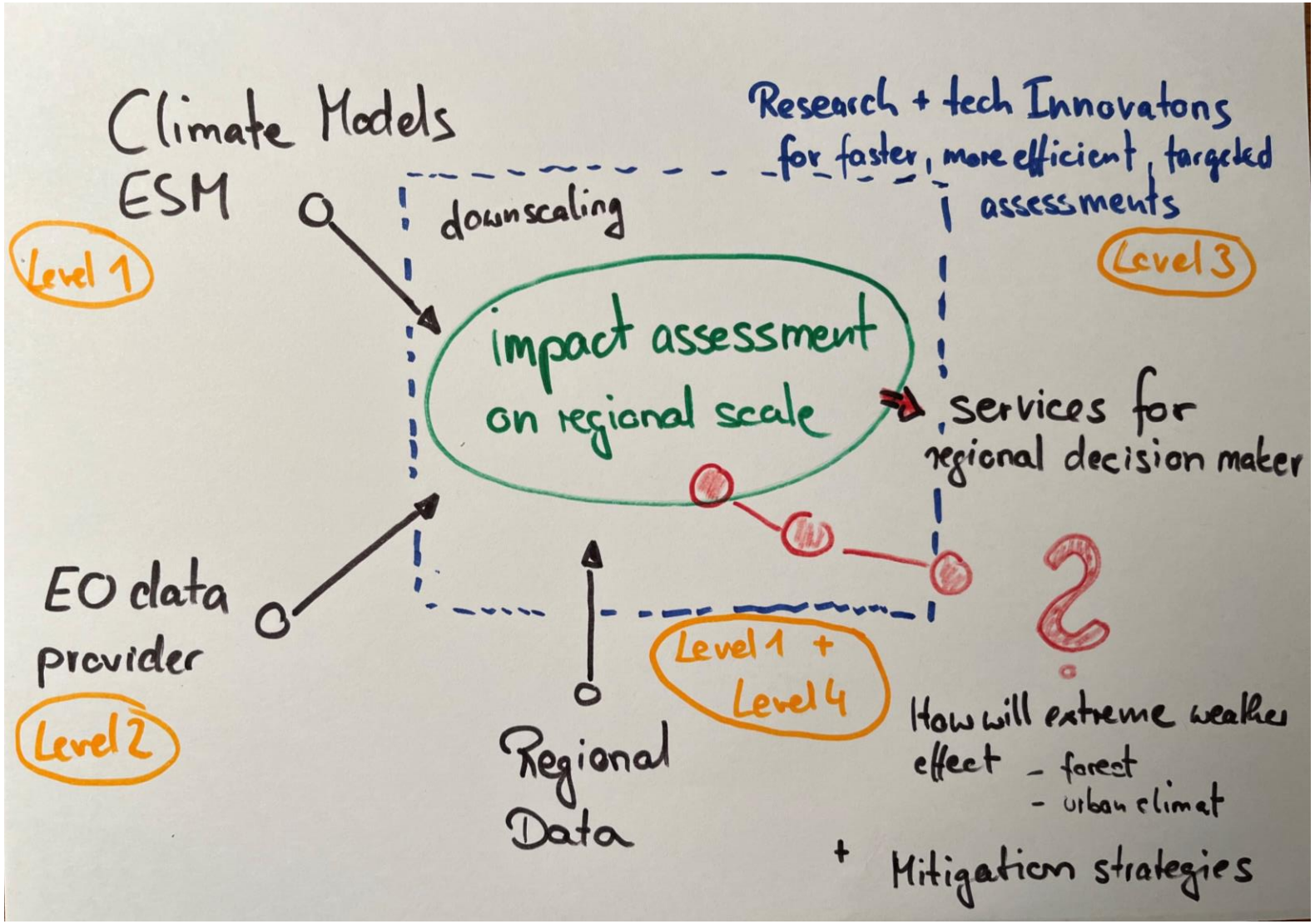
IN COLLABORATION WITH ATC, CORE|IC, FORTH, ERMIONI ICCS, DEMOKRITOS, NEUROPUBLIC

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

New data brings a lot of answers and in same time opens new challnages

HOW TO USE CLIMATIC MODELS TO THE REGIONAL / LOCAL SCALE



DATAWEEK
JOIN.LEARN.SHARE.GET VALUE
ORGANISED BY BDV BIG DATA VALUE ASSOCIATION
HOSTED BY COSMOTE
IN COLLABORATION WITH ATC ICCS core|IC FORTH DEMOKRITOS NEUROPUBLIC

Pilots

Urban Forest

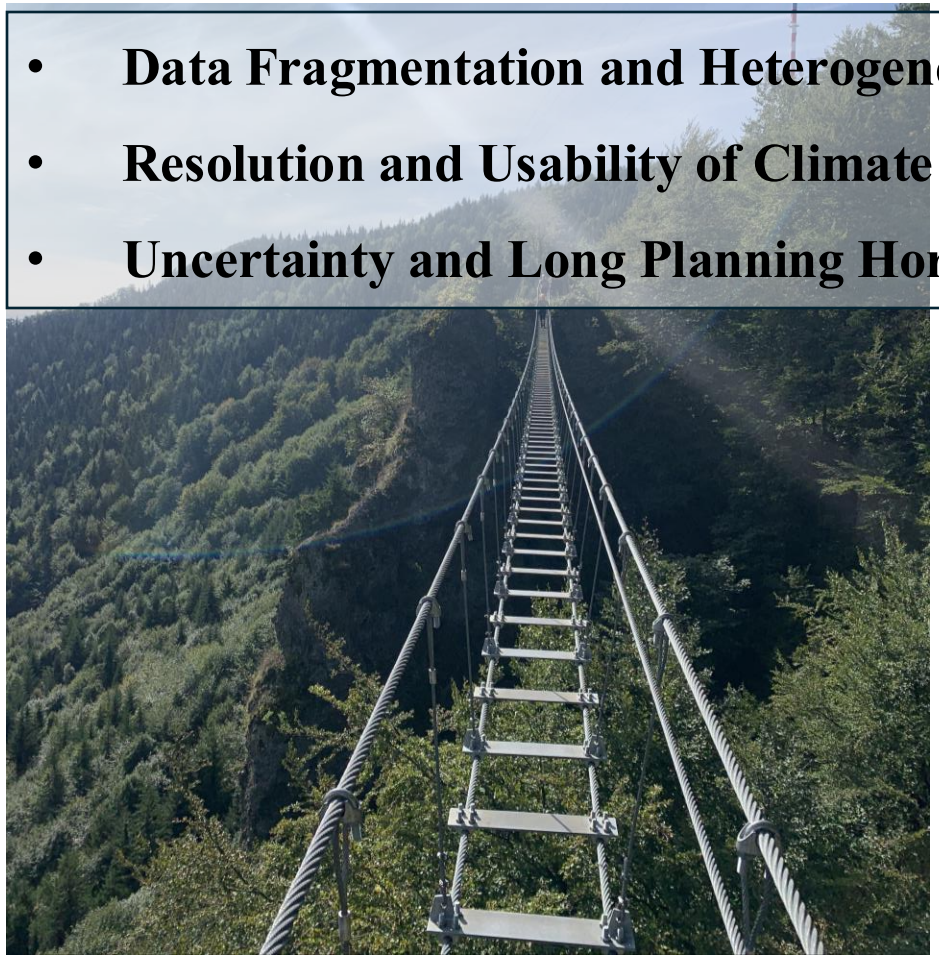
Forest pilot

support climate-resilient forest management by developing scalable, data-driven adaptation and mitigation tools. These tools are aimed at national agencies, forest managers, and local stakeholders, providing support for evidence-based decisions across both public and private forest lands.

These objectives will be implemented through co-designed workflows, supported by artificial intelligence (AI), machine learning techniques, and modular decision-support tools.

Challenges

- **Data Fragmentation and Heterogeneity**
- **Resolution and Usability of Climate Projections**
- **Uncertainty and Long Planning Horizons**



Solutions

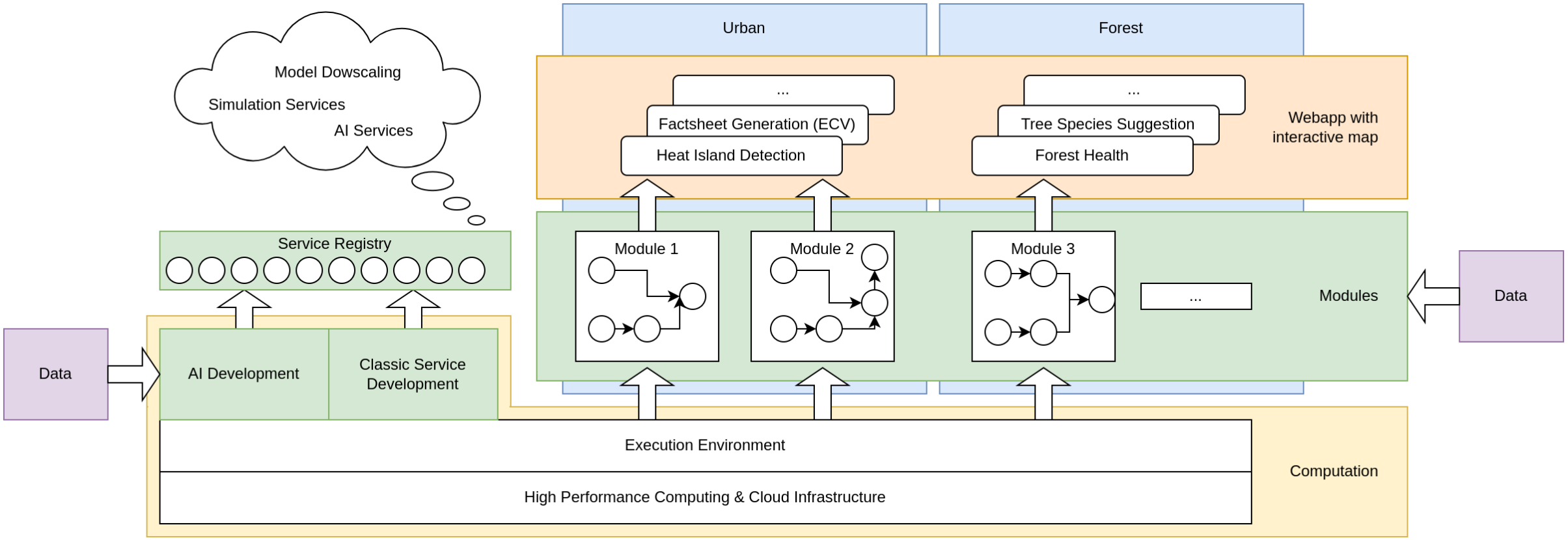
- **Data harmonization** across climate, soil, vegetation, and pest datasets
- **AI-based modeling** to increase predictive accuracy and automate risk detection
- **Modular platform design** enabling adaptation to different user needs and data availability
- **Participatory engagement** to ensure outputs are usable, transparent, and relevant to forest practitioners

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE
ORGANISED BY BDV BIG DATA VALUE ASSOCIATION
HOSTED BY COSMOTE
IN COLLABORATION WITH ATC CORE|IC FORTH DEMOKRITOS NEUROPUBLIC

Forest pilot

- Suitability of Tree Species
- Heat Stress in Forest Stands
- Prediction of Threatened Stands
- Water Logging in Forest Soils
- Updating Climate Area Maps
- Alerting Forest Owners to Stand Changes
- AI for Forest Land Management
- Modeling Long-Term Soil Property Changes
- Timber Supply Forecasting
- Modeling Timing of Forestry Operations
- Verifying the Basis of Forestry Planning

Focal project architecture



Innovative AI Technologies to Be Delivered Through the Project

Downscaling – super-resolution

- Climatic projections
- EO data

Tree grow modelling

- Input:
  - EURO-CORDEX downscaled to 100 m
  - for different **representative concentration pathways** (RCPs)
- True output / target values:
  - historical data on tree growth (time series?)
- Baseline method:
  - standard analytical model of tree growth.
- AI model of tree growth.

Thank you very much

Radovan Hilbert

https://www.focal-euproject.eu
https://yms.sk/en/homepage/

neuralio
ARTIFICIAL INTELLIGENCE

Stelios Kotsopoulos

CEO of Neuralio A.I.

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Multiagent Solution for Agriculture: EO-Informed Multi-level ABM for DT Simulations (ESA TRANSITION)

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

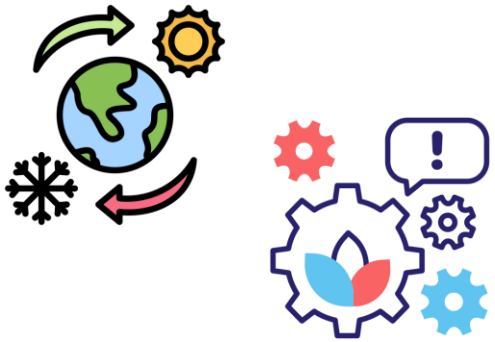
HOSTED BY COSMOTE

IN COLLABORATION WITH ATC, CORE|IC, FORTH, DEMOKRITOS, NEUROPUBLIC

The challenge in sustainable agriculture

Global agriculture faces escalating pressures:

- Climate change impacts: Extreme weather, shifting land suitability.
- Resource scarcity: Water, arable land.



Need for transformative, data-driven approaches beyond incremental improvements.

Limitations of traditional planning:

- Difficulty merging large-scale EO data with economic/social factors.
- Inability to simulate dynamic feedback loops in complex systems.

Goal: **Advanced Decision Support for sustainability and food security.**

ESA TRANSITION project: A Digital Twin for agriculture

Objective | Develop a sophisticated Decision Support System (DSS) for climate-adaptive land-use transitions.

Core Methodology | Earth Observation (EO) data + Multi-Level Agent-Based Models (MABM) → Digital Twin (DT) simulations for agriculture.

Enables | Dynamic simulation and proactive strategy testing, rather than reactive responses.

transition
EO-INFORMED AGENT BASED MODELS FOR DIGITAL TWINS APPLICATIONS

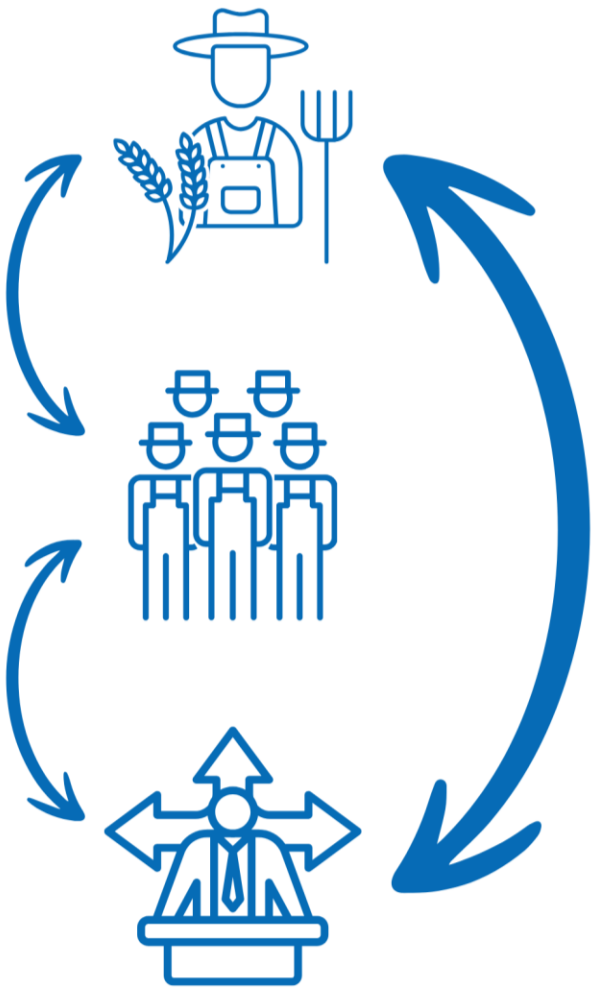
esa

Project Aims

Develop a user-centered platform for agriculture, renewable energy, and environmental sustainability.

Simulate complex socio-environmental dynamics to inform policy and sustainable land-use planning.

Core Innovation | Multi-Level Agent-Based Modeling (MABM)



MABM Definition: A computational approach simulating actions and interactions of autonomous agents (e.g., farmers, landowners, policymakers) across hierarchically structured levels.

Three Interaction Levels

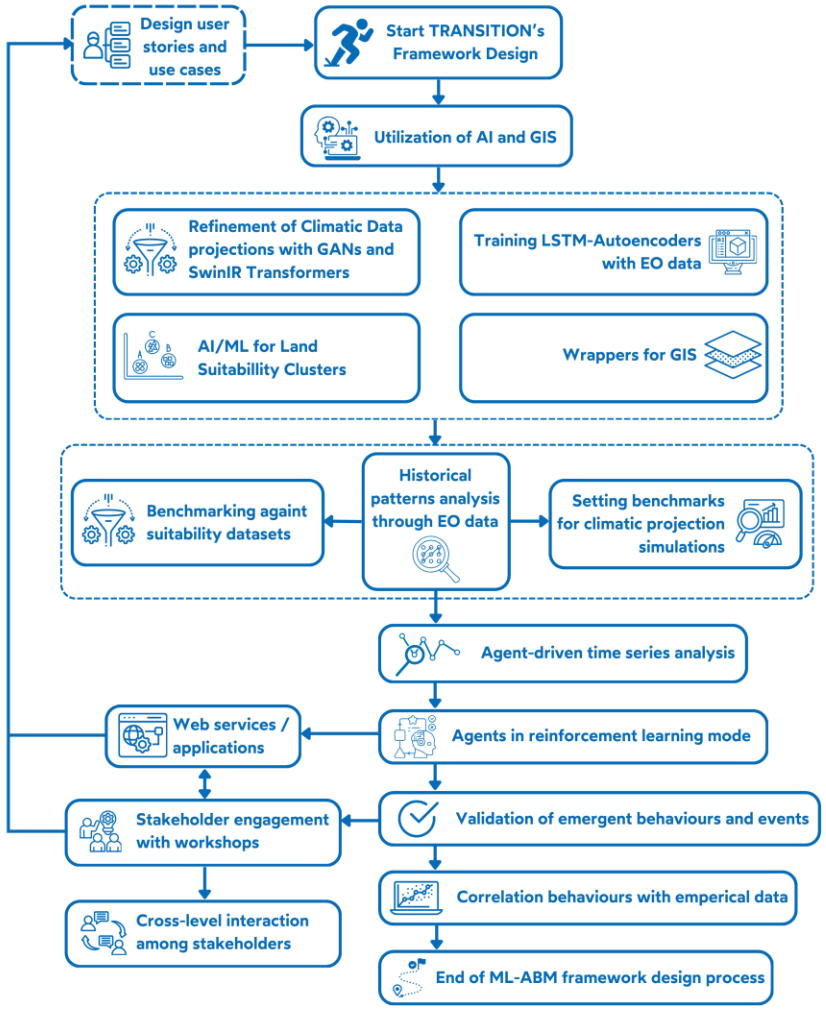
Micro-level: Individual farmers/landowners making land-use decisions (e.g., crop choice, PV adoption) based on local conditions, profit motives, and behavioral traits.

Meso-level: Regional dynamics, including agricultural cooperatives, community interactions, and the aggregation of individual behaviors into broader patterns.

Macro-level: Systemic factors and policy interventions, such as subsidies, environmental regulations, climate goals, and overarching economic conditions.

PECS Framework Integration: Incorporates Physical, Emotional, Cognitive, and Social (PECS) attributes for enhanced behavioral realism. Agents consider factors like stress levels, social influence, and policy awareness, moving beyond purely economic motivations.

Core Innovation | Earth Observation (EO) Driven Insights



Role of EO Data | Grounds the MABM simulations in real-world, dynamic environmental context, crucial for the fidelity of the Digital Twin.

Key EO Data Sources Utilized

Sentinel Satellites (Sentinel-1, -2, -3): Provide data for land cover classification, vegetation health monitoring, Land Surface Temperature (LST), surface roughness and soil moisture content.

Climate Reanalysis: Offers consistent historical climate conditions for model calibration and trend understanding.

Climate Projections (e.g., RCP 2.6, 4.5, 8.5): Used to simulate future "what-if" climate scenarios for impact assessment.

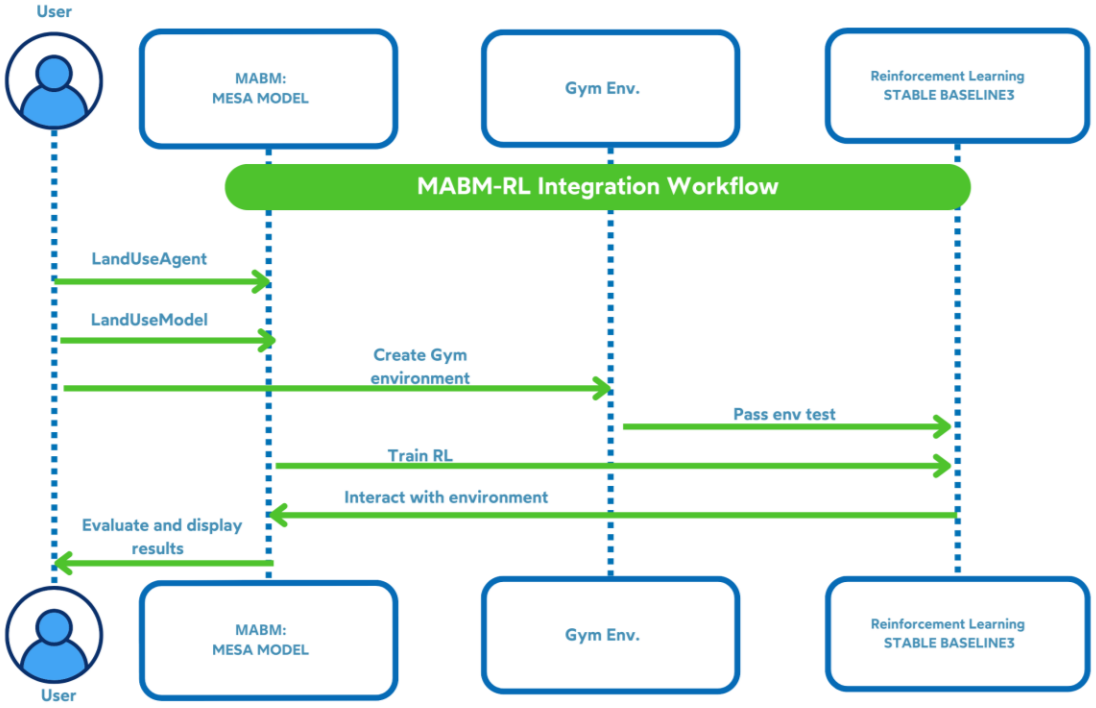
Geospatial Data: Copernicus DEM, SOILGRIDS, CORINE land cover inventory for topography, soil properties, and existing land use.

Informs Key Agricultural Parameters

Dynamic Crop Suitability: Biophysical simulation tools like **AquaCrop** estimate crop yields, irrigation needs, and soil fertility impacts under various climate scenarios.

Photovoltaic (PV) Suitability on Farmlands: Assesses optimal PV locations considering solar radiation, land slope, ambient temperature, infrastructure proximity, and land cover.

Core Innovation | Adaptive Decisions via RL



Role of RL: Enables agents (e.g., farmers) to learn and adapt their land-use strategies dynamically over time in response to changing conditions.

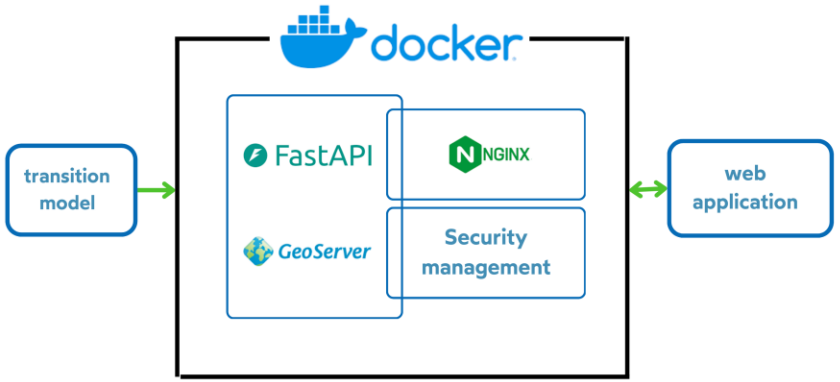
Algorithm Employed: Proximal Policy Optimization (PPO), a state-of-the-art actor-critic RL algorithm chosen for its stability and proficiency in handling complex decision spaces (both discrete and continuous actions).

Key RL Components in Agricultural Context:

Observation Space: Defines what an agent perceives (e.g., current crop/PV suitability scores, its own PECS attributes like stress level, prevailing policy incentives, climate-related loss factors).
Action Space: Delineates the set of actions an agent can take (e.g., discrete choice: continue farming vs. install PV; continuous: amount of land to allocate or investment level).
Reward Function: Critical for guiding the learning process; incentivizes or penalizes actions. Rewards can be based on economic profit (farming/PV), psycho-social factors (satisfaction from PECS), achievement of policy targets, or stress minimization. The structure varies by simulation scenario (Base, PECS, Full).

Technical backbone and Simulation scenarios

Service-Oriented Technical Architecture | Designed for modularity, scalability, and efficient data management.



Core Simulation Libraries

MESA: Foundational framework for developing Agent-Based Models.

Gym (OpenAI Gym): Standardized interface for Reinforcement Learning environments.

Stable Baselines3: Provides reliable implementations of RL algorithms like PPO.

Key Simulation Scenarios: **Base** (economic returns), **PECS-Enhanced** (behavioral factors),

Full Scenario (Green Credit Policy integration), **Climate Change Adaptation** (climate-driven losses and resilience targets).

Validation and Explainability (XAI)

Model Validation for Reliability & Credibility

Ensemble Modeling: Running simulations multiple times with different random seeds to assess decision variability, model robustness to initial conditions, and quantify uncertainty in outcomes.

Sensitivity Analysis: Assessing how variations in input parameters (e.g., crop prices, PV costs, policy levels) affect simulation outcomes to identify critical leverage points.

(**Future Goal**: Comparison with empirical data for benchmarking).

Explainable AI (XAI) for Transparency

SHAP (SHapley Additive exPlanations) Analysis: A game theory-based approach to quantify the importance of each feature in an agent's decision to switch land use.

Decision Plots: Visualizations illustrating aggregated decision trends over time.

Impact and future outlook

Potential Impact of TRANSITION DSS

Informing sustainable land-use planning and agricultural policy design.

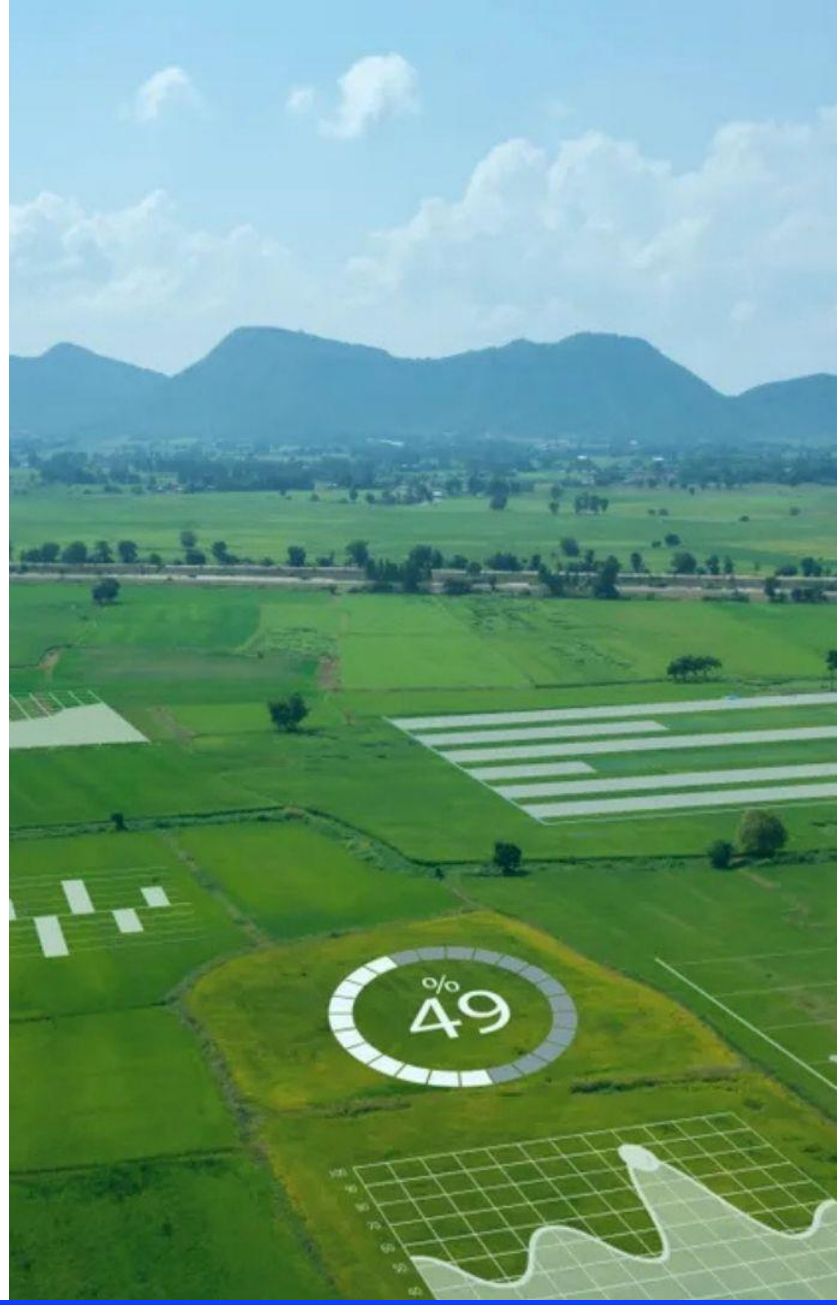
Optimizing climate adaptation strategies for agricultural sector resilience.

Facilitating stakeholder engagement and collaborative decision-making.

Key Future Directions

Enhanced integration with broader European platforms like Destination Earth (DestinE) for operationalization and access to higher-resolution data.

Continued iterative refinement of model components, simulation scenarios, and user interfaces based on ongoing stakeholder feedback and evolving needs.



Thank you very much

Stelios Kotsopoulos, CEO of Neuralio A.I.

https://transition.neuralio.ai

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

Upscaling agricultural data from farm level to regional datasets (ScaleAgData)

Dr. Odysseas Sekkas
Head of Technical Project Management
Neuropublic S.A.

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION
HOSTED BY COSMOTE
IN COLLABORATION WITH ATC, CORE|IC, FORTH, DEMOKRITOS, NEUROPUBLIC

Water productivity

Crop management

Yield monitoring

Soil health

Grasslands

Dairy

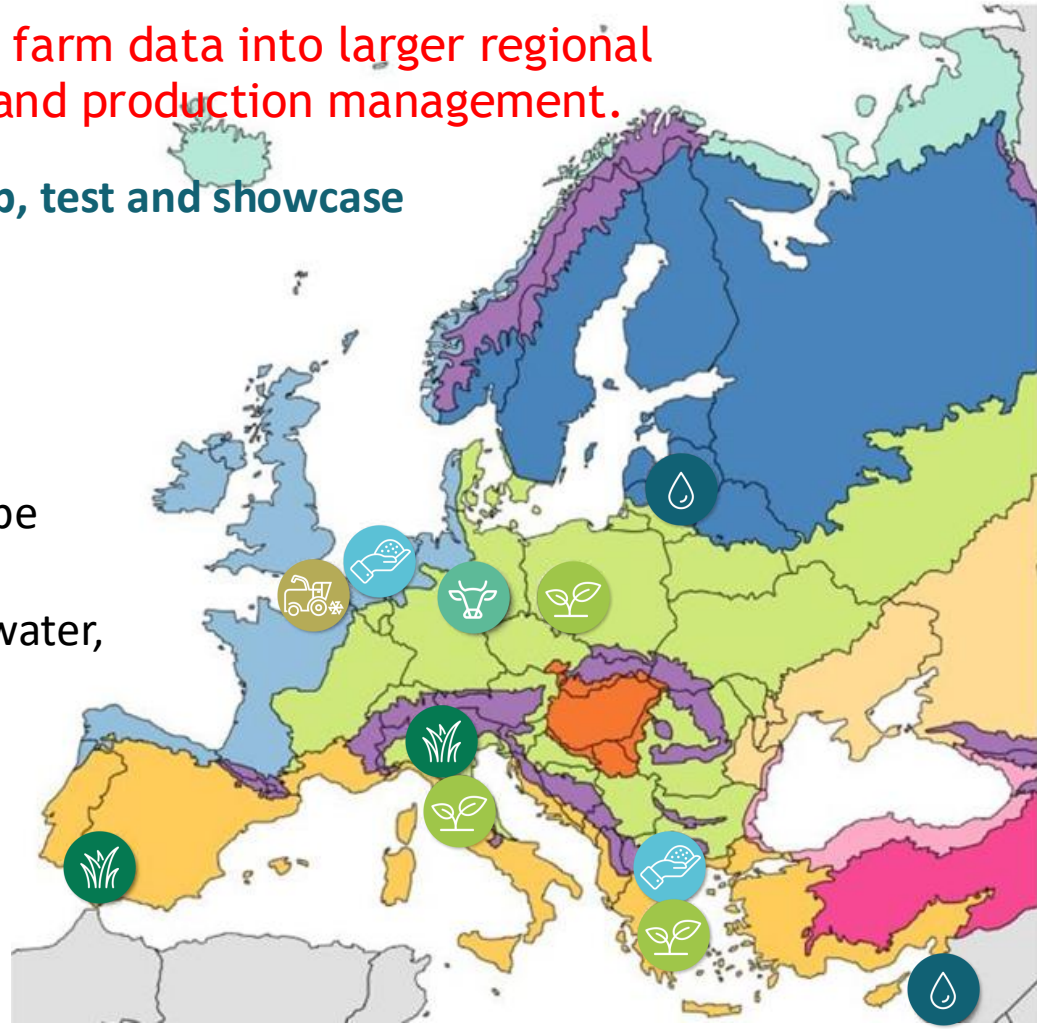
ScaleAgData aims to understand and refine how complex farm-level data streams are governed, organized and managed.

It also focuses on technologies to scale local farm data into larger regional datasets for agri-environmental monitoring and production management.

Six (6) Research and Innovation Labs will **develop, test and showcase** how integrating near real-time (NRT) sensor data can provide improved monitoring capabilities.

These RILabs will **drive innovation**, covering

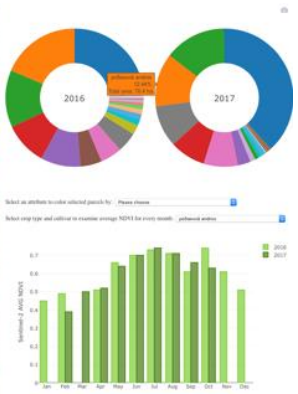
- different biogeographical zones of Europe
- the 4 dimensions of environment: soil, water, air, living organisms (crops & livestock)
- available sensor data & products



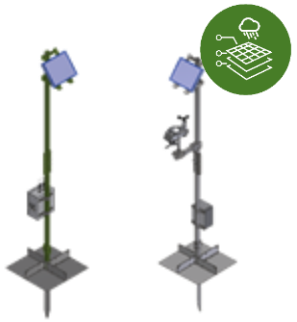
Crop management RIL goals

- To create a mechanism and a governance plan for collecting IoT and farm log data at the **parcel level** and aggregating them at a **regional level**
- User-friendly navigation from municipal overviews down to individual parcels
- Role-based insights for **farmers**, **agronomists** and **policymakers**
- Aggregated metrics for **CAP performance** (Common Agricultural Policy)

Local data → Regional performance

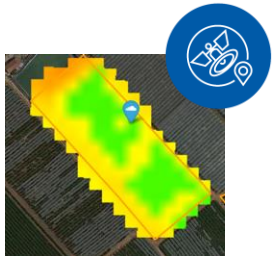


Data sources and methodology



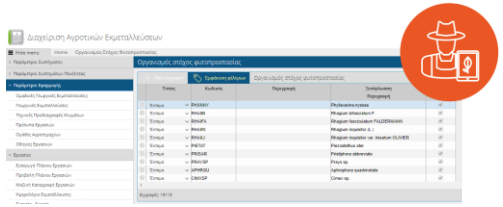
Atmospheric stations

Soil stations



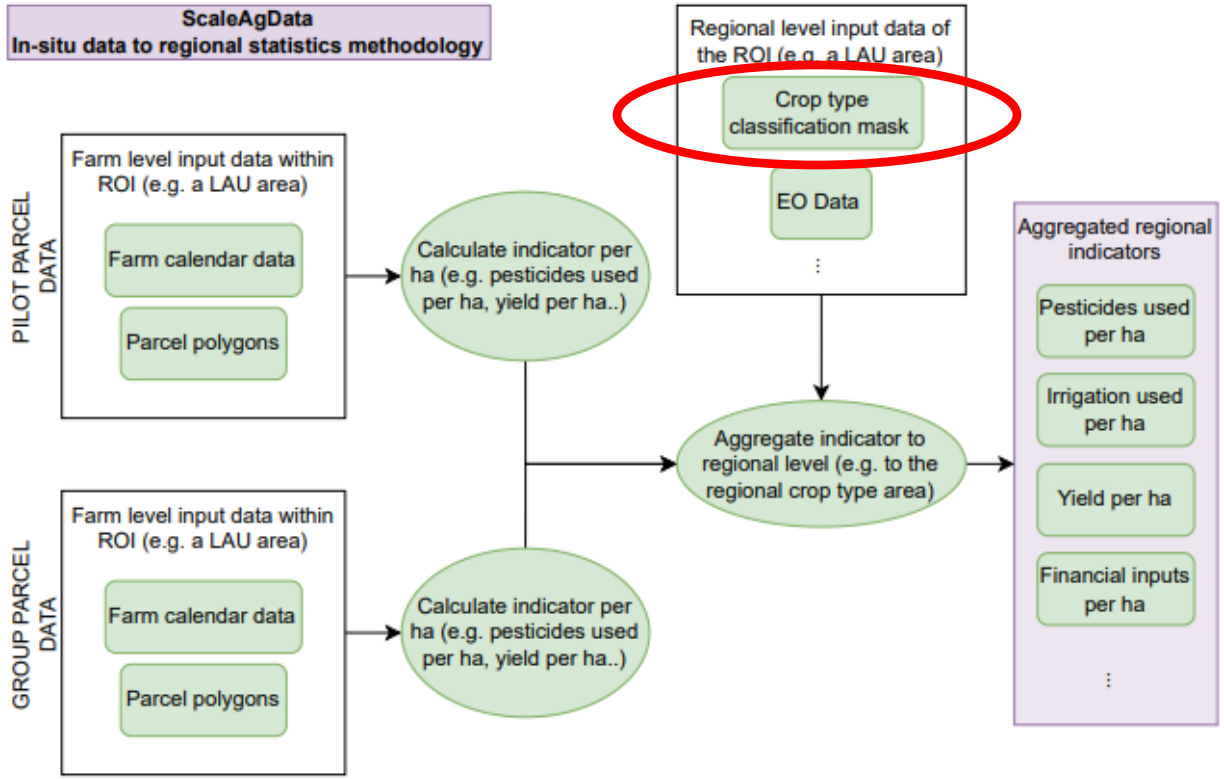
Satellite data

Advisor & farmer field visits



Farm calendars from FMIS

External data sources



ML Classification algorithm

Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel

To train the classifier and produce the crop classification we used the following data:

Parcel data

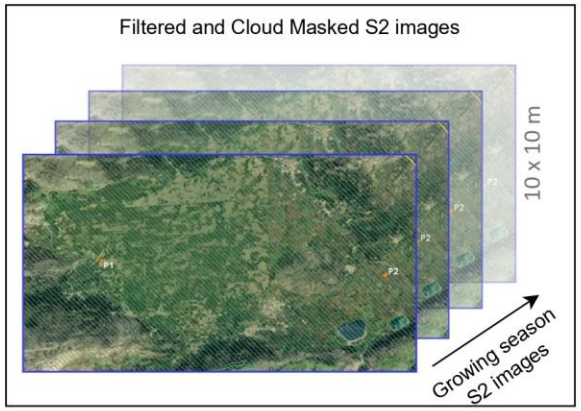
- Polygon geometries
- Crop type
- Growing season start
- Growing season end

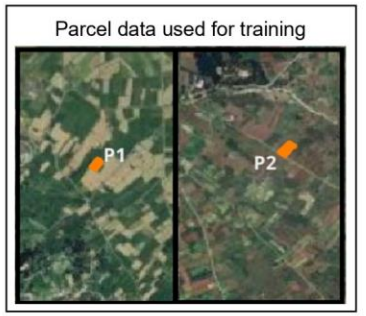
EO data

- Filtered Cloud masked Sentinel-2 images over the growing period
- Selected bands: 'B2', 'B3', 'B4', 'B5', ‘B6’

Other data

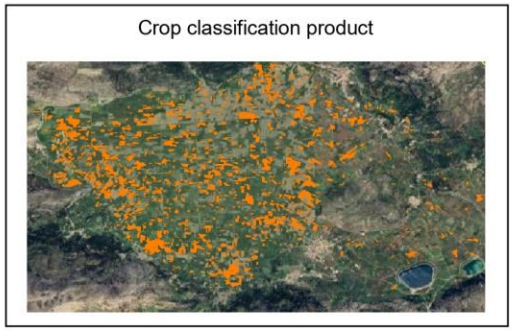
- Local Administrative Unit (LAU) regional boundary
- Land use data (for identifying water, bare land, forest, urban area etc.)





Crop classification engine

- **Classifier Type:** Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel.
- **Parameters:** Gamma set to 0.6 and cost set to 10.
- **Training:** Use the sampled training data and selected bands to train the classifier.



L1

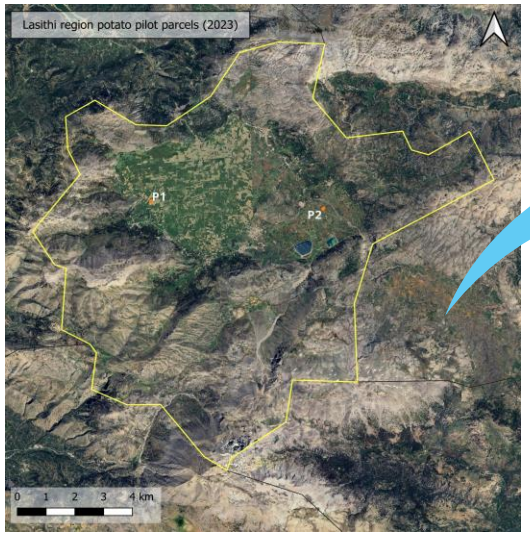
Pilot parcels

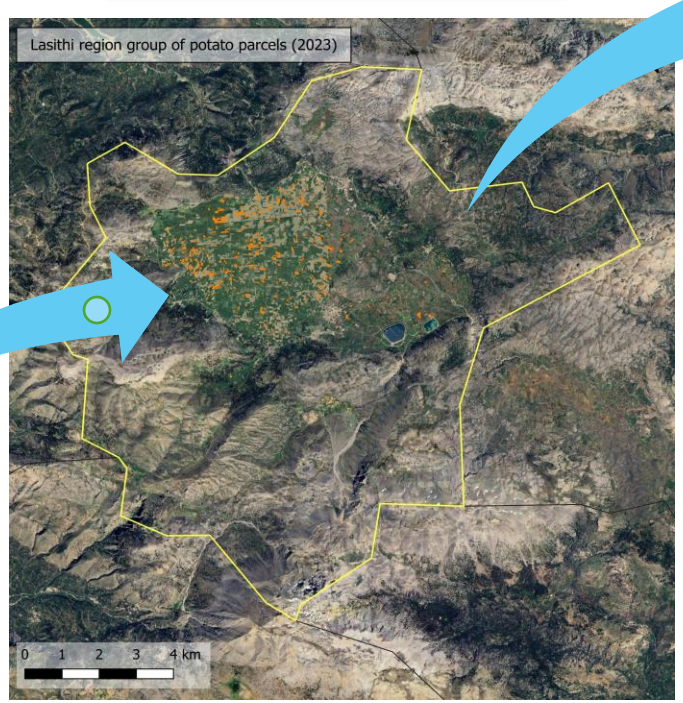
L2

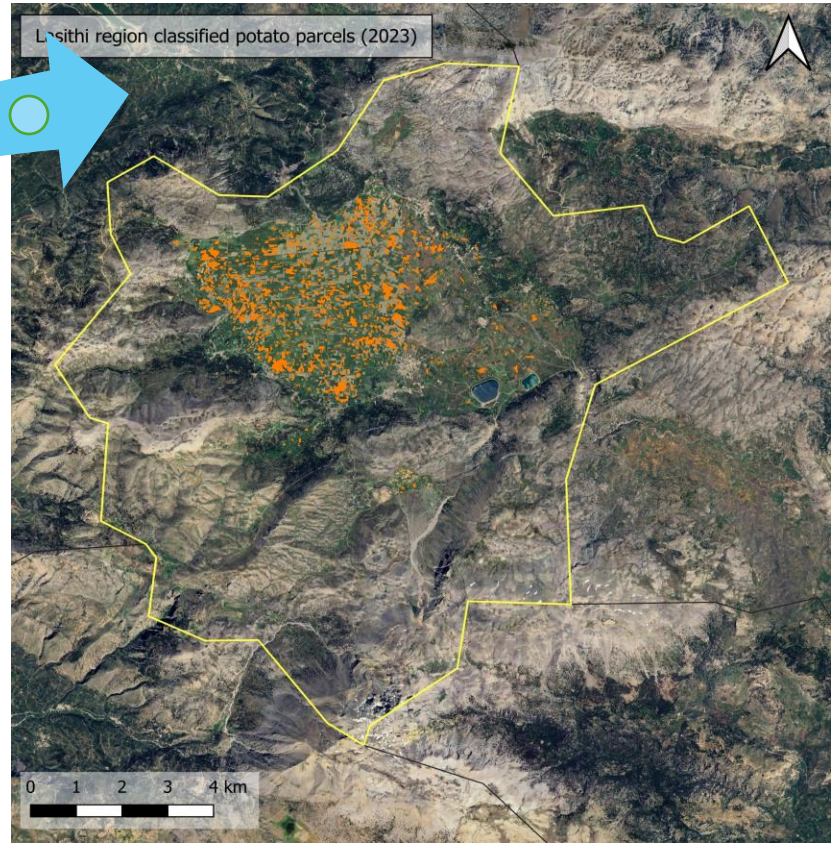
Group of parcels with crop of interest (can be used as control parcels)

L3

Areas within the administrative region where the crop of interest is cultivated







Considerations and Future work

- Accuracy may differ depending on the area size and croptype
- Train on multiple croptypes within the area of interest fore better results
- Include in the workflow the phenological stages (leaf development, flowering, etc.) information of the crop
- Investigate additional AI/ML methods and extrapolation algorithms
- Validation in large regions of different agroclimatic zones

Thank you!

Discussion