ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
AGRICULTURAL UNIVERSITY OF ATHENS

Evaluation of digital maps of top-soil properties compared to large-scale laboratory soil data and synergies towards a better European soils' delineation

Stergia Palli Gravani, Stylianos Gerontidis, Dimitrios Kopanelis, Orestis Kairis, Konstantinos Soulis, and Dionissios Kalivas

Agricultural University of Athens, Department of Natural Resources Development & Agricultural Engineering, Greece

Geographical Information Systems Research Unit | AGRICULTURAL UNIVERSITY OF ATHENS

This project is carried out within the framework of the National Recovery and Resilience Plan Greece 2.0, funded by the European Union – NextGenerationEU (Implementation body: HFRI), https://greece20.gov.gr

Greece 2.0 NATIONAL RECOVERY AND RESILIENCE PLAN

Funded by the European Union NextGenerationEU

H.F.R.I. Hellenic Foundation for Research & Innovation

Introduction

European soil maps are compared with Greek soil data to evaluate their accuracy across key properties like texture, organic carbon, pH and Cation-exchange capacity (CEC) using spatial and statistical comparisons.

In this presentation we focus on soil texture

Objective :

- Evaluate the accuracy and representativeness of European soil datasets, incorporating statistical methods to optimize analysis and draw useful conclusions.
- Explore spatial correlation and identify spatial discrepancies and their causes.
- Highlight the value of integrating national soil data for improved continental-scale mapping.
- Offer recommendations for future global soil data initiatives.

Soil mapping techniques

Comparison of Traditional vs Predictive Soil Mapping

Criteria	Traditional Soil Mapping	Predictive Soil Mapping
Data Collection	Relies on field surveys and soil sampling, often resulting in limited data points	Utilizes extensive datasets, including remote sensing and geospatial data, allowing for broader coverage.
Accuracy and Reliability	Subject to human error and biases, with accuracy dependent on the expertise of the surveyor	Generally, more accurate but can suffer from data limitations and mathematical errors.
Applications	Best suited for localized studies or regions with extensive field data	Ideal for large-scale assessments and cross-boundary issues, essential for regional and national planning.
Cost and Time Efficiency	More time-consuming and expensive due to the need for fieldwork	More cost-effective and faster as it leverages existing datasets and computational techniques.

Origin of evaluated international data

ISRIC (International Soil Reference and Information Centre - SoilGrids)

Global soil properties prediction system with 250 m resolution, updated in 2016.

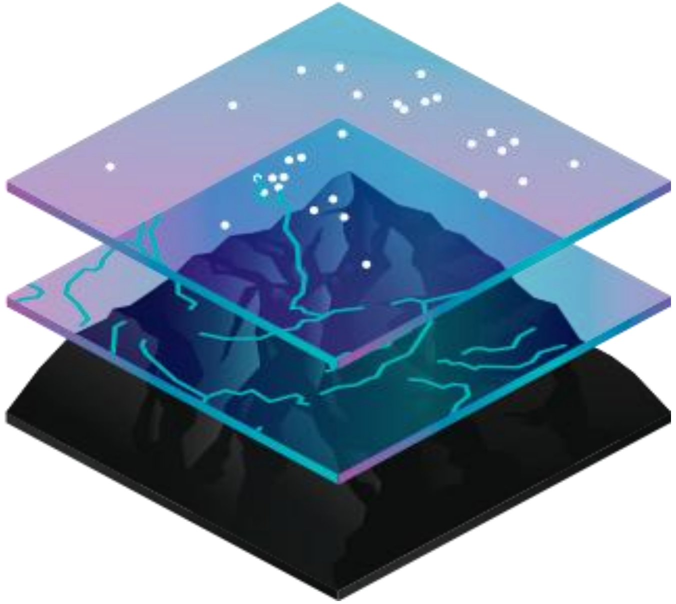
- Provides data for Organic Carbon, pH, soil texture, (6 depth levels)
- Was created based on 150,000 soil profiles, satellite data from MODIS, SRTM DEM & climate images, 158 spatial variables
- Machine learning methods: Random Forest Gradient Boosting Implemented in R
- Generated 280 raster layers, enhancing the mapping of global soil features.

ESDAC (European Soil Data Centre)

Includes soil properties raster files with a resolution of 500 m.

- Created by interpolating the LUCAS 2009 file (surface soil samples) using hybrid approaches like regression kriging.
- Regression models were fitted using, along other variables, remotely sensed data coming from the MODIS sensor.
- Provides rater layers for many soil properties such as: % of sand, silt and clay etc.

Origin of Greek observed data



Greek Soil Map (GR)

- Digital map providing detailed information about the soils of Greece, which includes Physical, chemical, mechanical properties.
- Uses & applications: good agricultural practice, land management, agricultural decision making
- Includes 10,058 sampling points

Data processing methodology

Spatial Soil Properties – ISRIC - ESDAC

- Subset multidimensional rasters to obtain the selected soil properties only for topsoil (0-30 cm).
- Target soil properties: Sand, Clay, and Silt percentages (soil texture classes).
- Resulted raster layers were clipped in the study area (Greece) and projected in the official Greek coordinates system (Greek Grid).

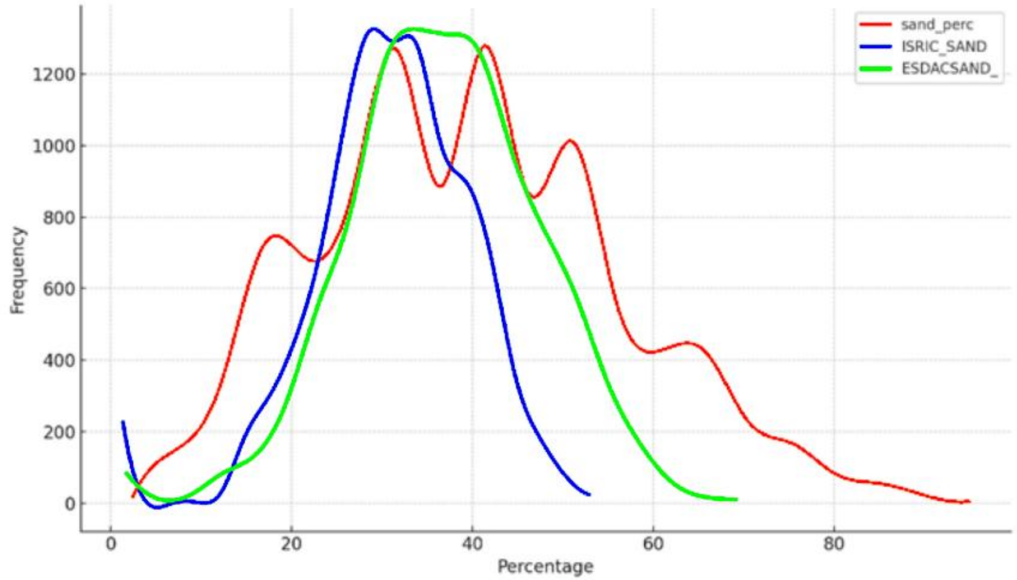
Greek Soil Map Data Processing

- Creation of a single point feature layer with the measured values for all the examined attributes.

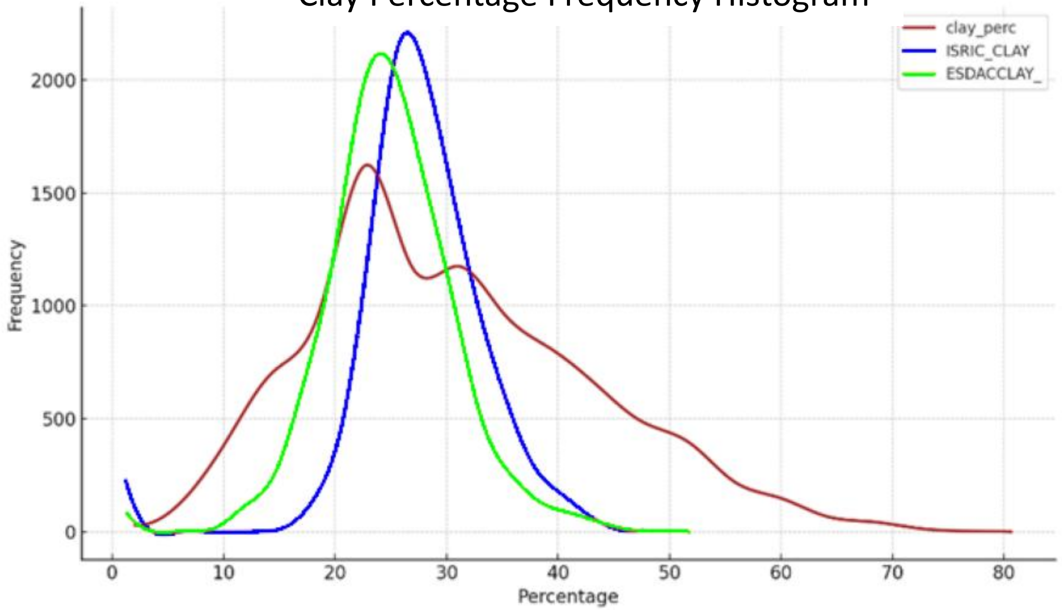
Analysis

- Sampling each soil property from the raster layer for the location of each point and joining with Greek soil map point layer
- Calculation of the differences between measured values and ISRIC and ESDAC predicted values for each soil property and for each point.
- Univariate statistics for each class (sand, clay, silt)
- Spatial Analysis of errors

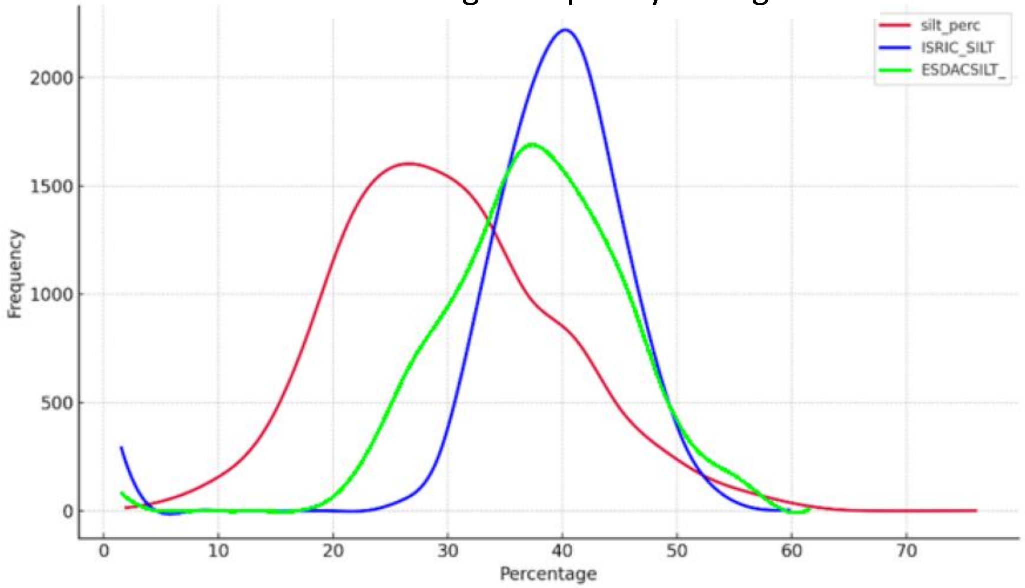
Sand Percentage Frequency Histogram



Clay Percentage Frequency Histogram

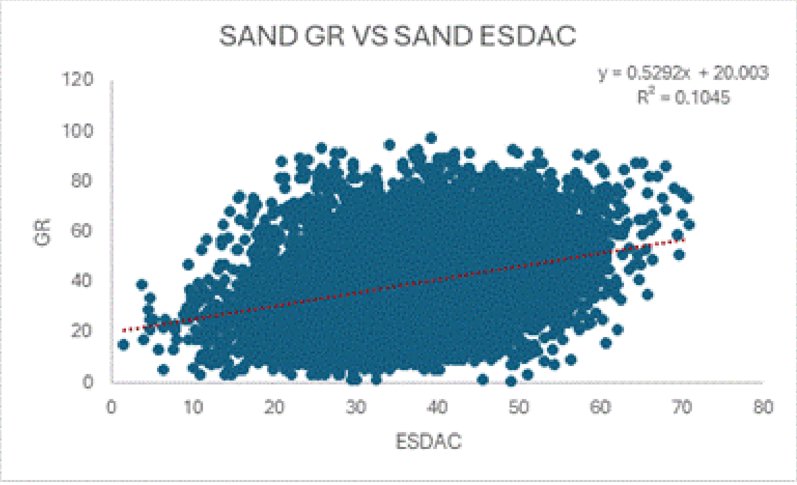


Silt Percentage Frequency Histogram

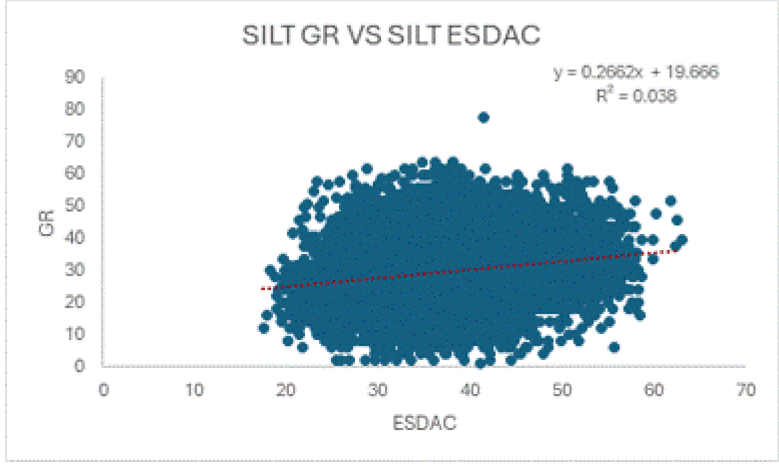


	SAND			CLAY			SILT		
	GR	ESDAC	ISRIC	GR	ESDAC	ISRIC	GR	ESDAC	ISRIC
Mean	39.69	37.06	31.76	30.60	25.08	28.26	29.70	37.86	39.98
Median	39.30	36.80	31.57	28.70	24.79	27.73	28.00	37.86	39.93

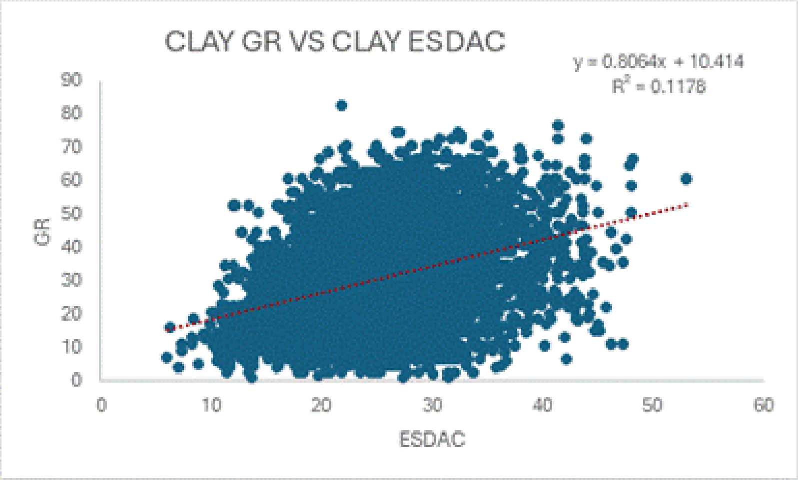
Relationship between the ESDAC dataset and the Greek Soil Map (GR)



$R^2 = 0.1045$
Slope = 0.5292



$R^2 = 0.038$
Slope = 0.2662

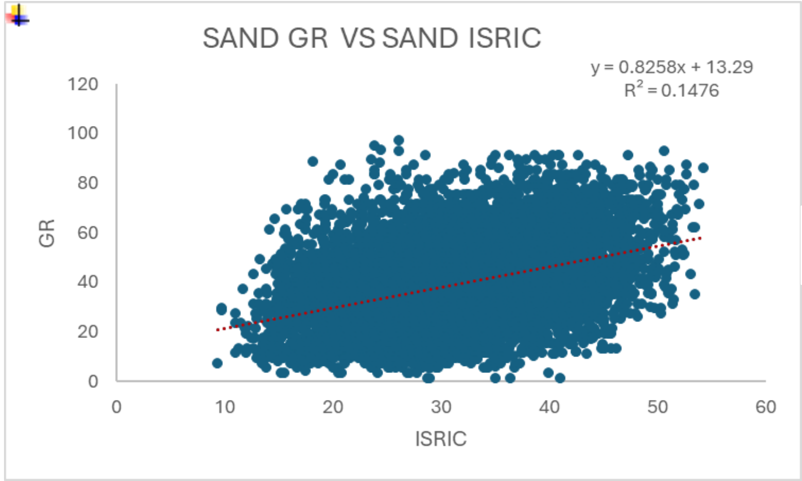


$R^2 = 0.1178$
Slope = 0.8064

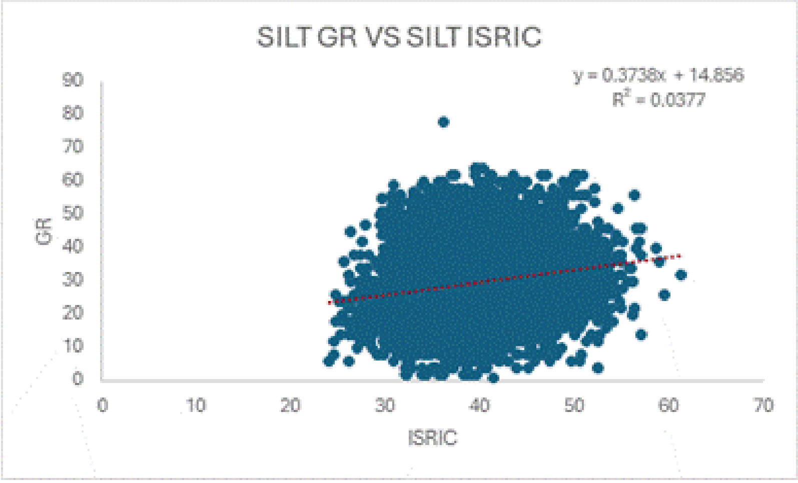
R^2 values very low → weak correlation between observed and measured data

Low slope values → overestimation in all properties

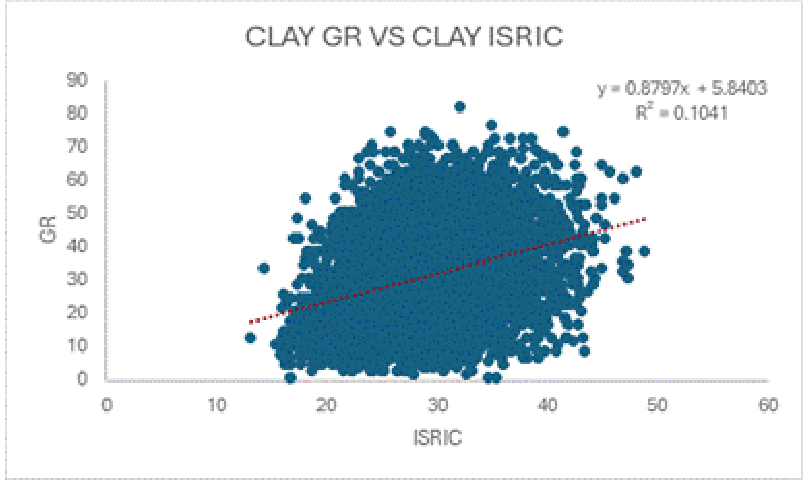
Relationship between the ISRIC dataset and the Greek Soil Map (GR)



$R^2 = 0.1476$
Slope = 0.8258



$R^2 = 0.0377$
Slope = 0.3738



$R^2 = 0.1041$
Slope = 0.8797

Coefficients of determination are relatively low, implying a weak correlation and substantial data variability

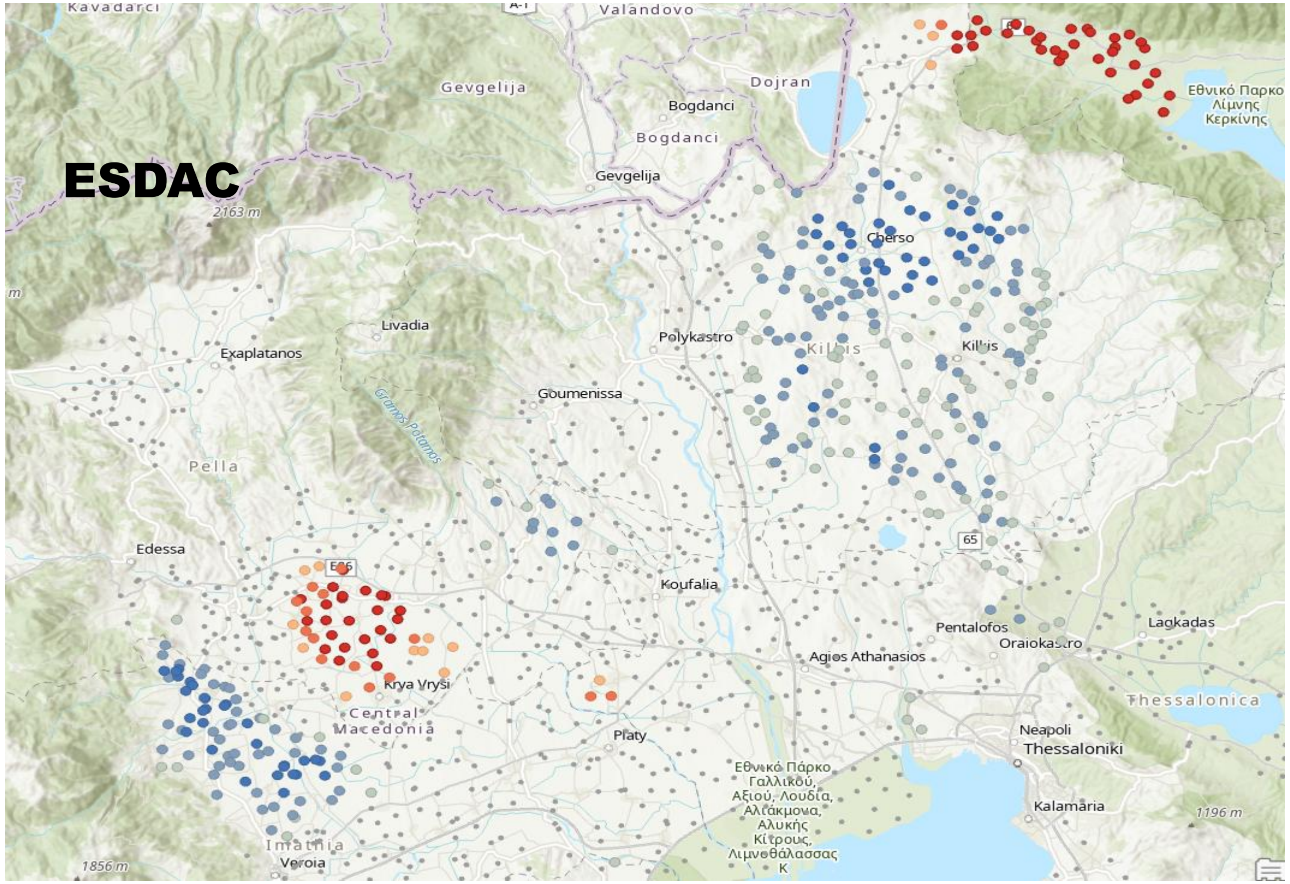
RMSE values were calculated for each soil property

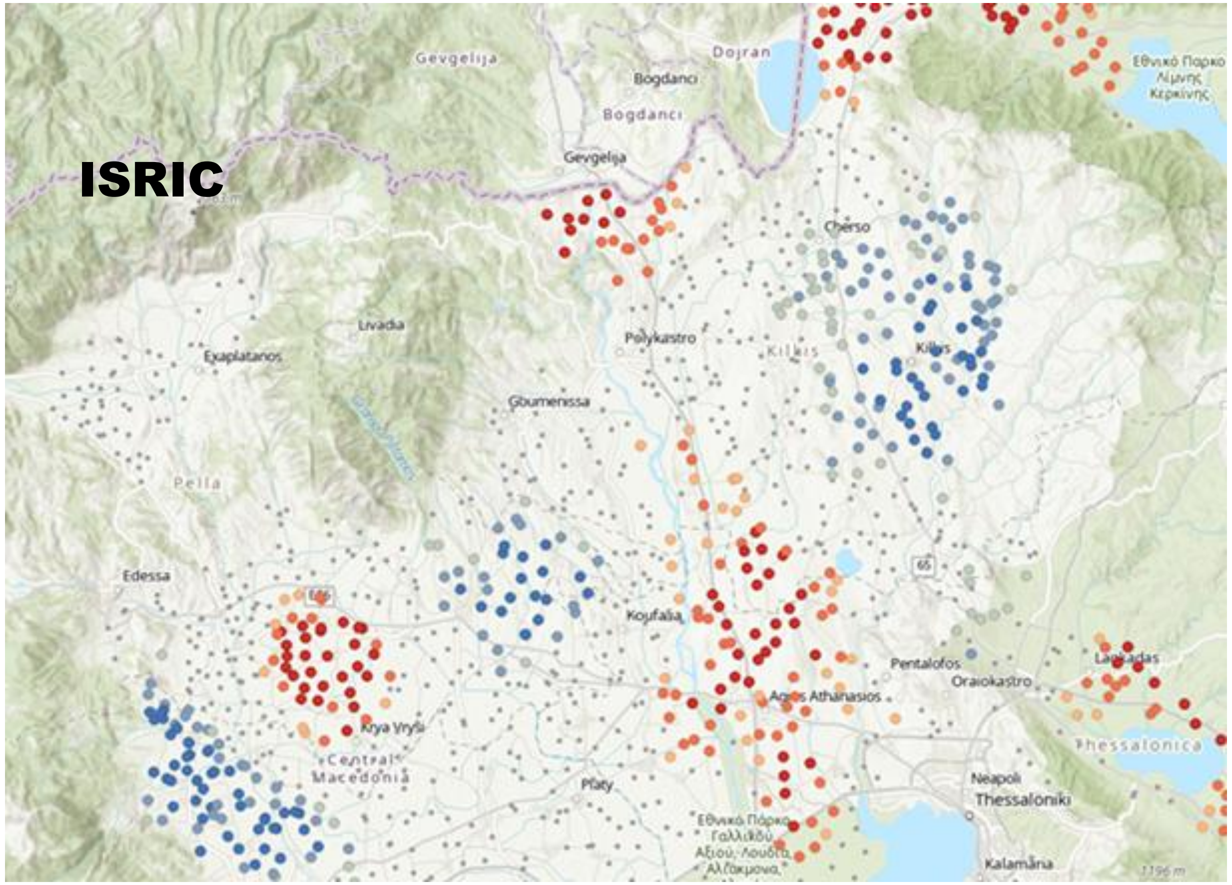
High RMSE values indicate substantial errors

	RMSE ISRIC	RMSE ESDAC
SILT	14.45	14.04
SAND	18.55	17.18
CLAY	13.20	13.69

Spatial distribution of the clustering of absolute errors

Hot Spot Analysis (Characteristic example) **SAND**



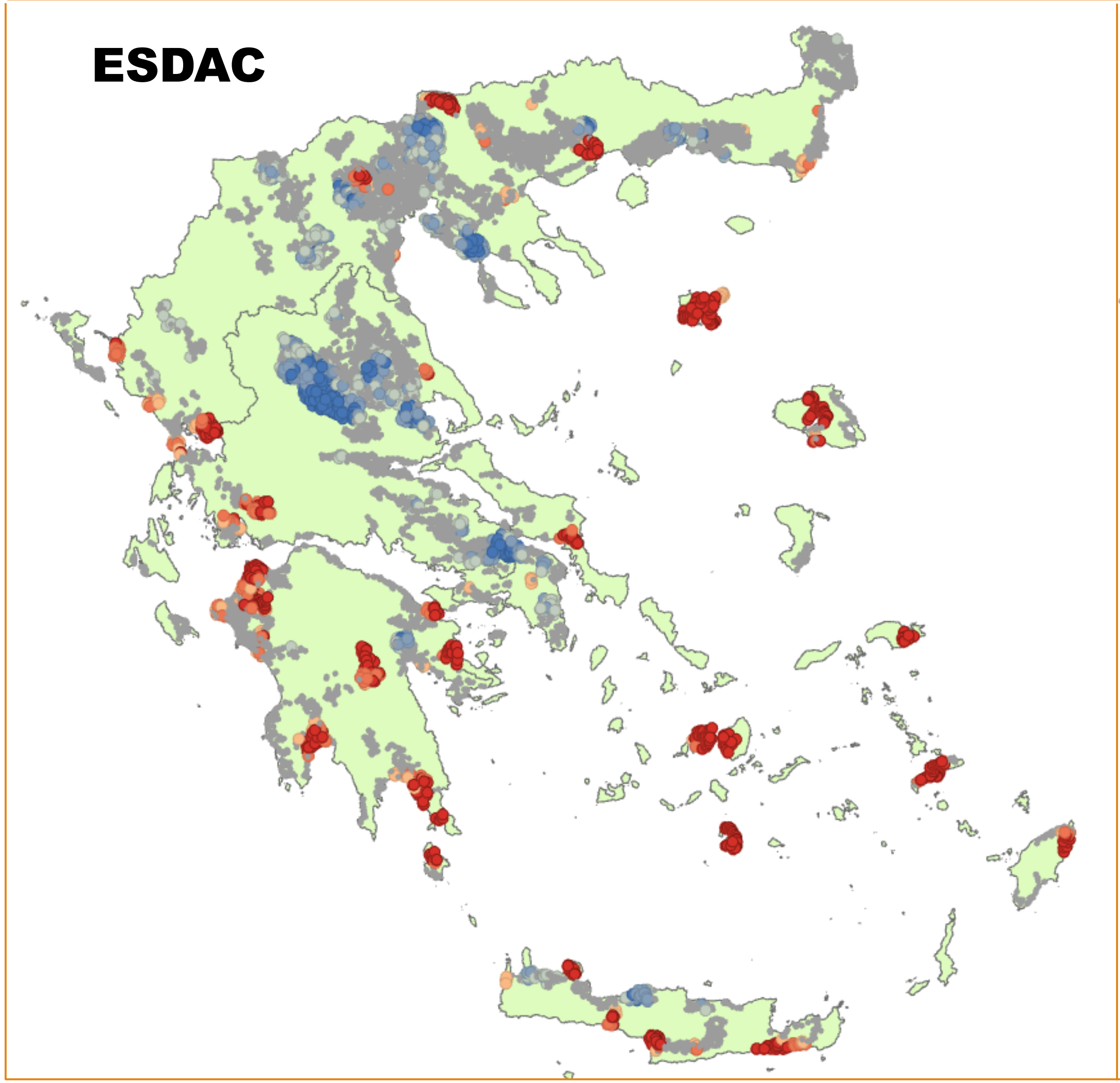


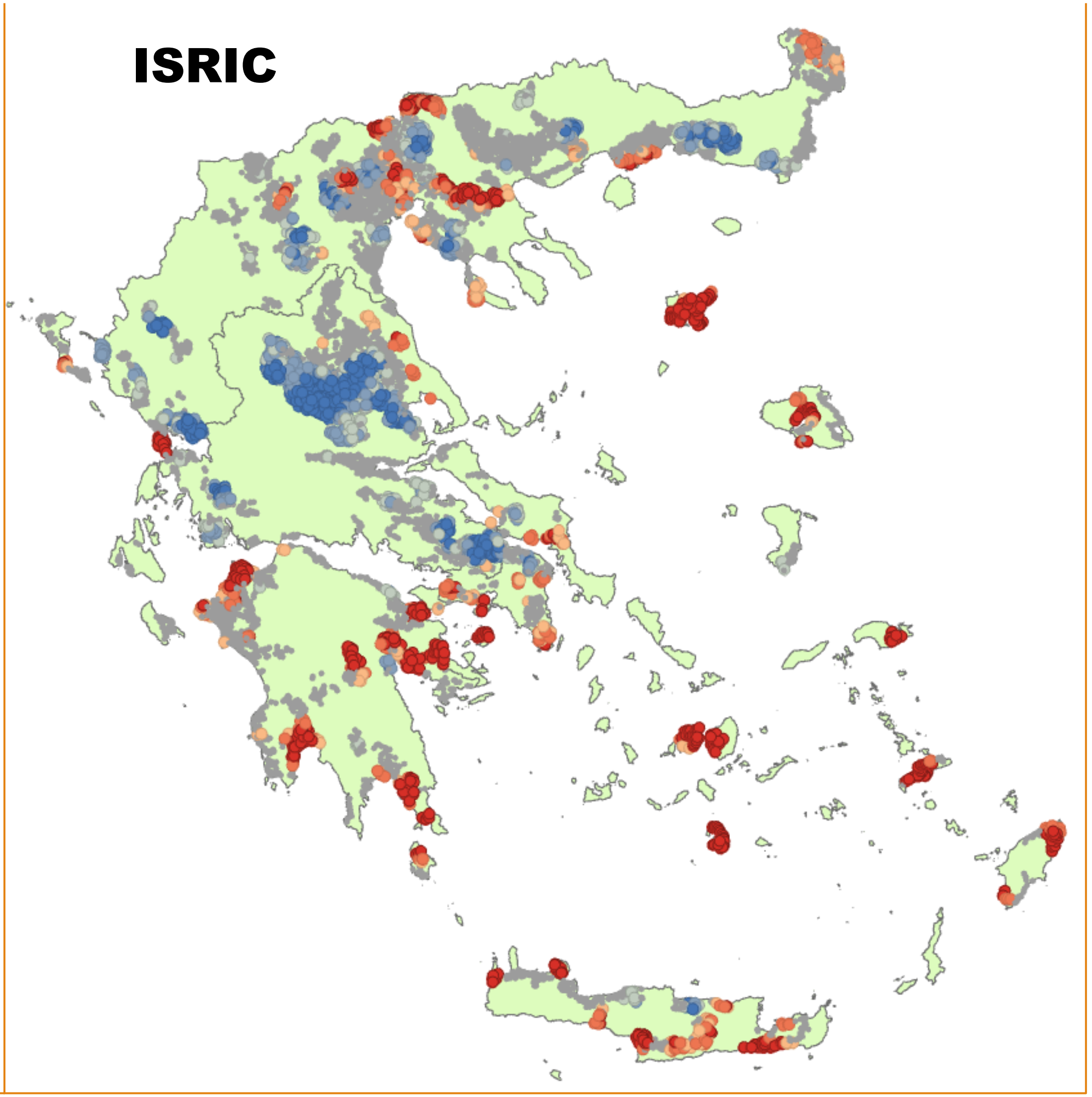
- Large difference between datasets
- Small difference between datasets

- Cold Spot with 99% Confidence
- Cold Spot with 95% Confidence
- Cold Spot with 90% Confidence
- Not Significant
- Hot Spot with 90% Confidence
- Hot Spot with 95% Confidence
- Hot Spot with 99% Confidence

Spatial distribution of the clustering of absolute errors

Hot Spot Analysis

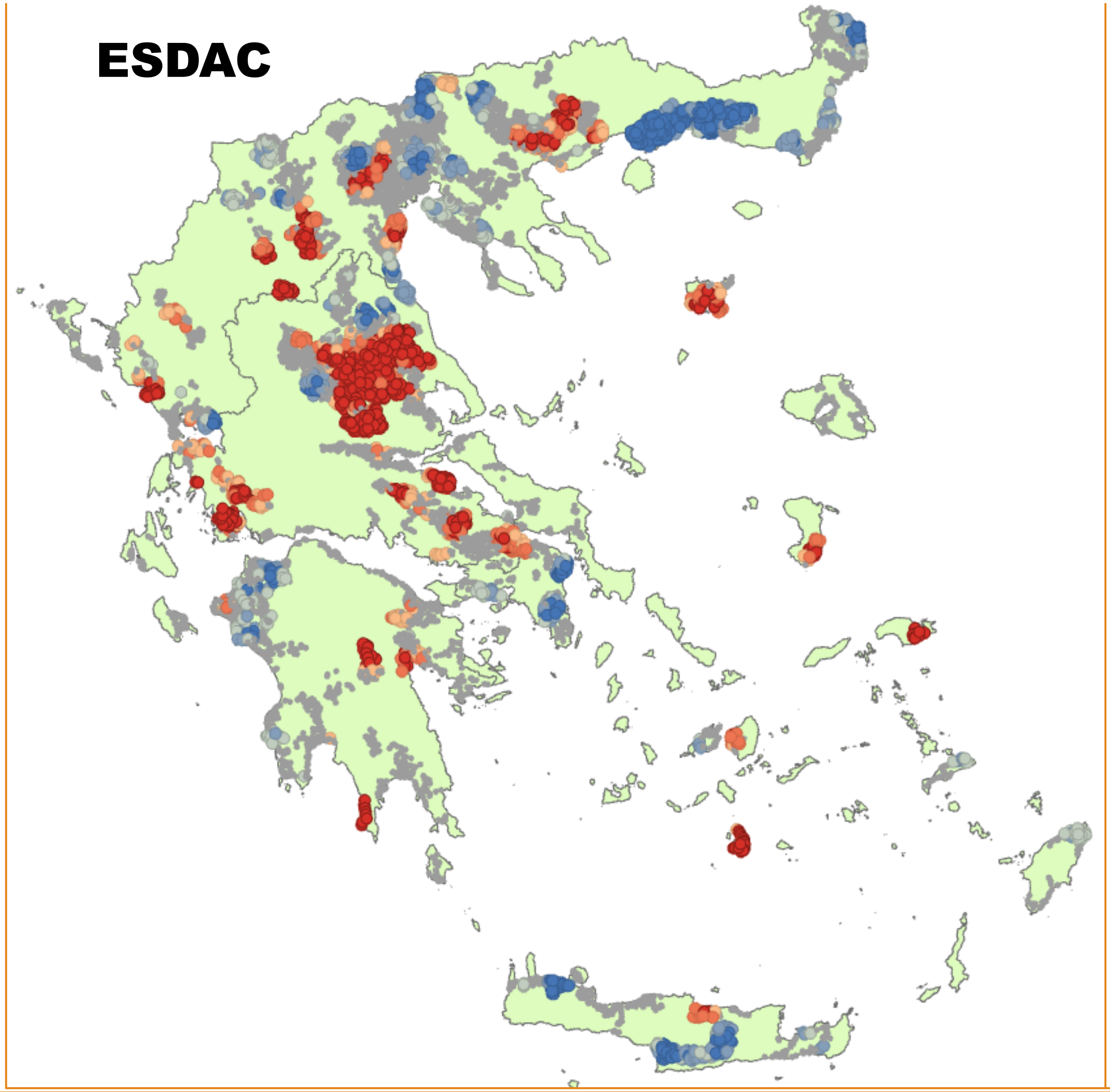


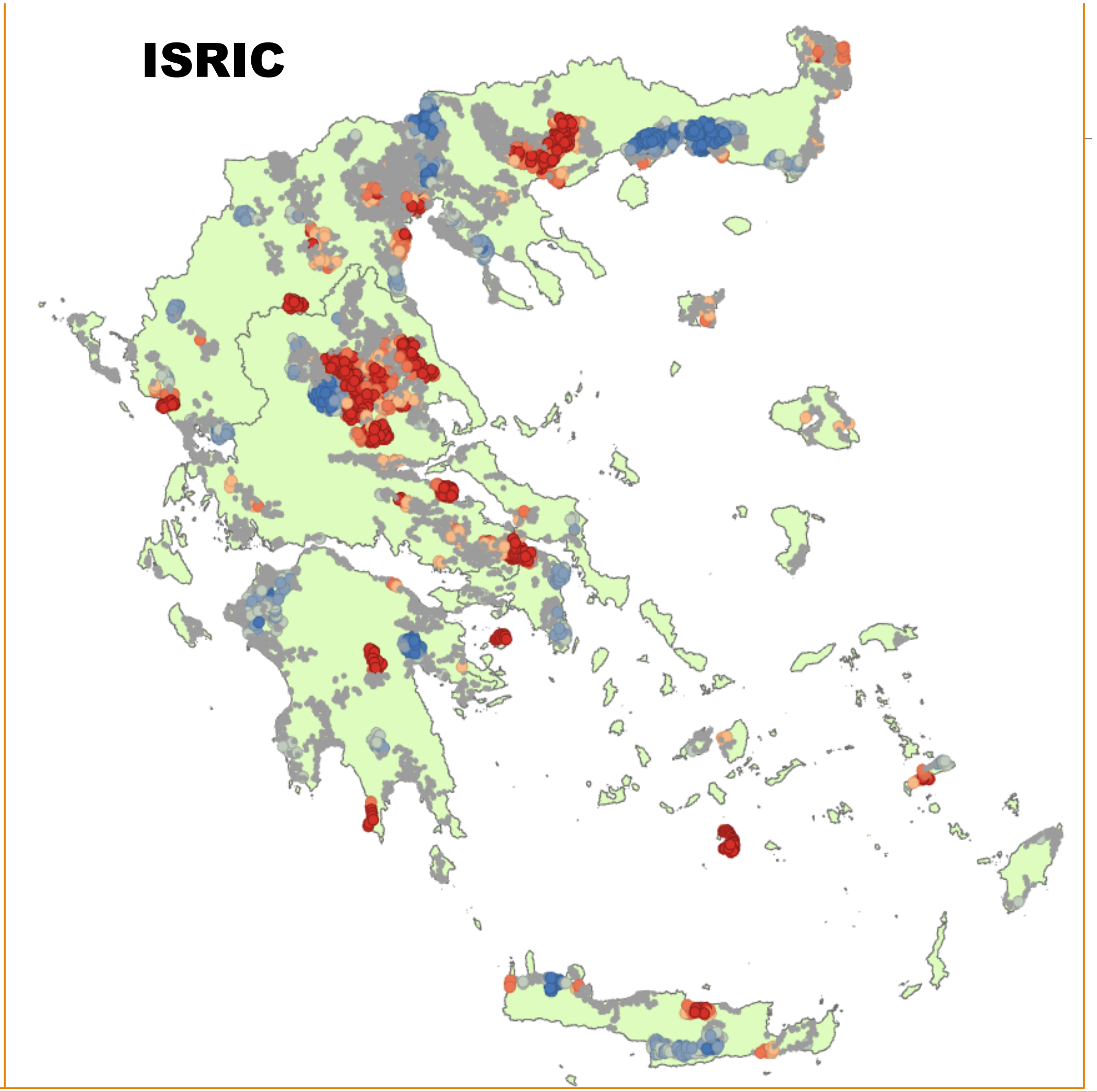


SAND

Spatial distribution of the clustering of absolute errors

Hot Spot Analysis

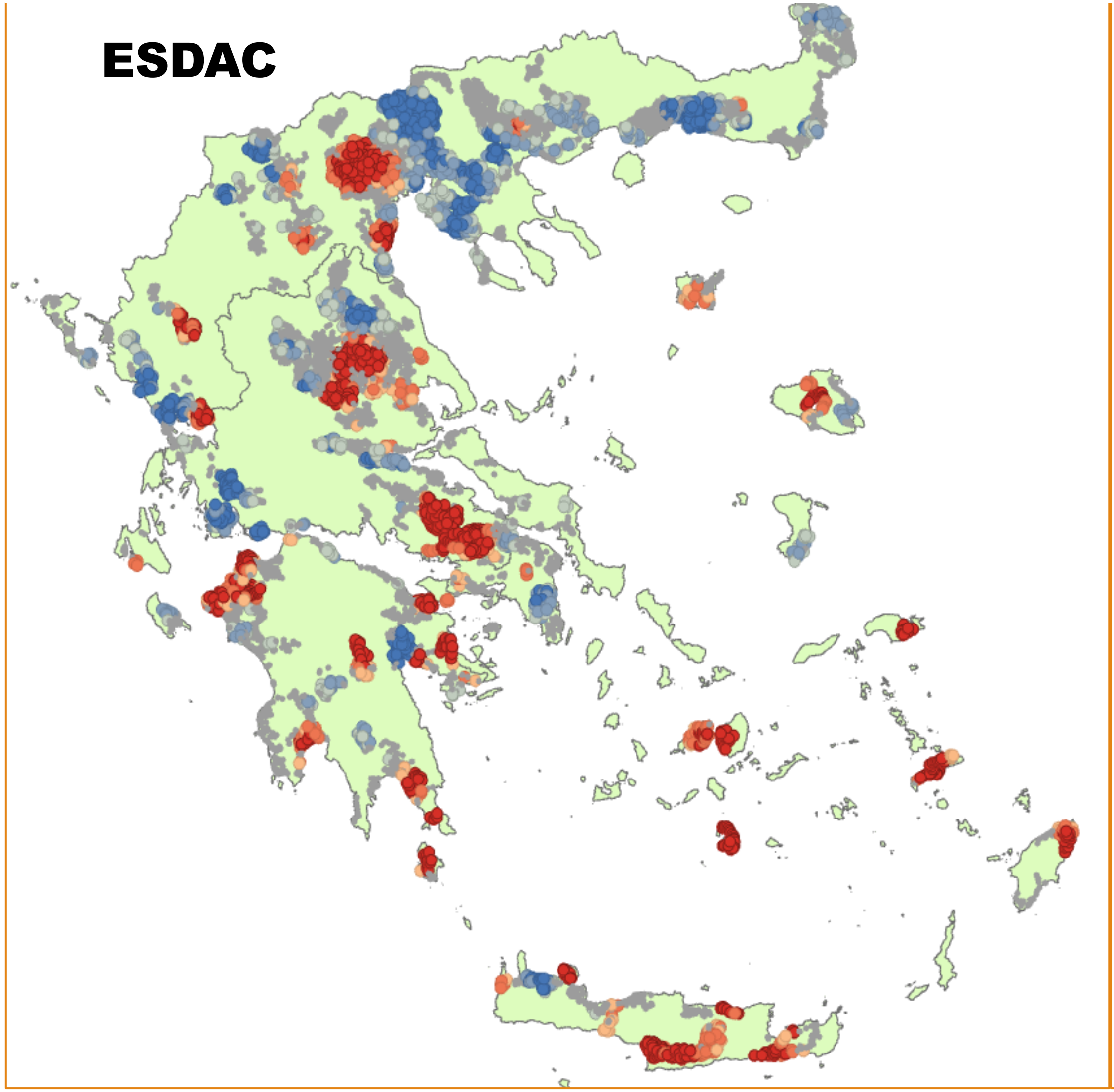


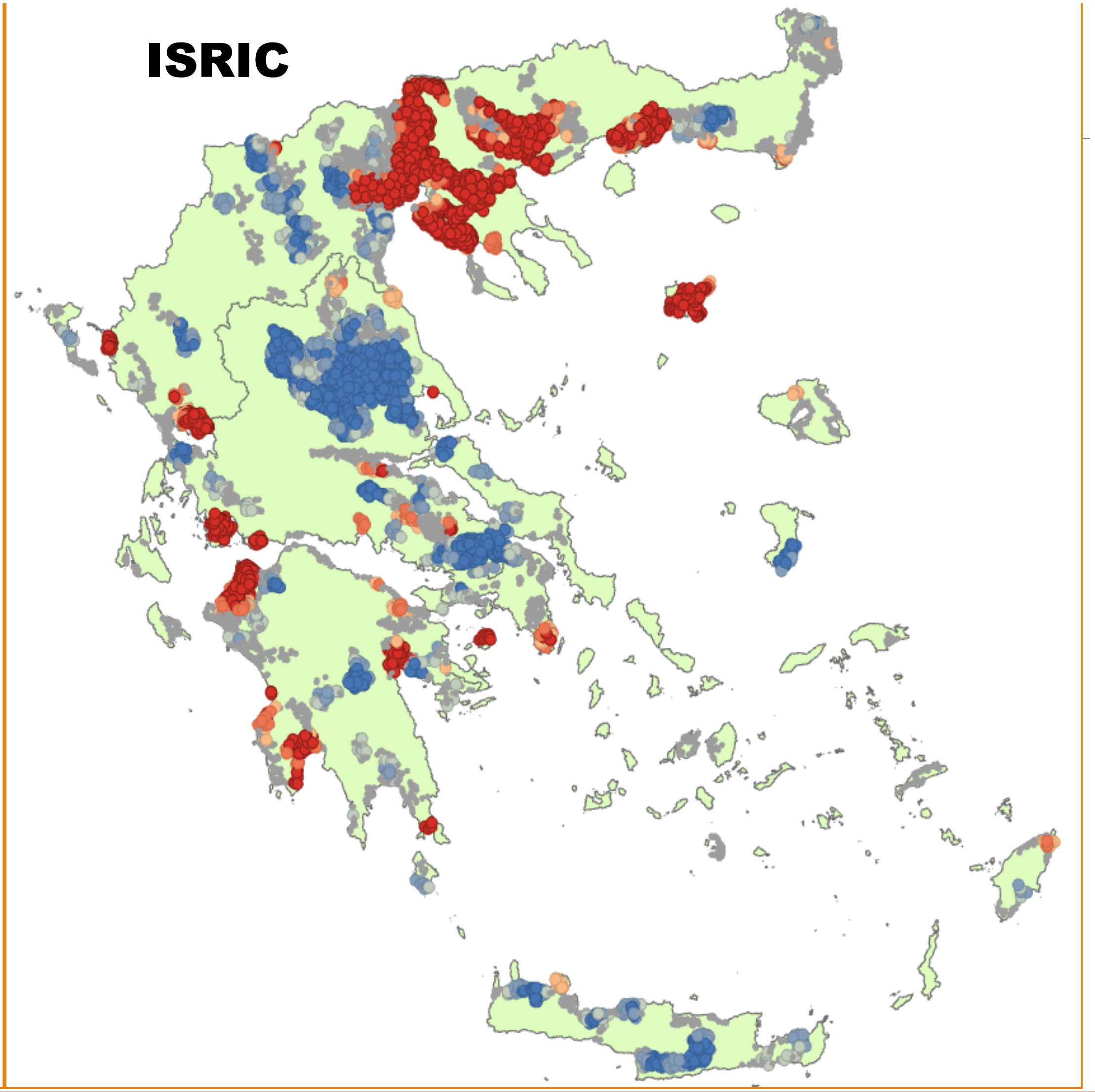


CLAY

Spatial distribution of the clustering of absolute errors

Hot Spot Analysis





SILT

Results – Accuracy of estimation of soil texture classes

GR VS ISRIC

Class	Producer's accuracy	User's accuracy	Overall accuracy
Clay	0.01602262	0.311926606	**0.188955558**
Clay Loam	0.410233789	0.248929336	
Loam	0.505322924	0.158151932	
Loamy Sand	0	0	
Sand	0.02173913	0.027027027	
Sandy Clay	0	0	
Sandy Clay Loam	0.018698579	0.265957447	
Sandy Loam	0.066031746	0.403100775	
Silt Loam	0.086734694	0.059233449	
Silty Clay	0.043478261	0.128205128	
Silty Clay Loam	0.093283582	0.034387895	

GR VS ESDAC

Class	Producer's accuracy	User's accuracy	Overall accuracy
Clay	0.007445323	0.228571429	**0.204095835**
Clay Loam	0.260232658	0.243941842	
Loam	0.530598053	0.161174482	
Loamy Sand	0.005934718	0.4	
Sand	0	0	
Sandy Clay	0	0	
Sandy Clay Loam	0.132034632	0.332727273	
Sandy Loam	0.283312578	0.325931232	
Silt Loam	0.085427136	0.049707602	
Silty Clay	0.008583691	0.042553191	
Silty Clay Loam	0.04029304	0.030726257	

Conclusions

- Initial results provide mixed picture with differences between the datasets greatly varying
- The comparison revealed **spatially variable differences** between the datasets
- **Discrepancies more pronounced** in areas in areas with distinct soil characteristics, such as fine-textured soils.
- Findings emphasize the **importance of integrating detailed national soil data** to enhance the accuracy of continental-scale digital soil maps.
- Results highlight the need for **data integration from multiple sources** to capture soil variability more effectively.
- The study contributes to the development of **more robust and reliable global soil datasets**.
- It offers **practical recommendations** for improving future soil mapping initiatives.
- Further research should focus on **interpretation of errors clusters** (e.g. geomorphology, geology, land cover, sampling density etc.) to improve datasets.

Thank you!

This project is carried out within the framework of the National Recovery and Resilience Plan Greece 2.0, funded by the European Union – NextGenerationEU (Implementation body: HFRI), https://greece20.gov.gr