4th MOPT CONFERENCE

Spatial Modeling for Sustainable Cities and Territories

Delimitation of soil districts: A new paradigm in Portuguese soil mapping and monitoring

JORGE ROCHA

MAY 22, 2025 IGOT | University of Lisbon, Portugal

LA/P/0092/2020
CUIDB/00295/2020
UIDP/00295/2020

Why is there a need for EU-wide soil monitoring?

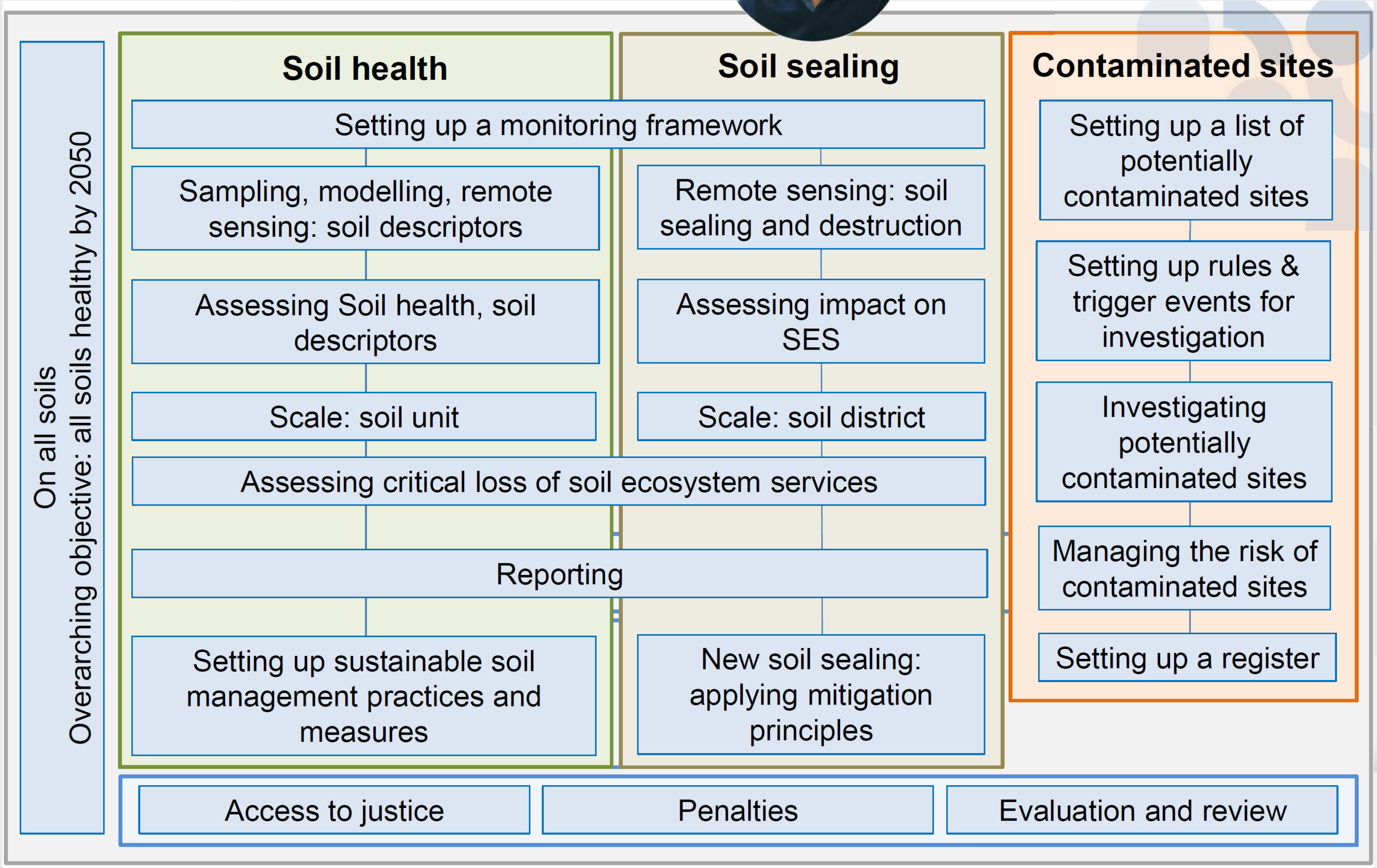
1. Unbiased estimation of the soil parameters on which the indicators are established;
2. Optimal coverage of different types of land use that drive soil degradation processes;
3. Optimal coverage of topographic, climatic, geological and biological variability of factors influencing soil characteristics;
4. Economic sampling.

4th MOPT CONFERENCE

Planning interventions to mitigate soil sealing impacts and adaptation to climate change in urban areas (UnSealingCities)

Monitoring the impact on Ecosystem Services through different soil management practices to inform sustainable land use and occupation policies (MonLand)

Co-participative Modelling of Soil Districts based on Machine Learning (ML-SOIL)



EUROPEAN COMMISSION

Brussels, 5.7.2023
COM(2023) 416 final

2023/0232 (COD)

Proposal for a

DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

on Soil Monitoring and Resilience (Soil Monitoring Law)

{SEC(2023) 416 final} - {SWD(2023) 416 final} - {SWD(2023) 417 final} - {SWD(2023) 418 final} - {SWD(2023) 423 final}

Article 4

Soil districts

1. Member States shall establish soil districts throughout their territory.

   The number of soil districts for each Member State shall as a minimum correspond to the number of NUTS 1 territorial units established under Regulation (EC) No 1059/2003.

2. When establishing the geographic extent of soil districts, Member States may take into account existing administrative units and shall seek homogeneity within each soil district regarding the following parameters:

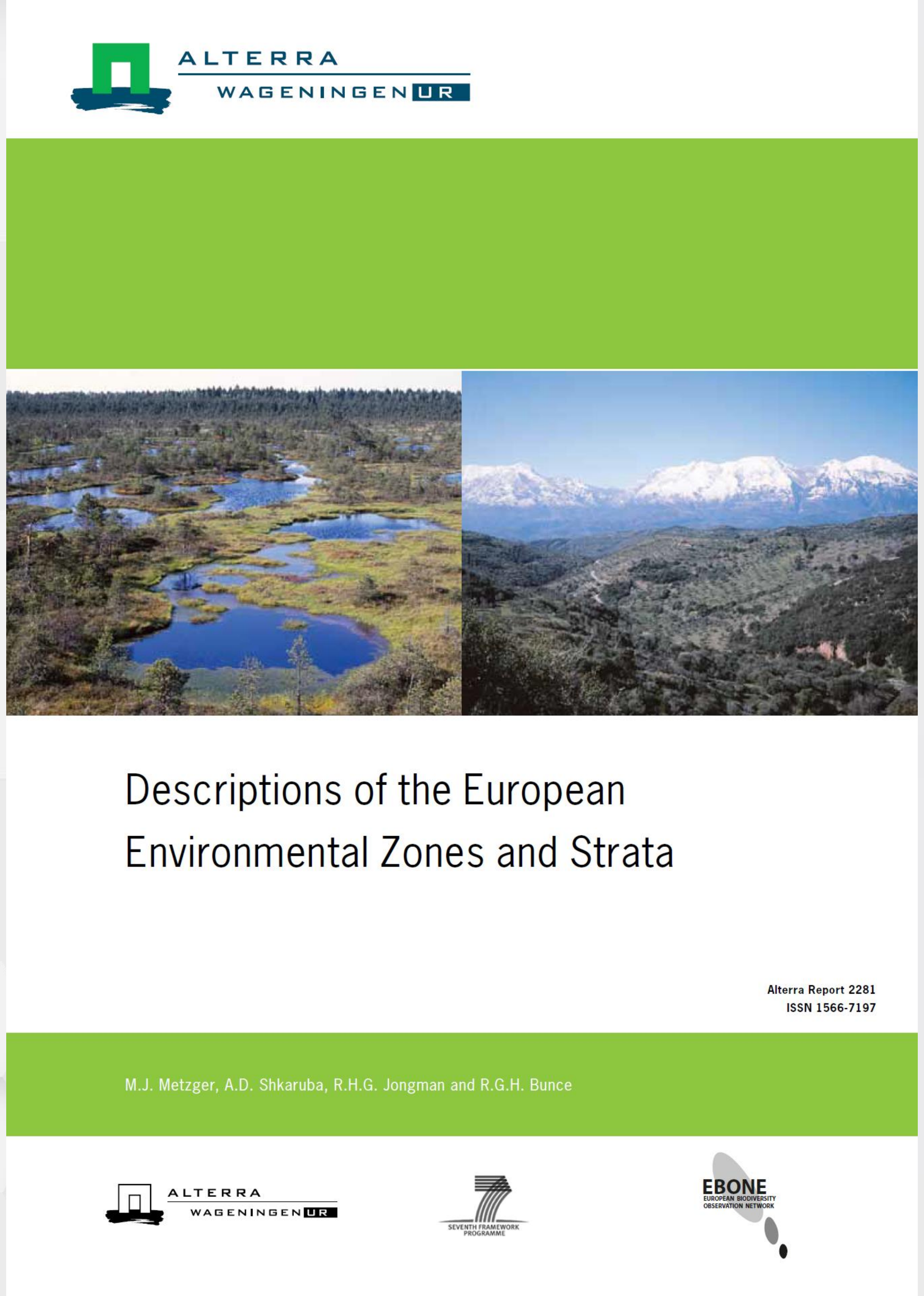
   (a) soil type as defined in the World Reference Base for Soil Resources[74];

   (b) climatic conditions;

   (c) environmental zone as described in Alterra Report 2281[75];

   (d) land use or land cover as used in the Land Use/Cover Area frame statistical Survey (LUCAS) programme.

4 th MOPT CONFERENCE



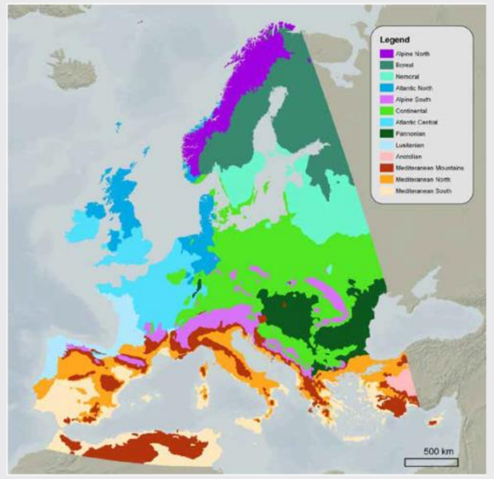
Twenty of the most relevant available environmental variables were selected, based on those identified by statistical screening (Bunce et al., 1996c). These were

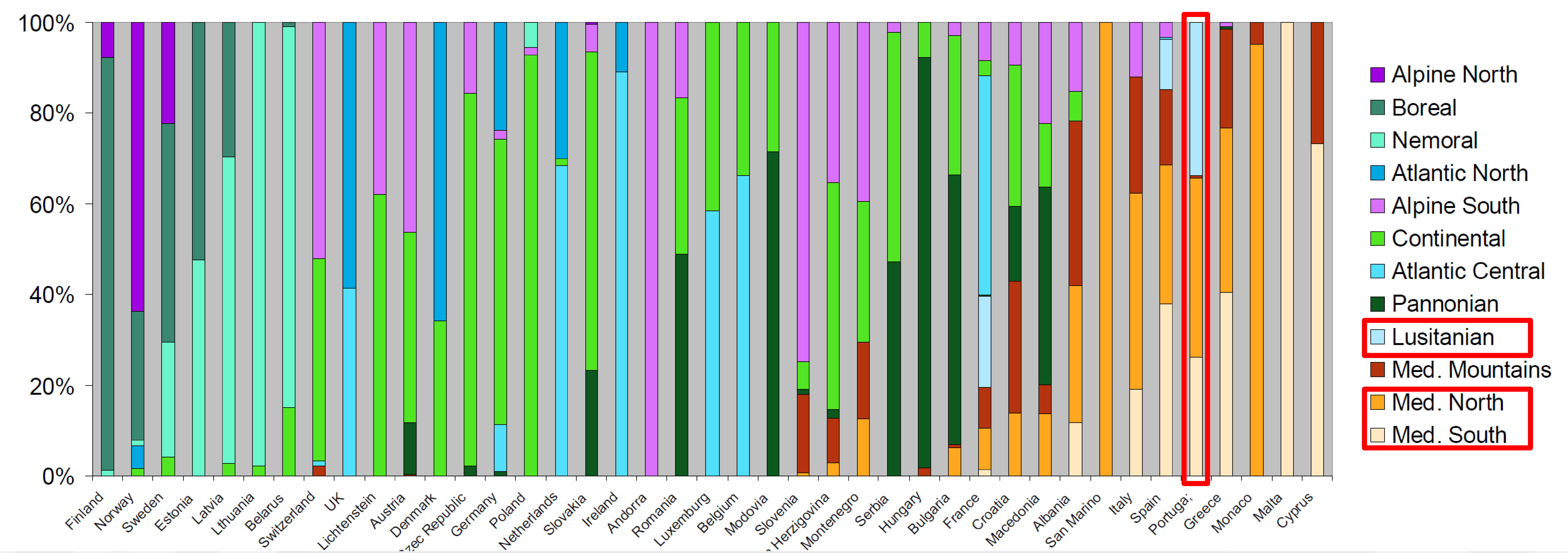
(1) climate variables from the Climatic Research Unit (CRU) TS1.2 dataset (Mitchell et al., 2004),
(2) elevation data from the United States Geological Survey HYDRO1k digital terrain model, and
(3) indicators for oceanicity and northing.

Principal Component Analysis (PCA) was used to compress 88% of the variation into three dimensions, which were subsequently clustered using an ISODATA clustering routine. The classification procedure is described in detail by Metzger et al. (2005a).

https://edepot.wur.nl/197197

4 th MOPT CONFERENCE





The LUCAS experience

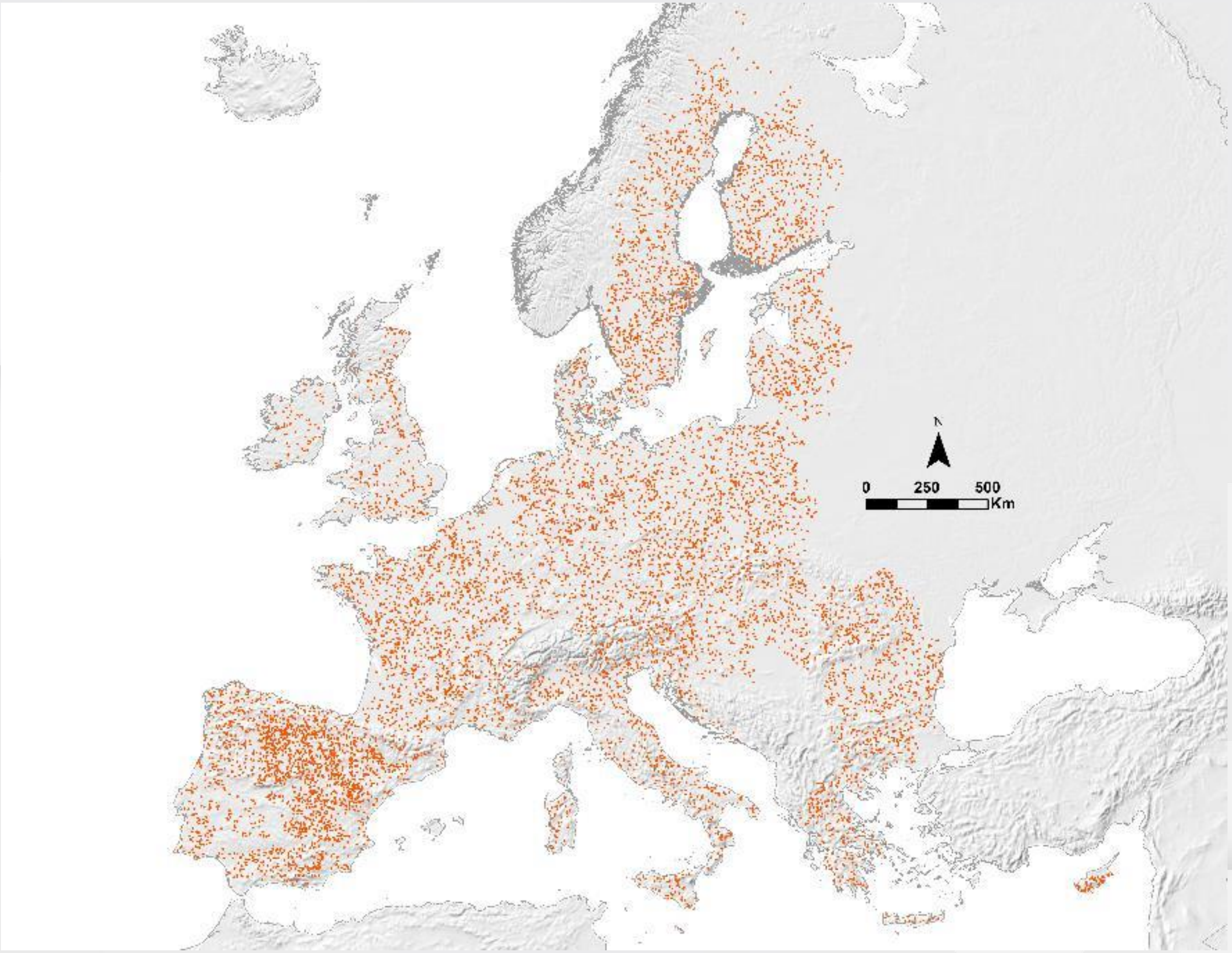
Stratified random sampling

Samples from 22,000 sites across the EU (2009/2012, 2015, 2018)

Samples from 40,000 locations across the EU (2020)

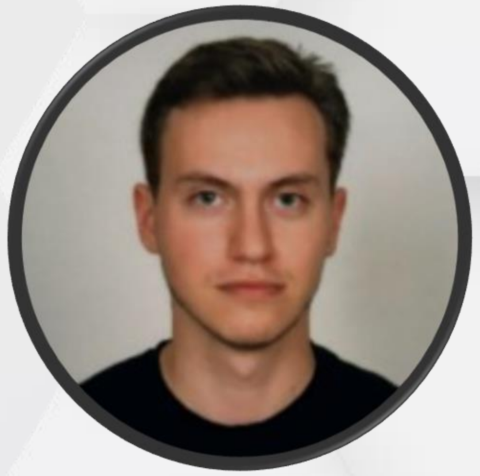
“Snapshot” survey: conducted in one year

Only harmonized soil data collection program for the EU (so far…)



4th MOPT CONFERENCE

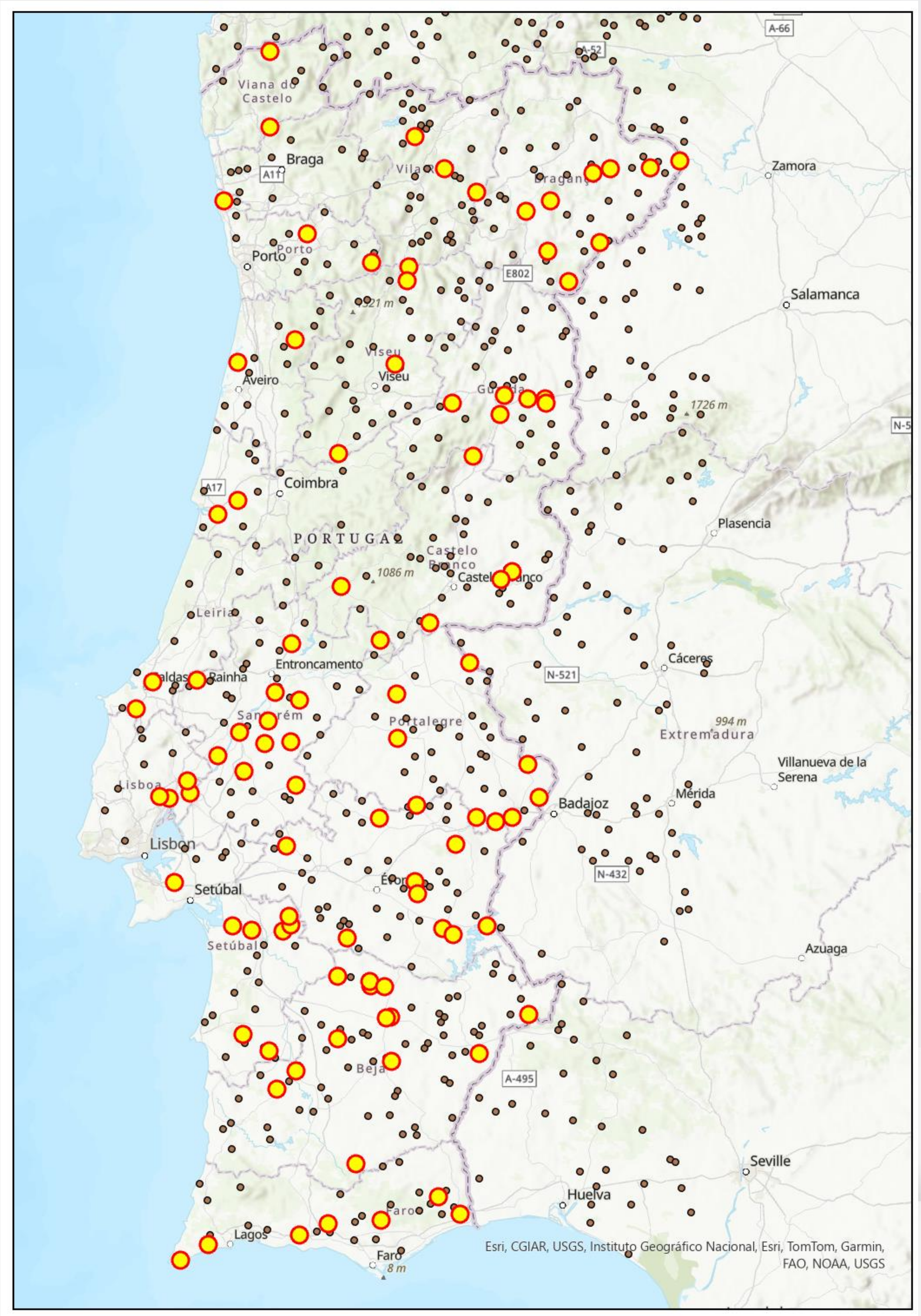
The Portuguese LUCAS experience



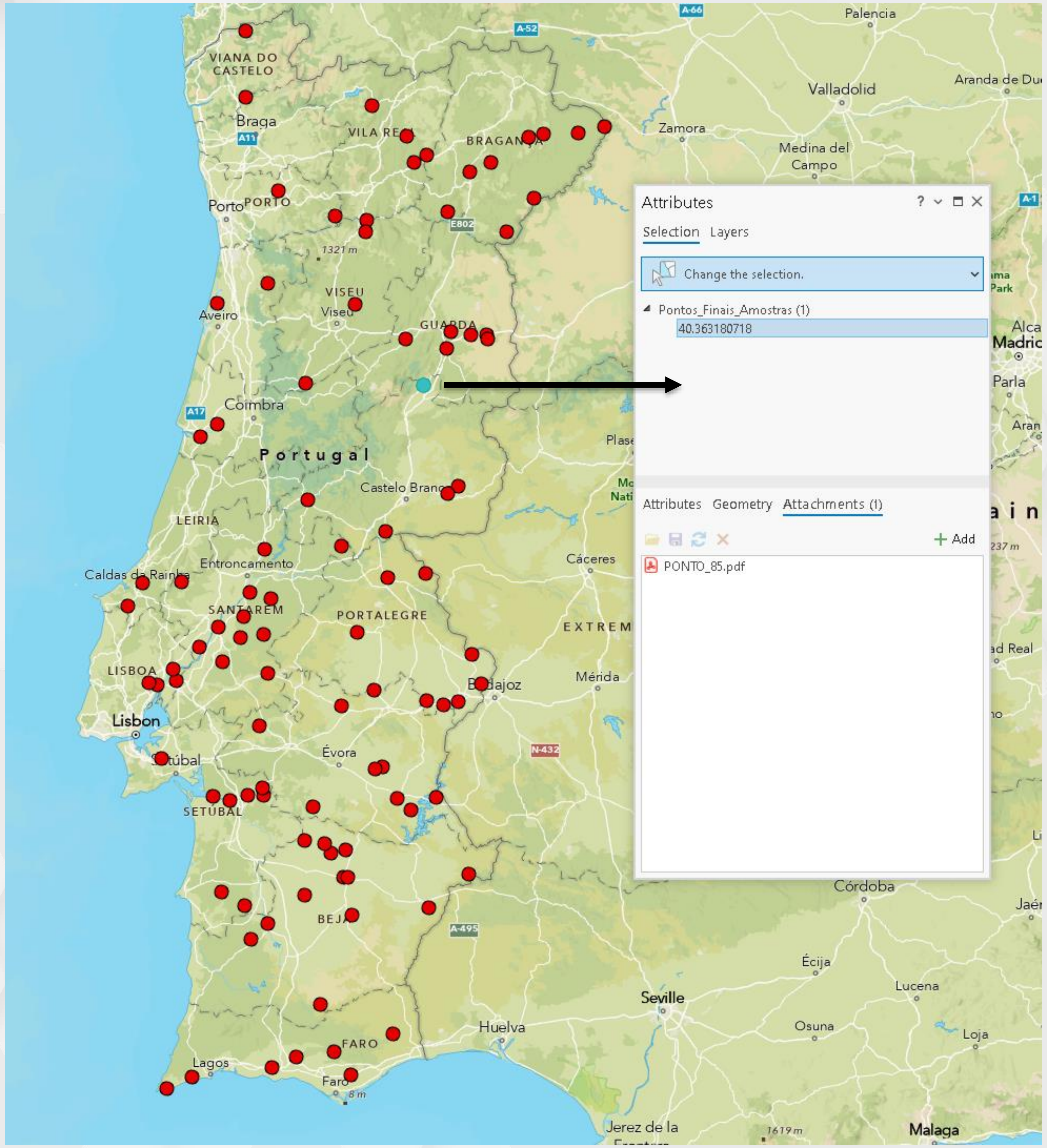
Field data and sample assembly manual

https://parceriaptsolo.dgadr.gov.pt/recursos/publicacoes

https://parceriaptsolo.dgadr.gov.pt/images/Manual_de_Colheita_de_Amostras_e_de_Dados_no_Campo_4.pdf



4 th MOPT CONFERENCE



FICHA DE CAMPO

Nº SEQUENCIAL associado ao ponto de amostragem georreferenciado identificado na TABELA / Nº ID	85		
Concelho:	**Covilhã**	Freguesia:	**União de freguesias de Vale Formoso e Aldeia do Souto**
A classe de cobertura de solo do ponto georreferenciado corresponde à indicada na TABELA	Sim □	Não X □	Qual? Oliveiras

Desvio na colheita da amostra:	Sim □	Não X □
Distância do ponto de amostragem georreferenciado identificado na TABELA ao ponto central de colheita da amostra	_m	
Motivo do desvio (problemas de acessibilidade, dureza do terreno, cobertura do solo – casas, árvores, etc.):		
Classe de cobertura do solo do ponto de amostragem:		

Desvio na colheita de uma ou mais subamostras:	Sim □		Não X □	
Qual ou quais as subamostras colhidas com desvio:	N□	S □	E □	O □
Distância do ponto de colheita da subamostra ao ponto predefinido:				
Motivo do desvio (excessiva quantidade de arbustos espinhosos, presença de rocha abaixo da superfície, presença de raízes, etc.):				

Colheita da amostra em local de difícil acesso (cultura agrícola) - Amostragem em padrão linear:	Sim X □	Não □
O ponto central corresponde ao ponto na entrelinha georreferenciado na TABELA		X □
O ponto central corresponde ao ponto na entrelinha mais próximo do georreferenciado na TABELA (localizado na linha). Distância: ______m		□
O ponto central corresponde ao ponto na entrelinha deslocado por o ponto georreferenciado na TABELA ser demasiado próximo da estrada. Distância: __________m		□
Percentagem da superfície coberta com resíduos de vegetação e pedras na área de 2 m de raio que serve de base para a recolha de solo		

0 -10 % X □	10 - 25 % □	25 – 50% □	> 50%

Colheita de amostras por método da escavação:	Sim □	Não X □
Motivo:		

Profundidade	Volume de solo escavado nos pontos:					
	C	N	E	S	O	Total
0 - 10 cm						
10 - 20 cm						
20 - 30 cm						

EROSÃO

Sinais de erosão	Sim □	Não X □			
Tipos de erosão	Laminar	□			
	Em sulcos	□ Nº:	< 5 □	5-10 □	>10 □
	Em ravinas	□ Nº:	< 5 □	5-10 □	>10 □
	Movimento de massas	□			
	Redeposição do solo	□			
	Eólica	□			
	Não aplicável	□			
Distância do ponto de amostragem	_______m				
Direção do ponto de amostragem	N □	S □	E □	O □	

PRÁTICAS PARA REDUZIR A EROSÃO

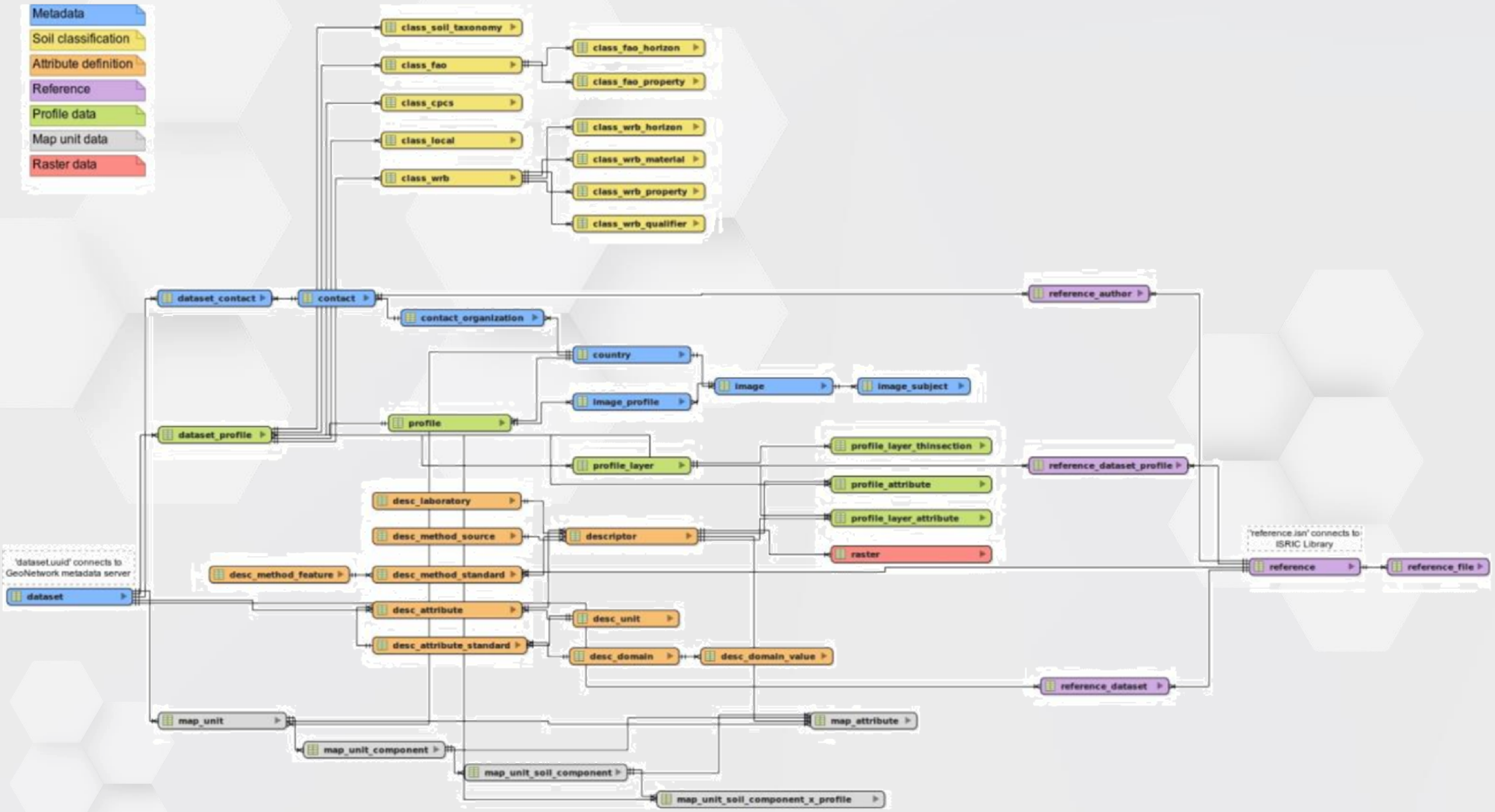
Direção da mobilização:	Transversal ao declive		□
	Ladeira abaixo		□
	Não aplicável		□
Declive do campo mobilizado:	Plano		□
	Declive ligeiro (sem esforço para subir)		□
	Declive acentuado (subir com esforço)		□
	Ondulante (declive em mais do que 1 direção)		□
Presença de resíduos da cultura	Sim □		Não □
Presença de bordaduras com vegetação	Sim □	< 1m largura □ > 1m largura □	Não □
Presença de muros de pedra	Sim □	Não mantidos □ Bem mantidos □	Não □

O proprietário pediu para receber uma cópia dos resultados analíticos X □

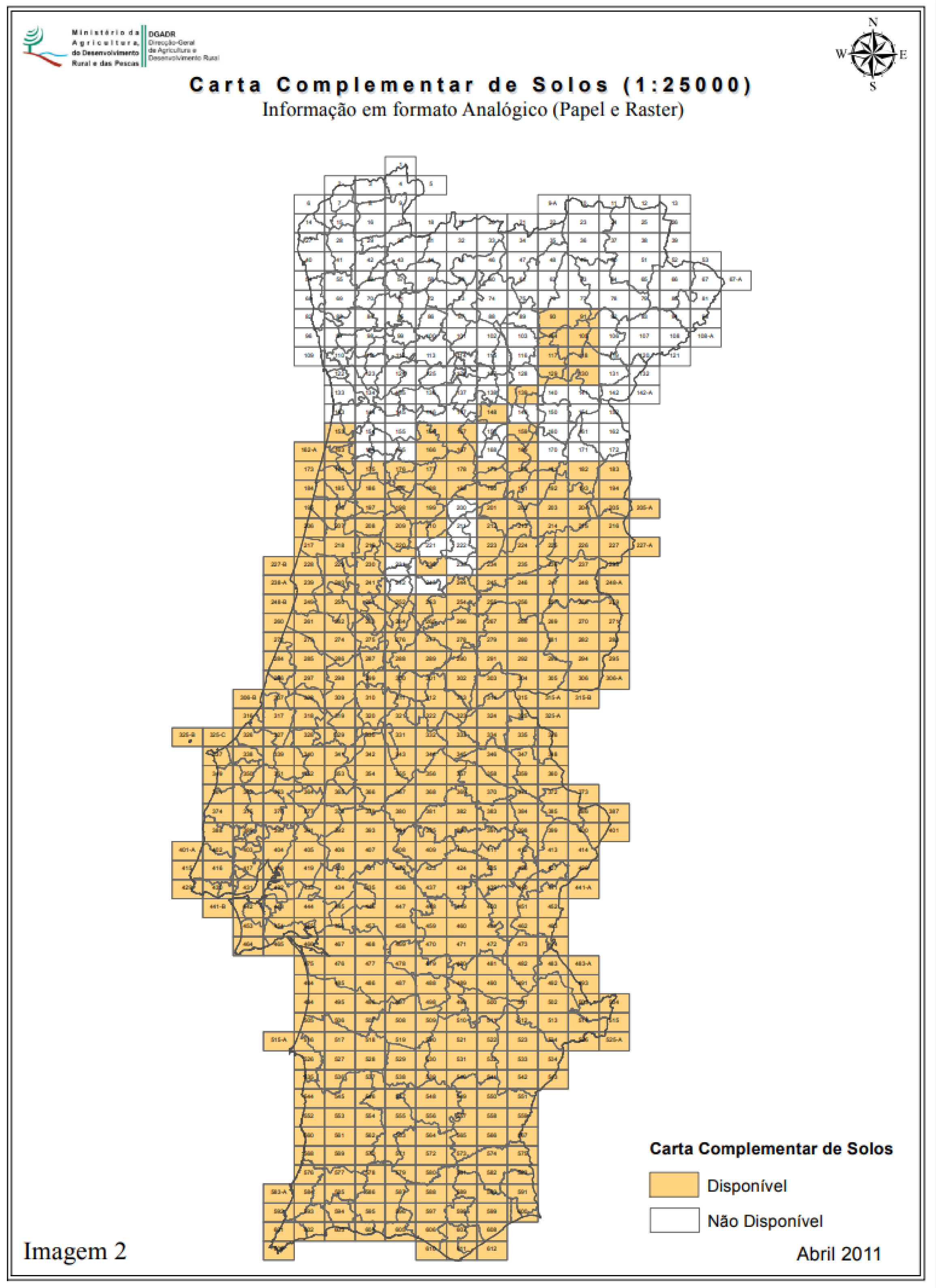
Contacto do proprietário:

4th MOPT CONFERENCE

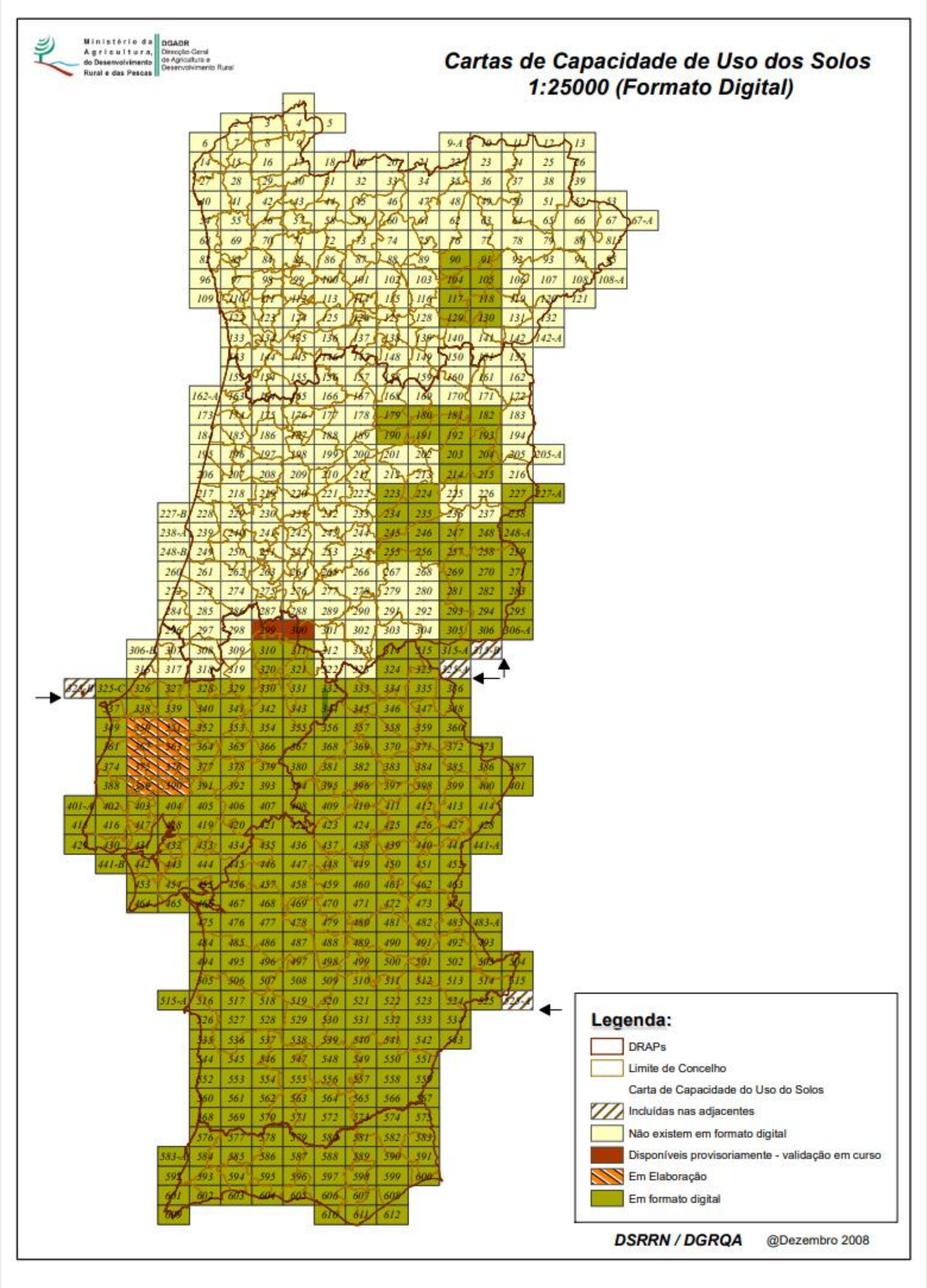
The database structure is that stipulated by the World Reference Base for Soil Resources(WoSIS) of 2018.



4th MOPT CONFERENCE

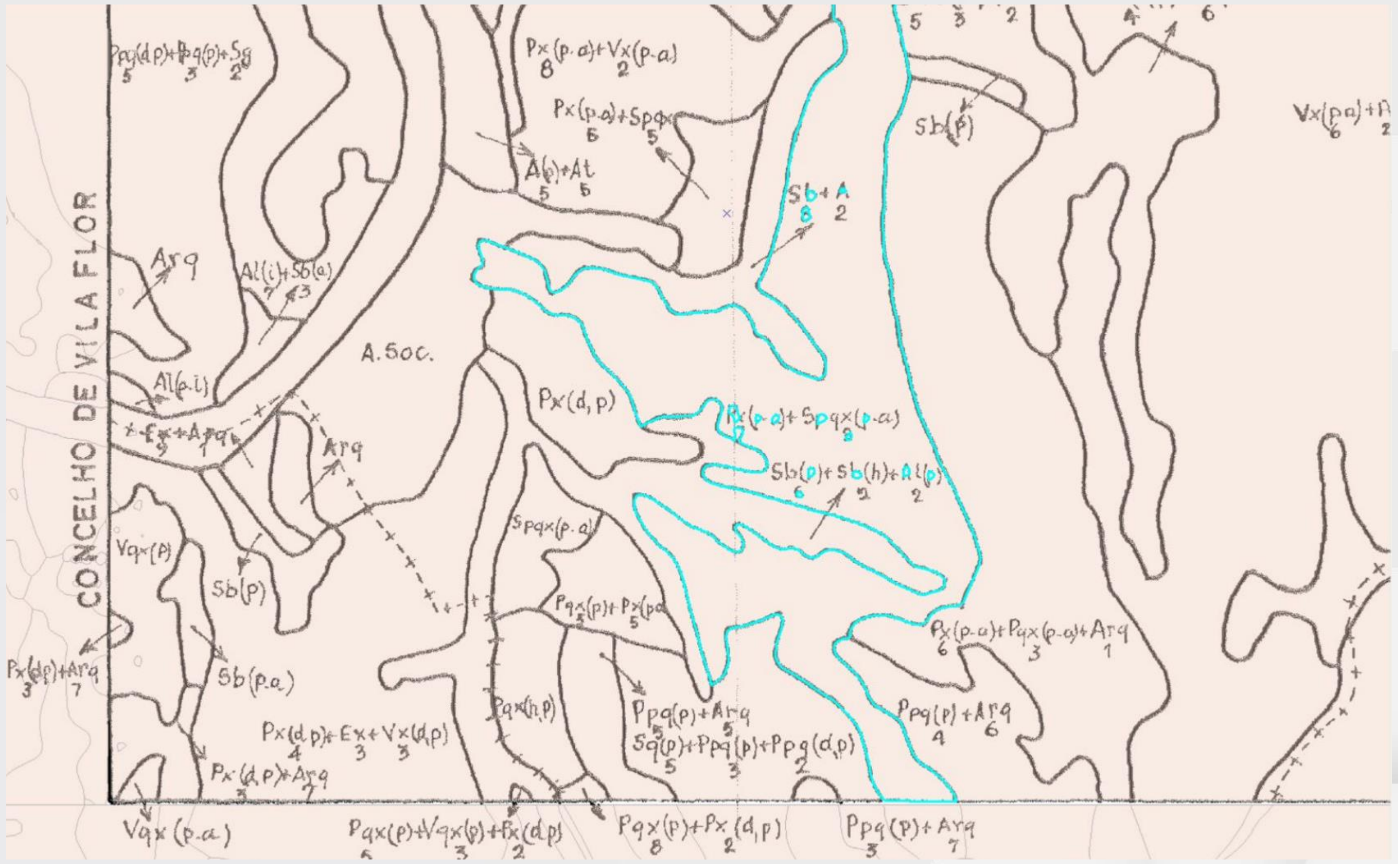


Analogic

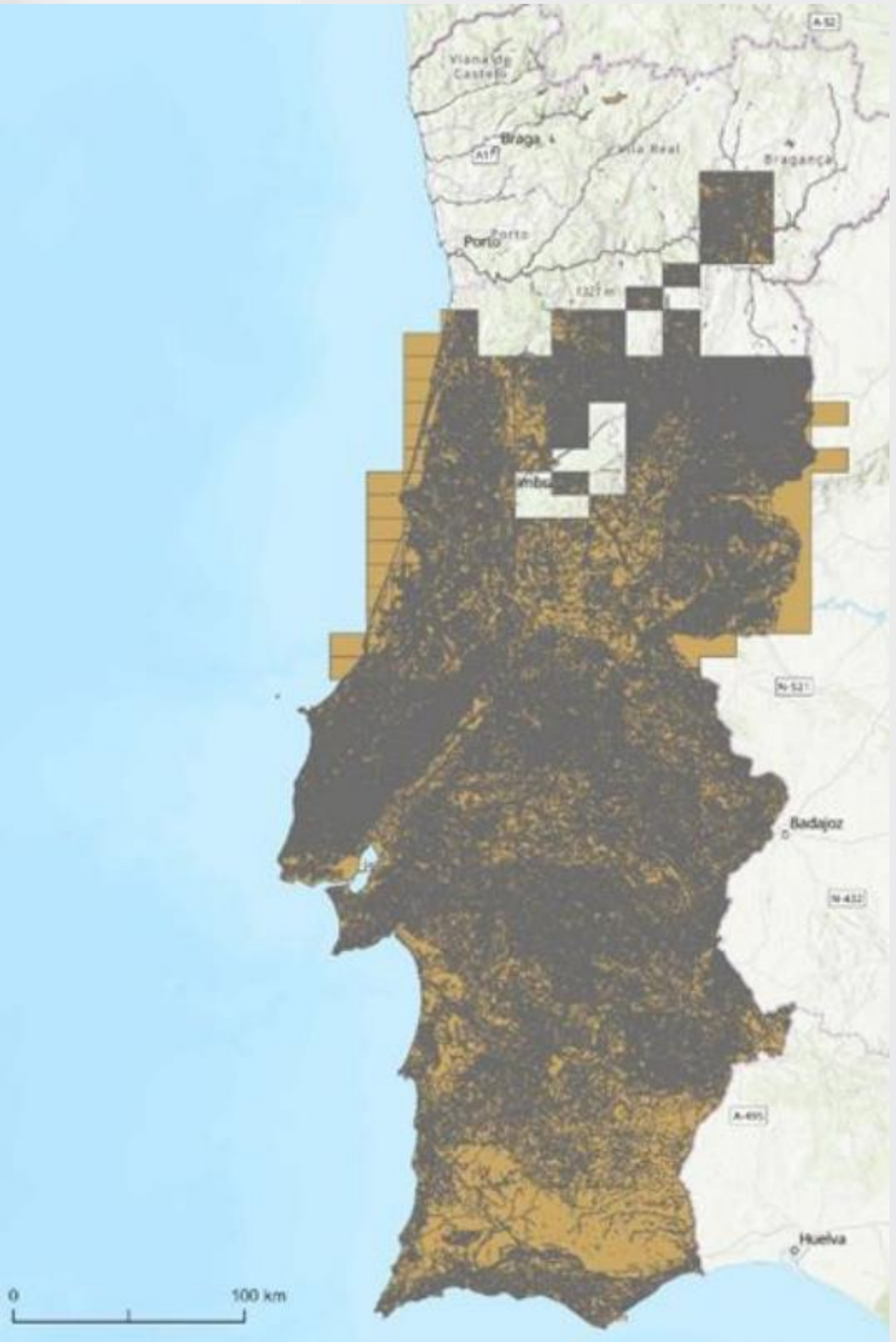


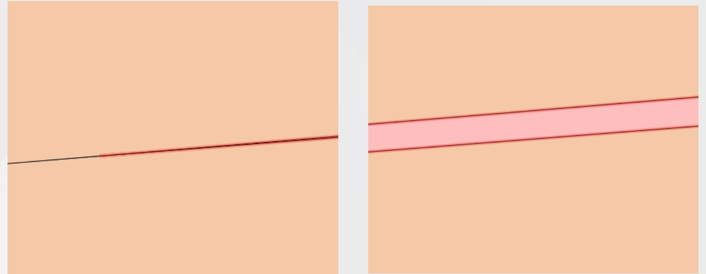
Digital

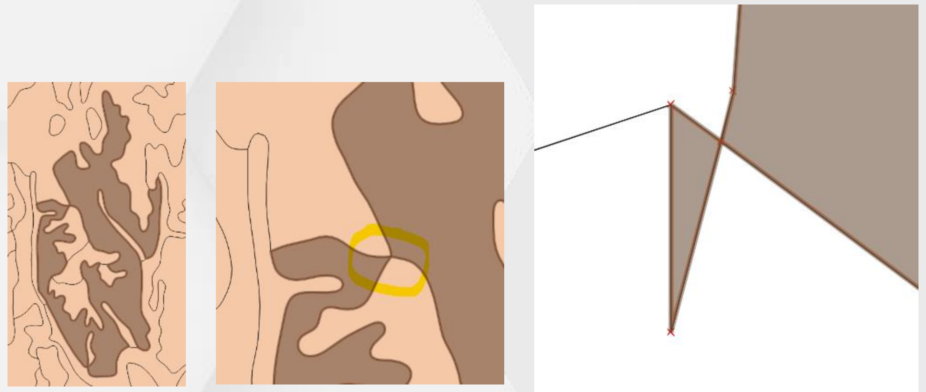
Soil Map at a 1:25,000 Scale

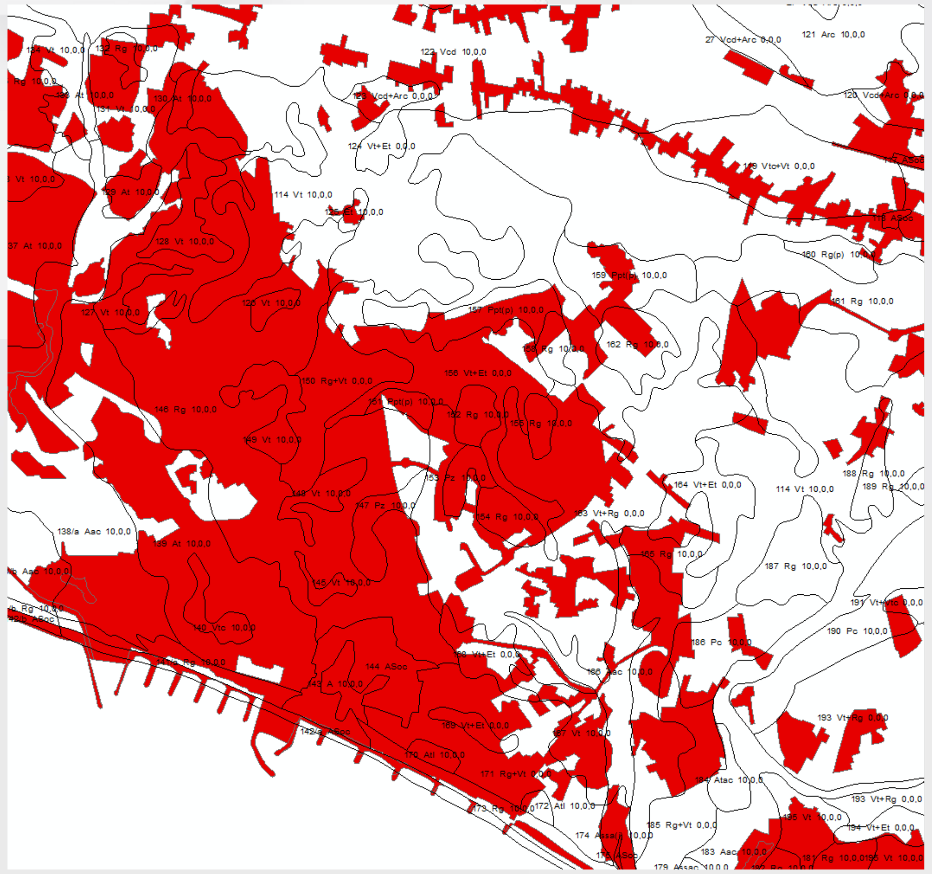


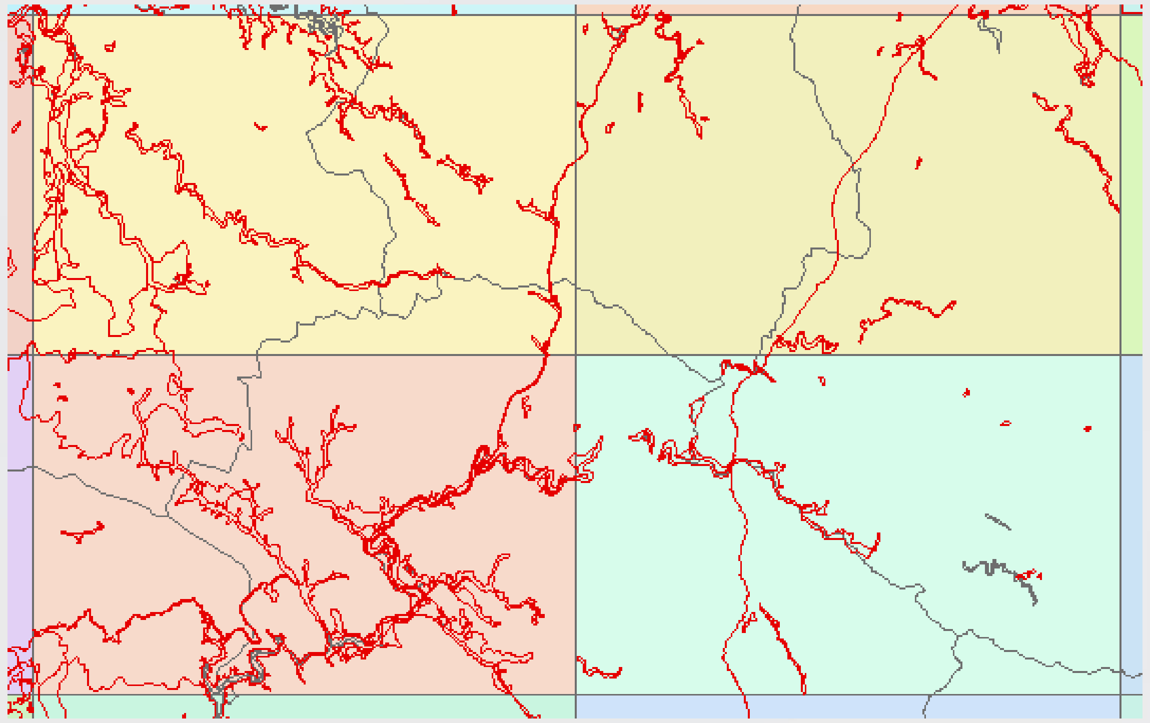
4th MOPT CONFERENCE

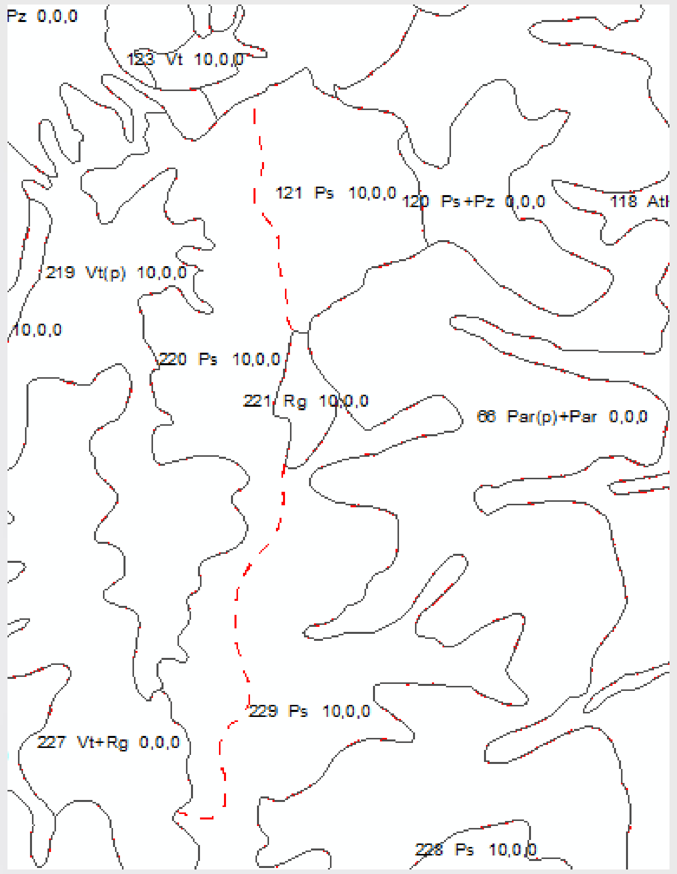


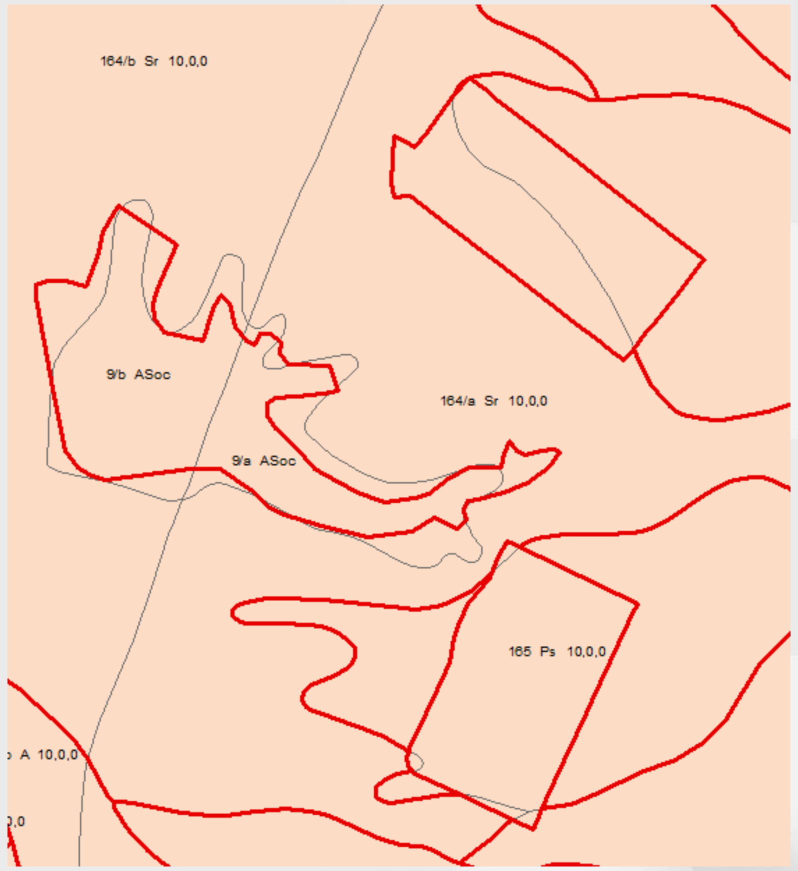






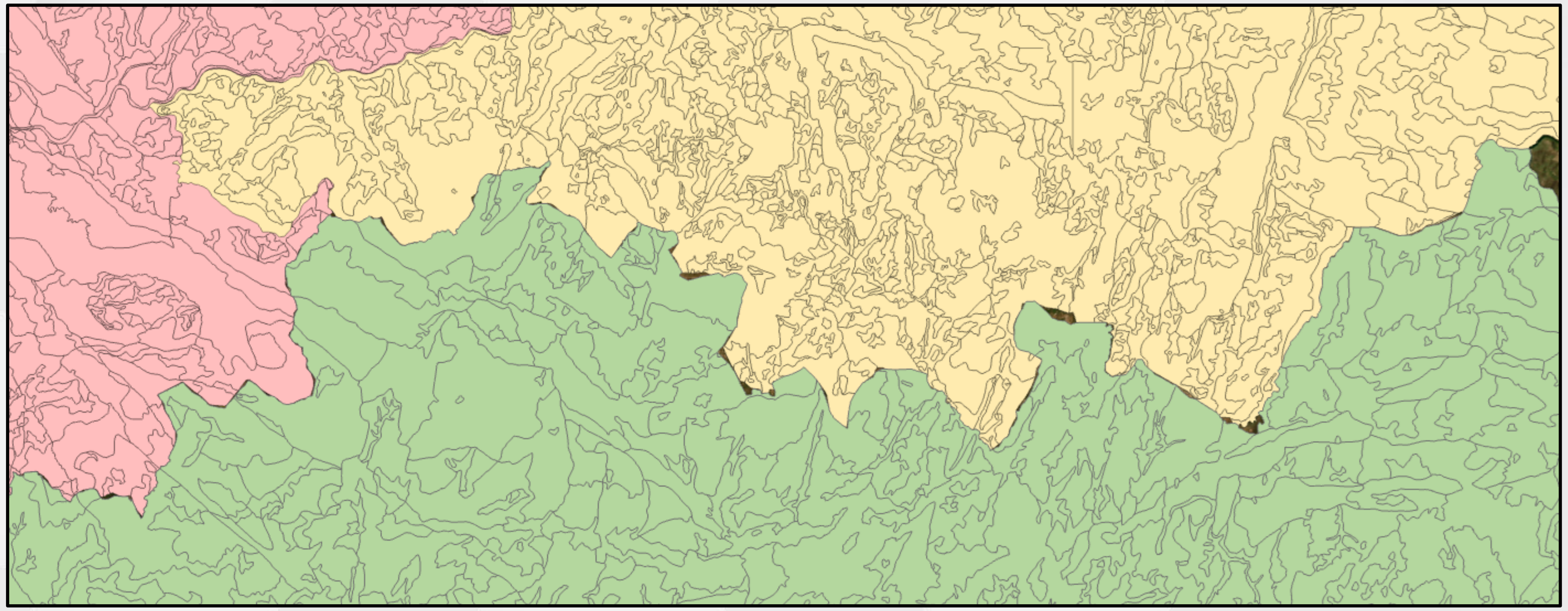






4th MOPT CONFERENCE

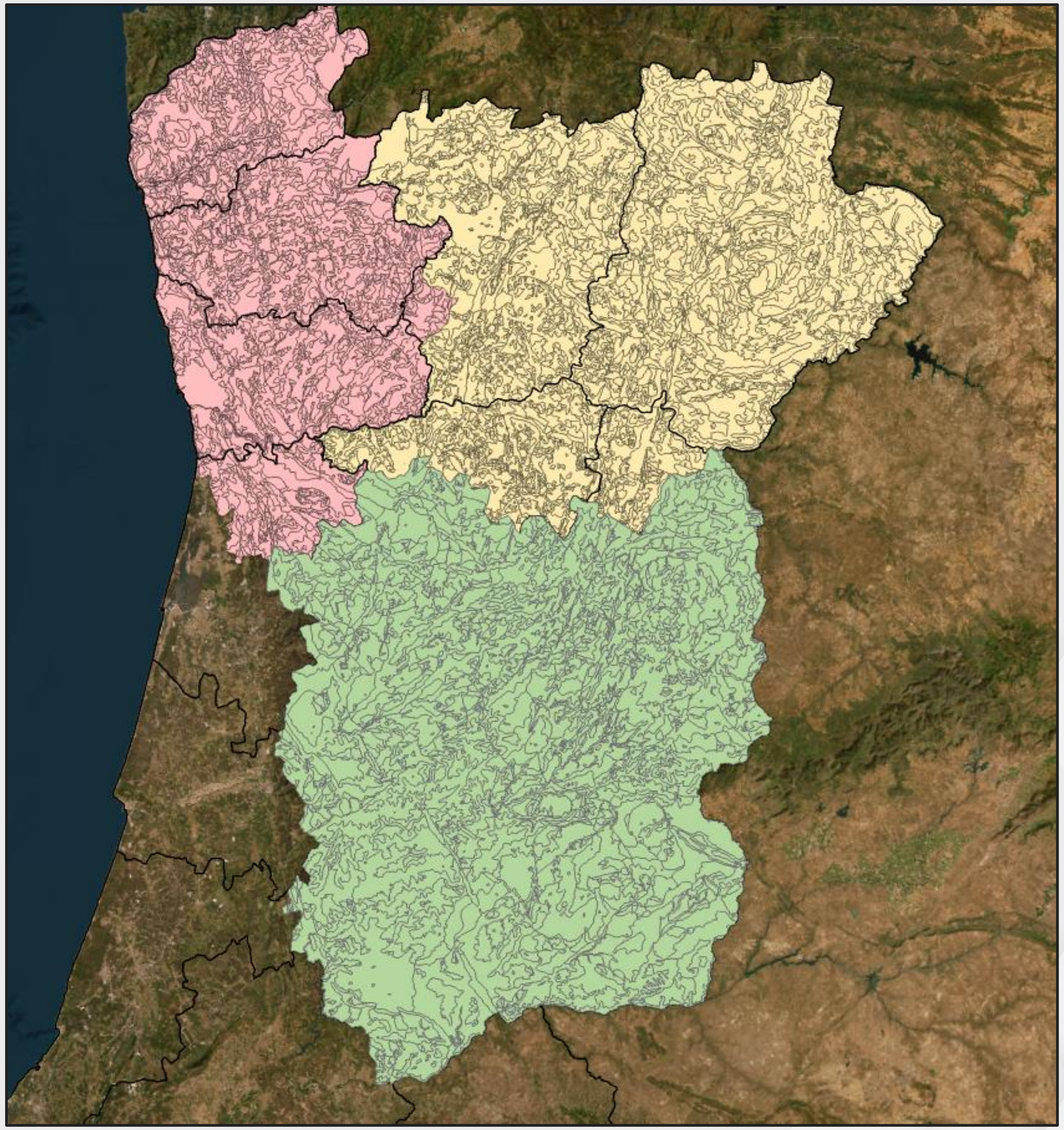
Soil Maps at a 1:100,000 Scale



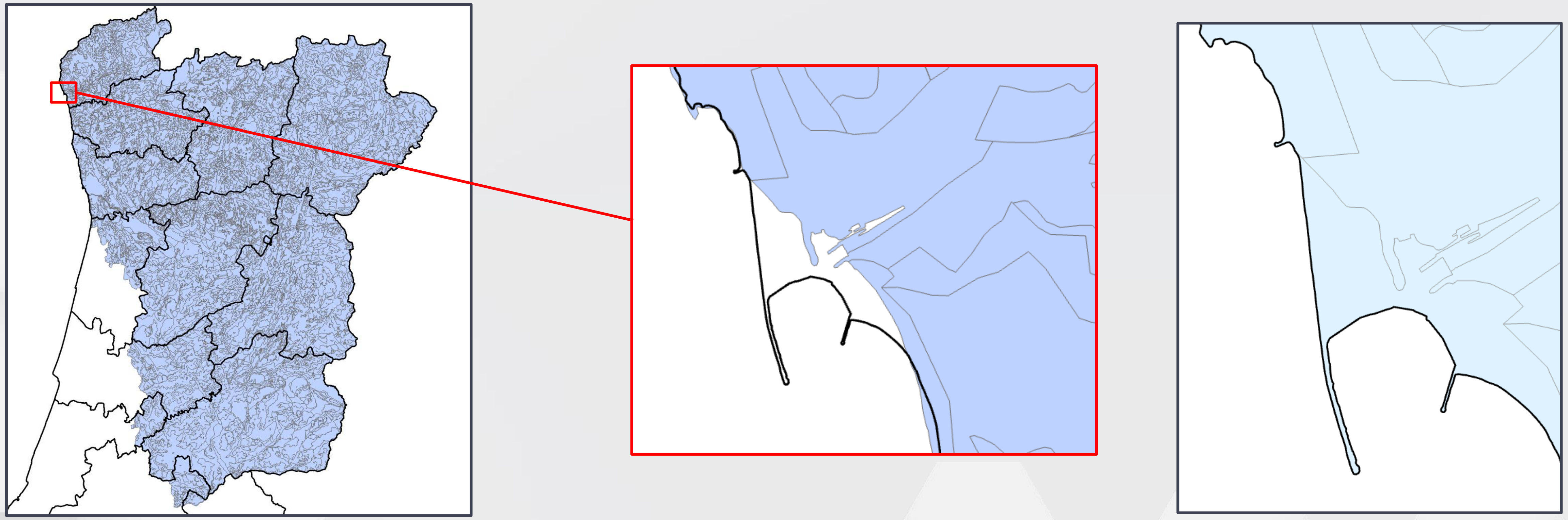
1 - Soil Map of the Entre Douro e Minho Region;

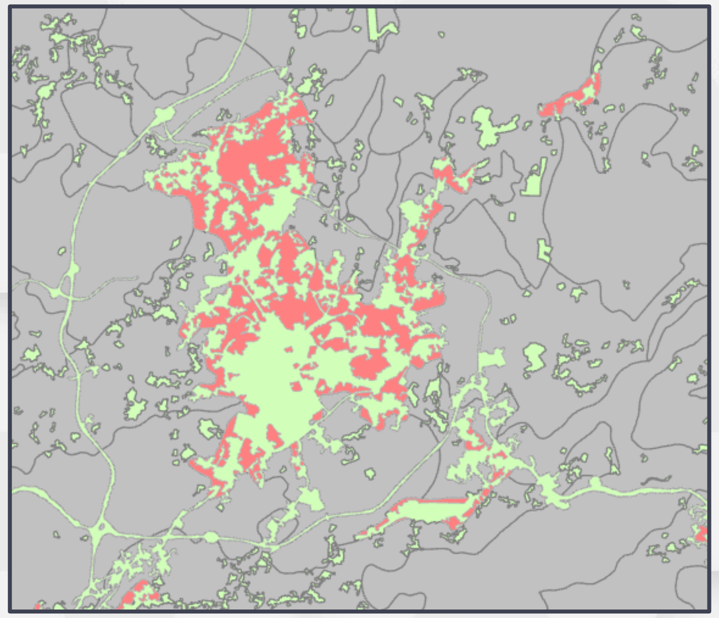
2 - Soil Map of Northeast Portugal;

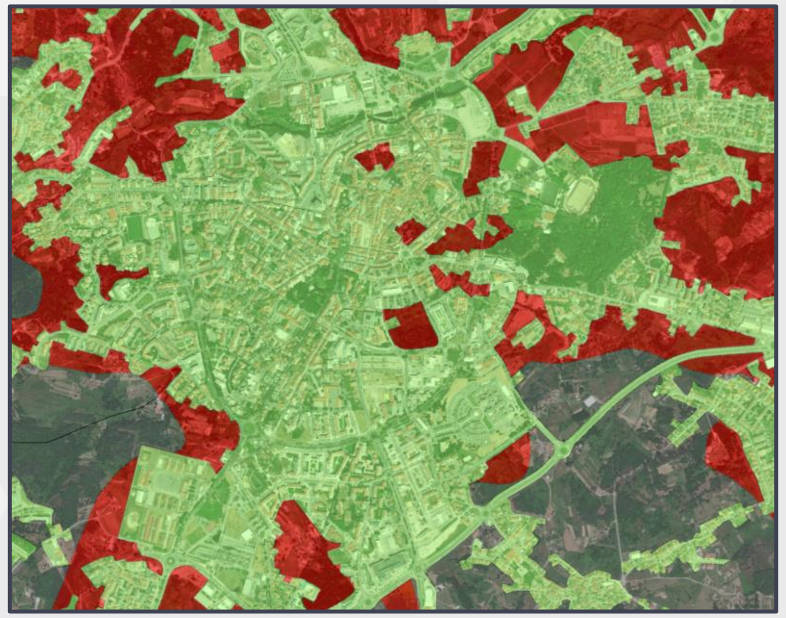
3 - Soil Map of the Central Interior Zone.



In the new Soil Map for mainland Portugal (1:100,000), there will be a **distinction between watercourses and artificialized areas**, as in the previous maps they were both grouped under **Social Area**.



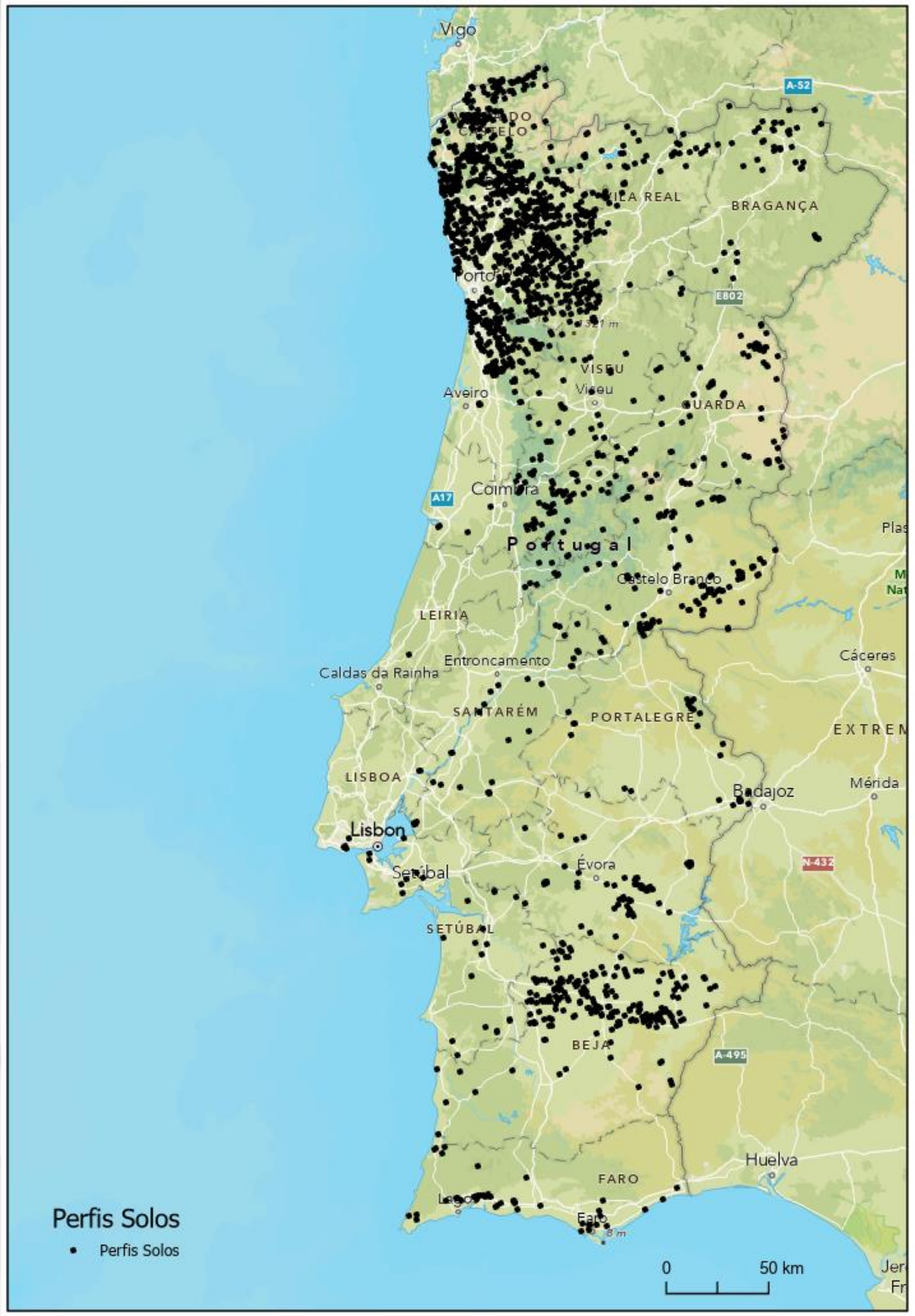




Displacement of watercourses

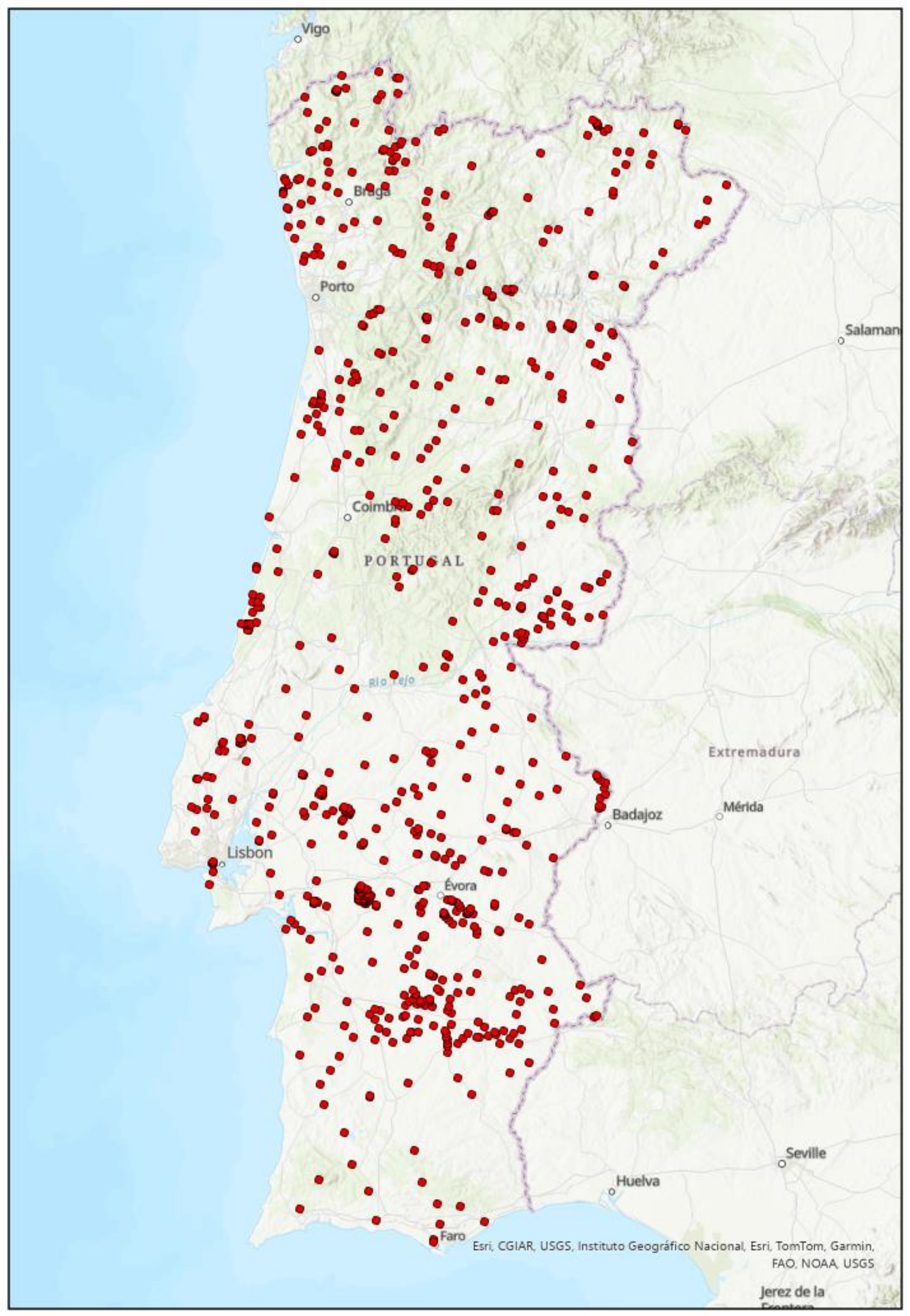
Correction

4 th MOPT CONFERENCE



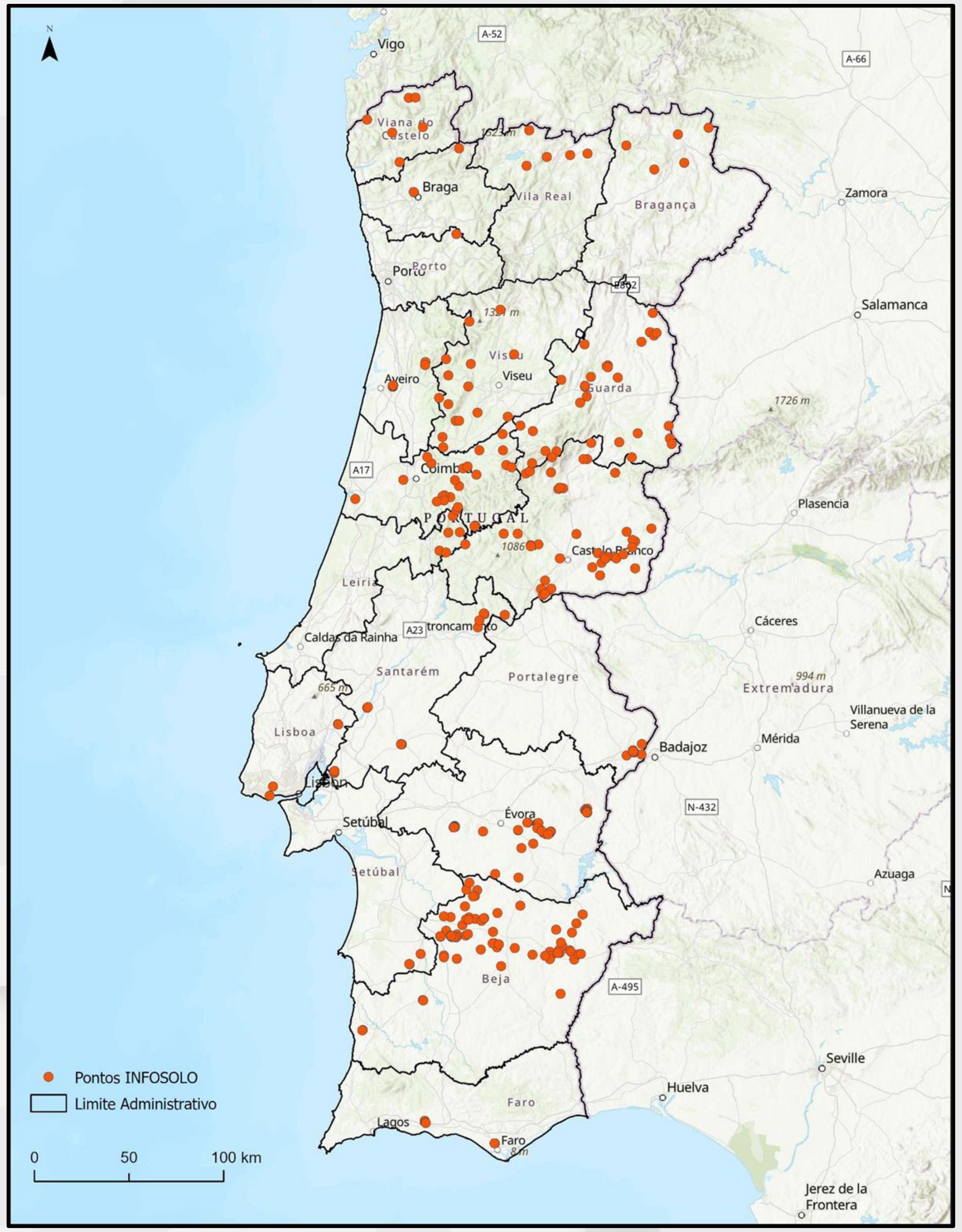
Field: Add Calculate Selection: Select By Attributes Zoom To Switch Clear Delete Copy

	FID	Shape *	Perfil	Regiao	Long_WGS4	Lat_WGS84	X_ETRS89	Y_ETRS89	LONG	LAT
1	0	Point	PT9A	ALT	-7.951236	37.045119	16179.187567	-291161.55957	7° 57' 4.450" W	37° 2' 42.430" N
2	1	Point	PT8	ALT	-7.951869	37.045933	16122.674627	-291071.342835	7° 57' 6.730" W	37° 2' 45.360" N
3	2	Point	SM16	ALT	-7.910506	37.113367	19784.803607	-283579.821669	7° 54' 37.820" W	37° 6' 48.120" N
4	3	Point	SM17	ALT	-7.664197	37.122428	41671.675419	-282494.497079	7° 39' 51.110" W	37° 7' 20.740" N
5	4	Point	SM18	ALT	-8.269447	37.129092	-12115.208623	-281849.152338	8° 16' 10.010" W	37° 7' 44.730" N
6	5	Point	TB 459	ALT	-8.697939	37.182092	-50156.588358	-275826.469754	8° 41' 52.580" W	37° 10' 55.530" N
7	6	Point	BS407	ALT	-7.799444	37.193056	29624.769651	-274706.95328	7° 47' 58.000" W	37° 11' 35.000" N
8	7	Point	TB 461	ALT	-7.928553	37.204547	18158.953015	-273464.136998	7° 55' 42.790" W	37° 12' 16.370" N
9	8	Point	SM14	ALT	-8.311031	37.262444	-15782.56418	-267043.34944	8° 18' 39.710" W	37° 15' 44.800" N
10	9	Point	BS346	ALT	-8.015556	37.300833	10422.204212	-262791.196075	8° 0' 56.000" W	37° 18' 3.000" N
11	10	Point	SM19	ALT	-8.594833	37.313808	-40929.517603	-261257.6943	8° 35' 41.400" W	37° 18' 49.710" N
12	11	Point	BS338	ALT	-8.408333	37.379167	-24376.080169	-254068.383662	8° 24' 30.000" W	37° 22' 45.000" N
13	12	Point	PT22	ALT	-7.950758	37.033436	16224.17598	-292458.077393	7° 57' 2.730" W	37° 2' 0.370" N
14	13	Point	BS340	ALT	-8.054444	37.440833	6961.374612	-247256.925148	8° 3' 16.000" W	37° 26' 27.000" N
15	14	Point	BS330	ALT	-8.452222	37.519167	-28210.611263	-238518.029454	8° 27' 8.000" W	37° 31' 9.000" N
16	15	Point	TB 460	ALT	-8.566428	37.646089	-38241.784302	-224390.631122	8° 33' 59.140" W	37° 38' 45.920" N
17	16	Point	SM15	ALT	-8.307139	37.678589	-15352.010034	-220857.527245	8° 18' 25.700" W	37° 40' 42.920" N
18	17	Point	BS323	ALT	-8.308056	37.688056	-15430.912148	-219806.665473	8° 18' 29.000" W	37° 41' 17.000" N
19	18	Point	SM13	ALT	-7.735011	37.692006	35111.661298	-219308.074116	7° 44' 6.040" W	37° 41' 31.220" N
20	19	Point	AG2	ALT	-8.584444	37.736389	-39783.54743	-214360.599224	8° 35' 4.000" W	37° 44' 11.000" N
21	20	Point	TB 456	ALT	-8.090956	37.742928	3715.269245	-213729.90336	8° 5' 27.440" W	37° 44' 34.540" N



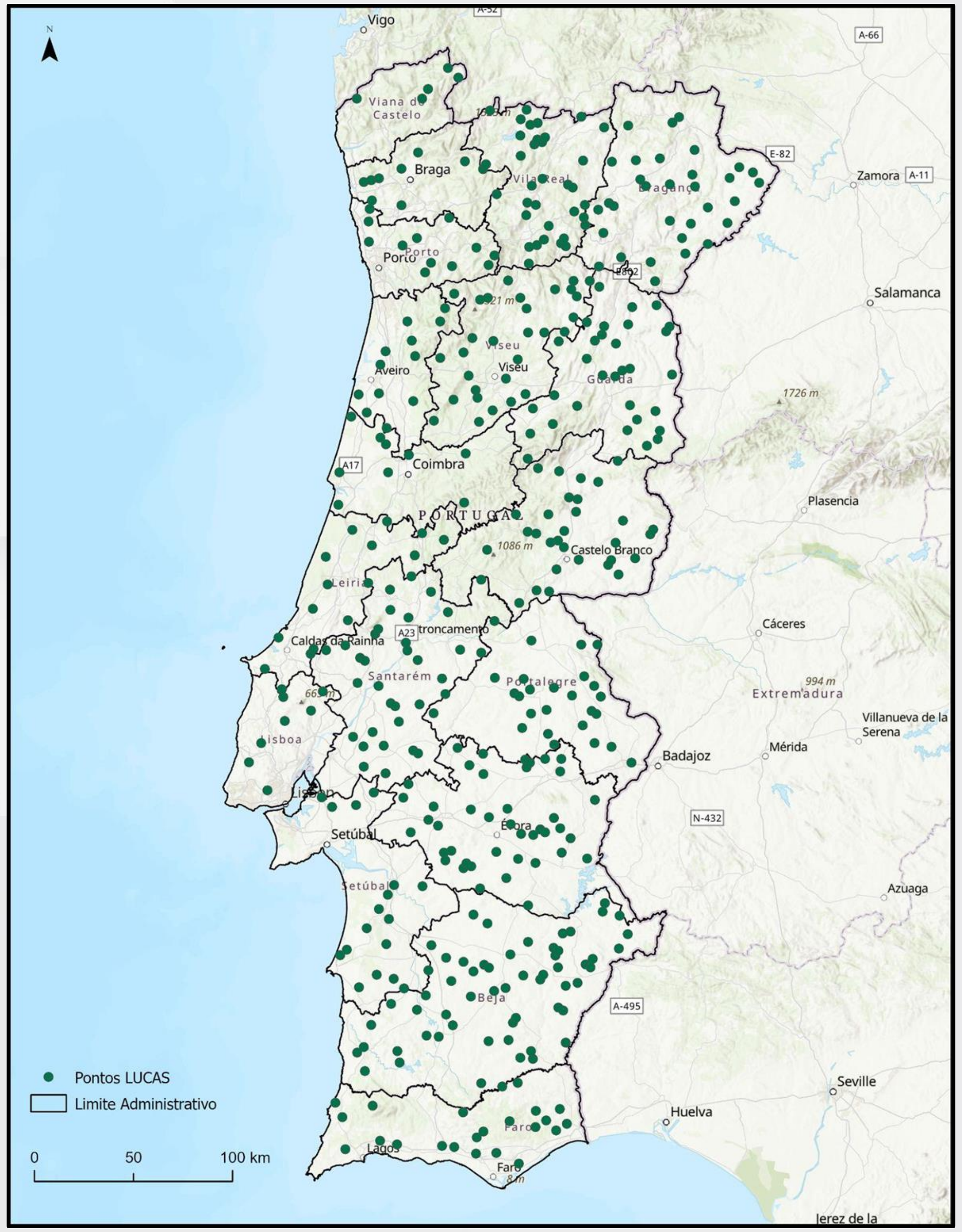
4 th MOPT CONFERENCE

INFOSOLO Points (309)

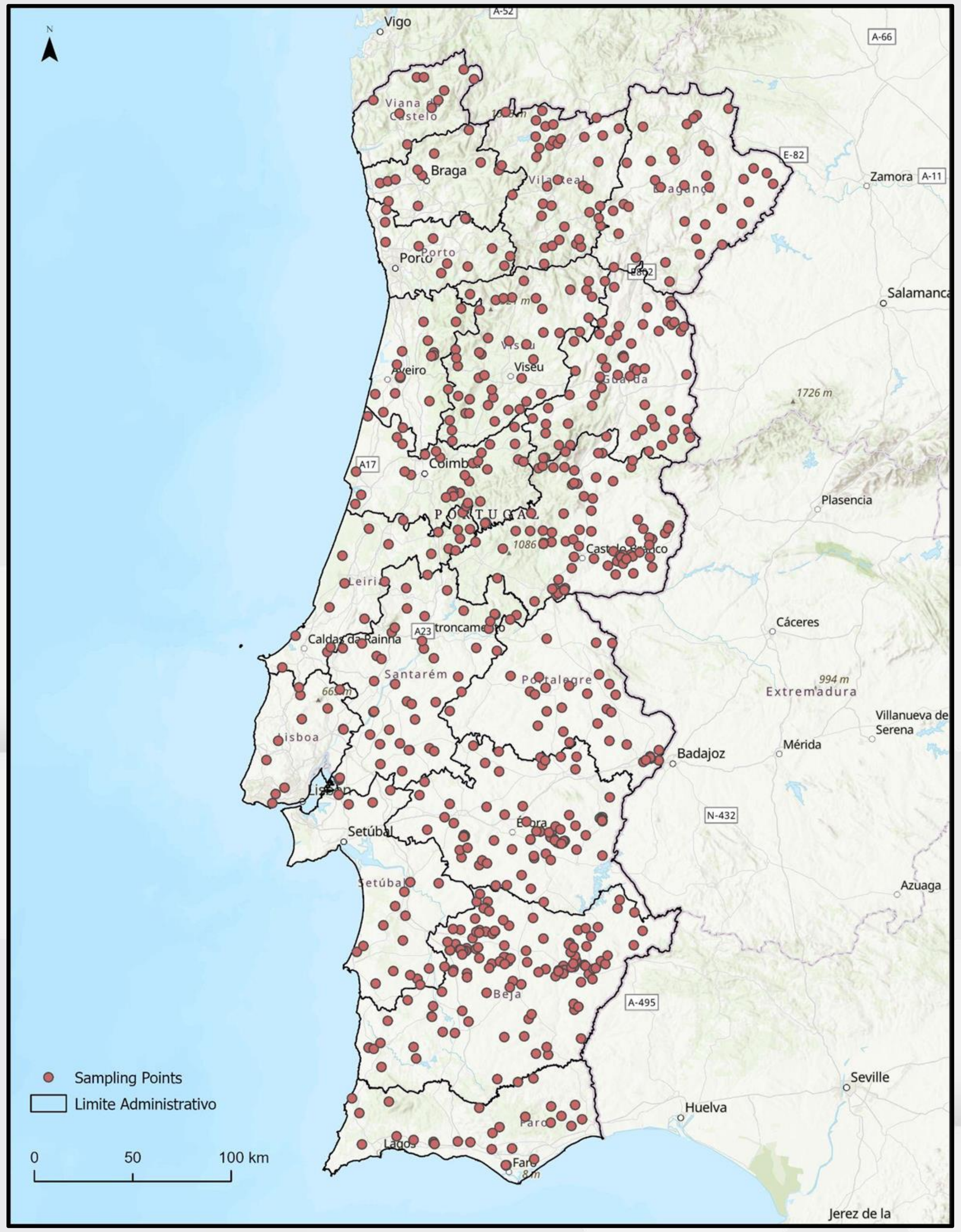


+

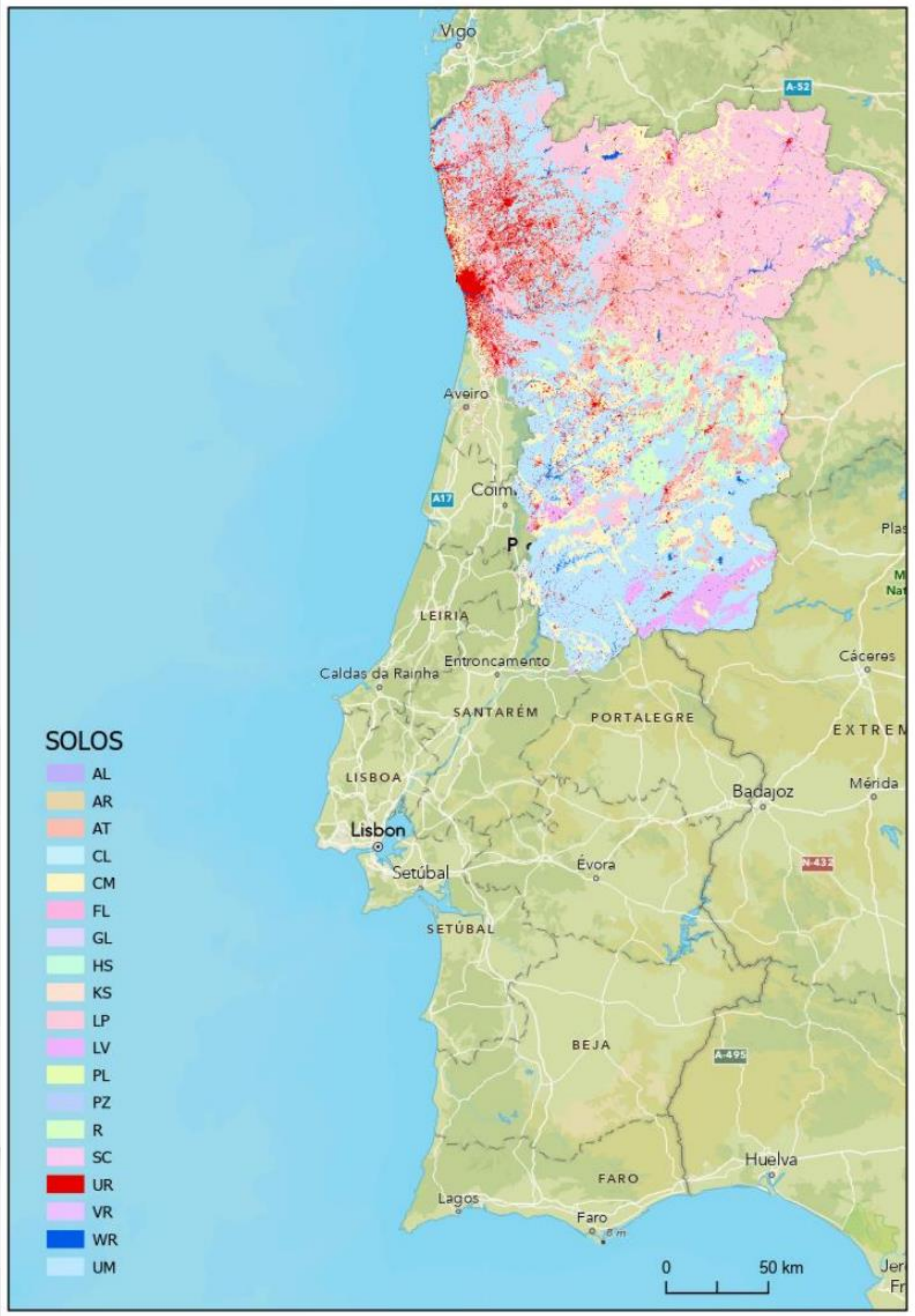
LUCAS Points (428)



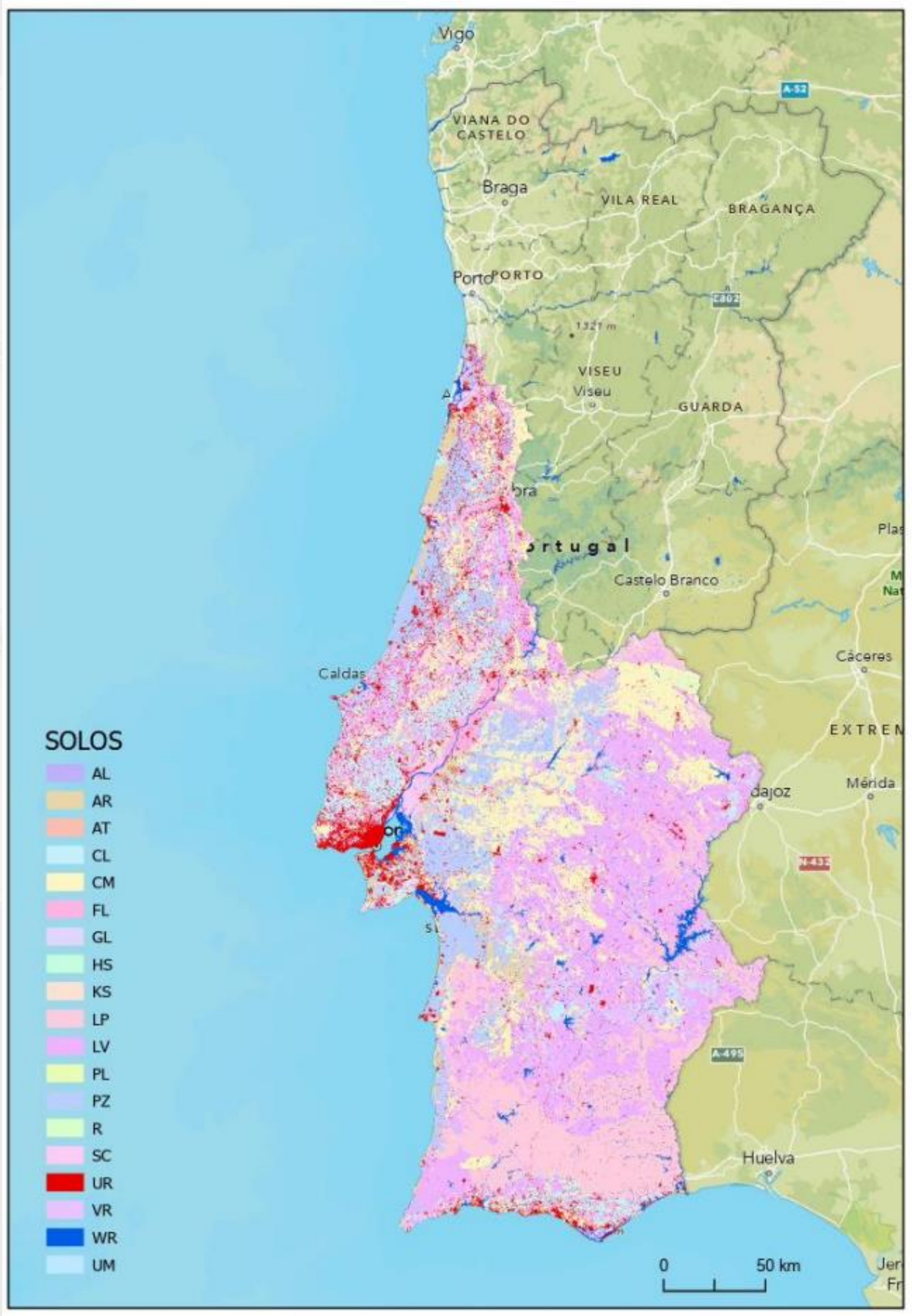
737 Sampling Points

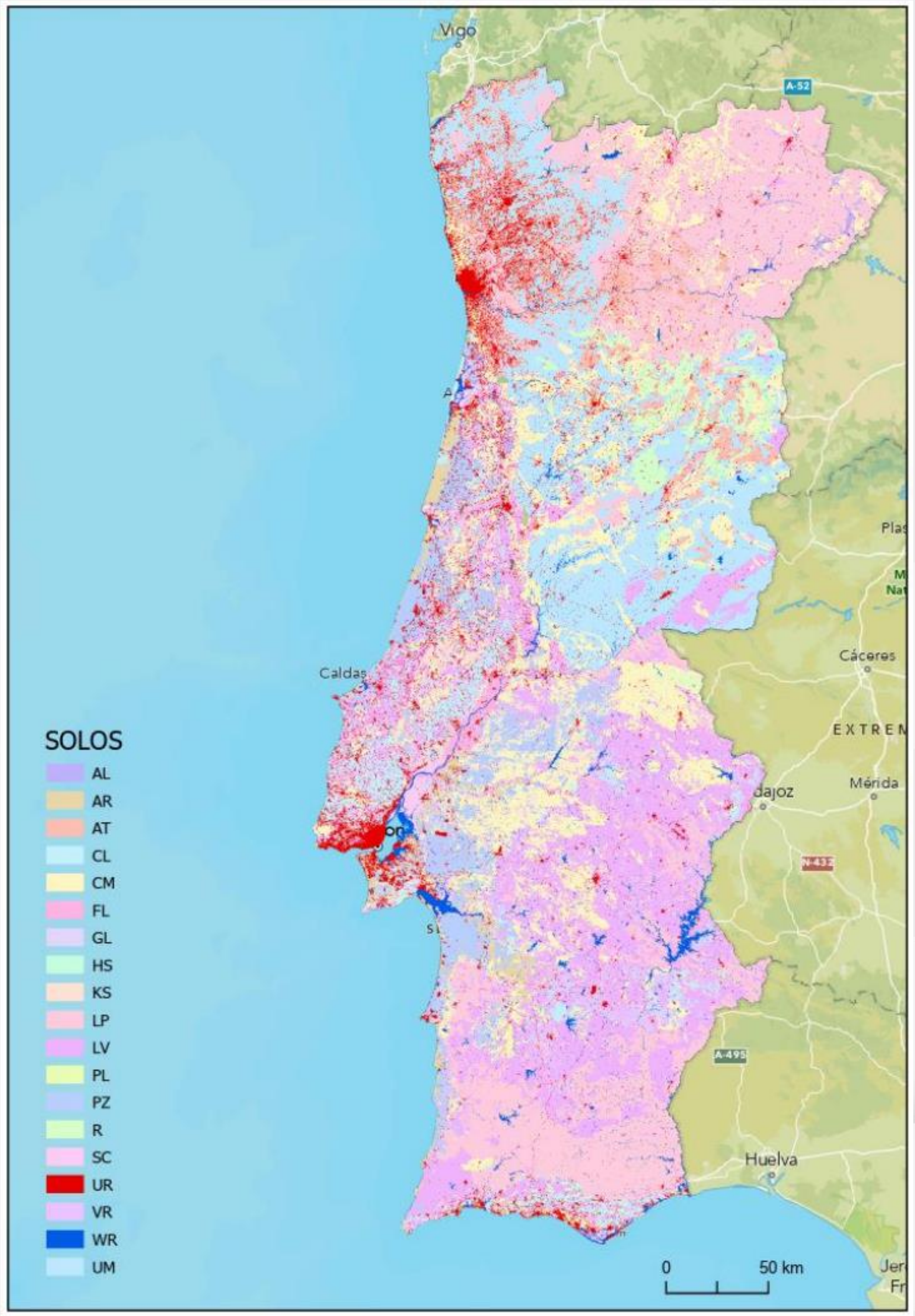


4th MOPT CONFERENCE



+





4th MOPT CONFERENCE

SOIL CHEMICAL PROPERTIES

Variable Name	Rescaling function	Rationale	Dataset Source	References
Soil pH	Near function with midpoint value	Higher values = values of pH of 6.4 and those closer	SOILGRIDS	
Soil Organic Carbon	Linear	Higher values = higher percentage organic carbon in soil		Gardi et al. , 2016
Carbon-Nitrogen Ratio	Near function with midpoint value	Higher values = values of Carbon-Nitrogen of 24:1 and those closer		Brust, 2019;
Cation Exchange Capacity	Linear	Higher values = higher values of CEC = good capacity		Chowdhury et al., 2021

SOIL PHYSICAL PROPRITIES

Variable Name	Rescaling function	Rational	Dataset Source	References
Soil Texture	Proportion of clay, sand and silt USDA soil texture classes	Higher Values = Loamy Soils	SOILGRIDS	Seation et al (2020); Yang et al (2023)
Bulk density	Small with midpoint value	Higher values = values of bulk density lower than 133 cg/cm3 (the used optimal value is for a medium textured soil)		Lehmann et al (2020); Montgomery & Biklé (2021);

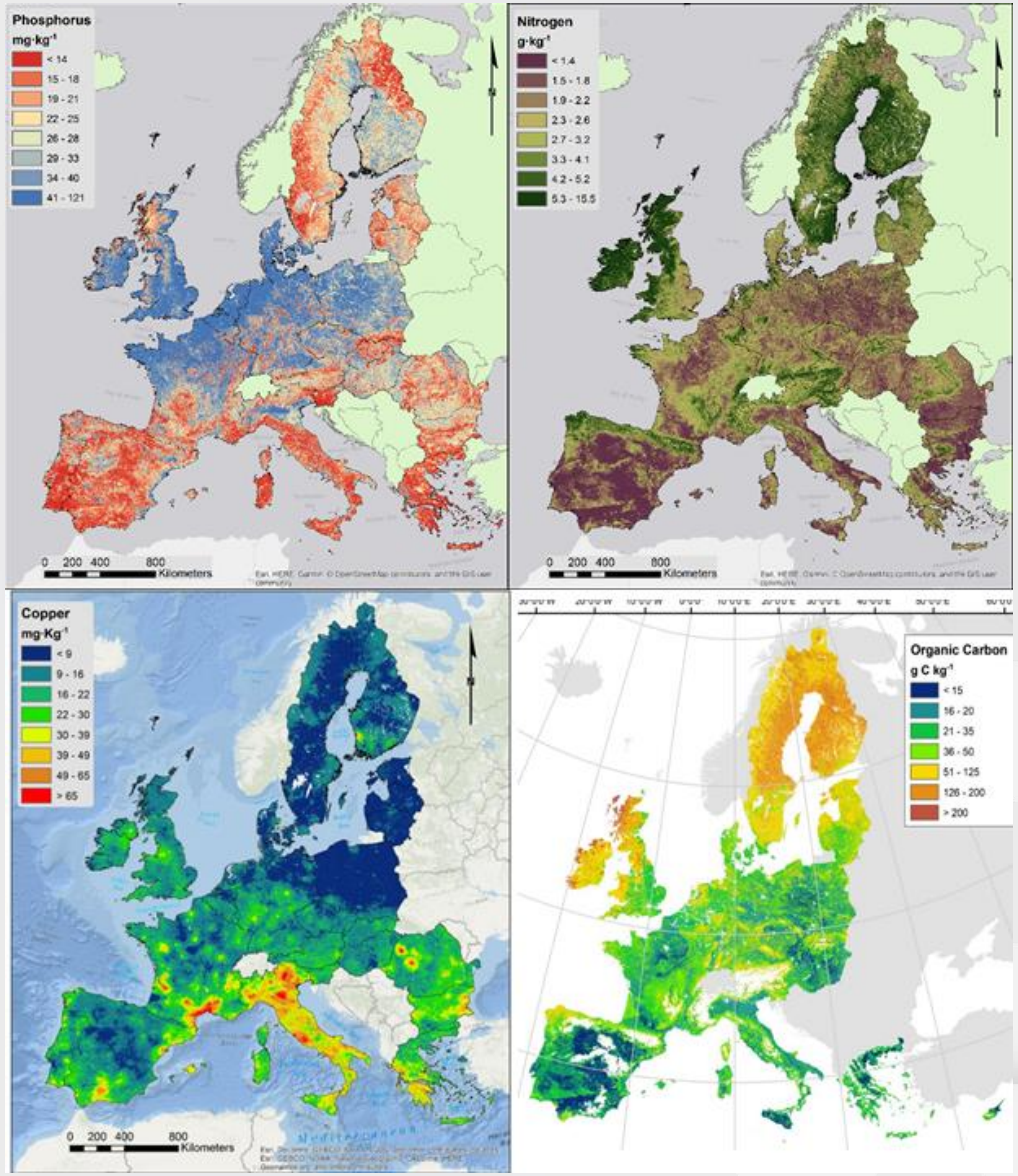
SOIL HEAVY METALS

Variable Name	Normalization function	Rationale	Dataset Source	References
As Cd Cr Cu Hg Pb Mn Sb Co Ni	Small with midpoint value (established thresholds)	Higher values = values lower than the established thresholds	EEA	Vácha (2021); Rashid et al (2021); Quinton and Catt (2007); Hu et al (2018); Kelepertzis (2014)

4th MOPT CONFERENCE

- Coarse fragments
- particle-size distribution (clay, silt, sand)
- pH
- Organic carbon
- Carbonate content
- Total nitrogen content
- Extractable potassium content
- Phosphorous content
- Cation exchange capacity

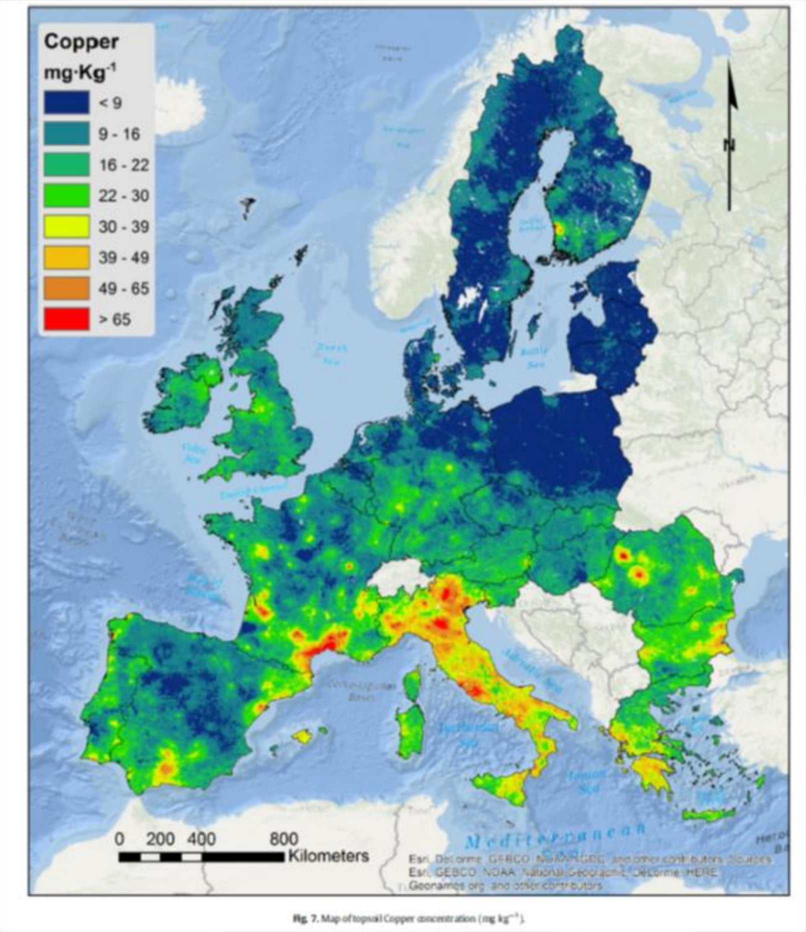
- Soil Biodiversity



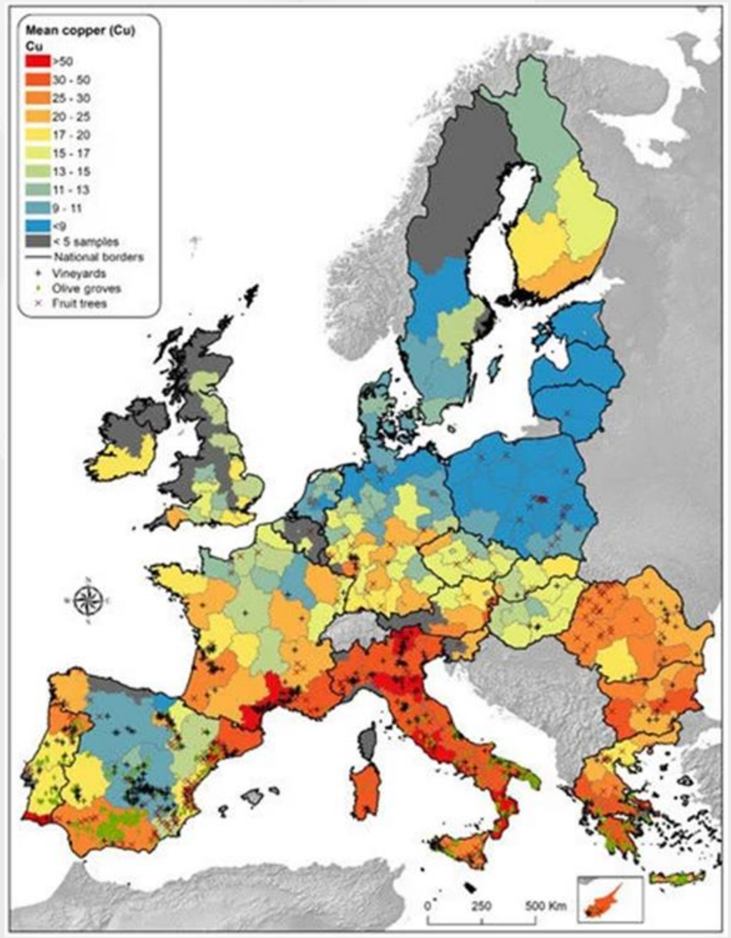
4 th MOPT CONFERENCE

Copper distribution in European topsoils: An assessment based on LUCAS soil survey

- Cu is correlated to soil properties (pH, texture, OC), climate, geology and management.
- Vineyards (49.3 mg kg^{-1}), olive groves (33.5 mg kg^{-1}) and orchards (27.3 mg kg^{-1}) show high [Cu] that may be affected by the application of Cu-based fungicides for controlling plant diseases



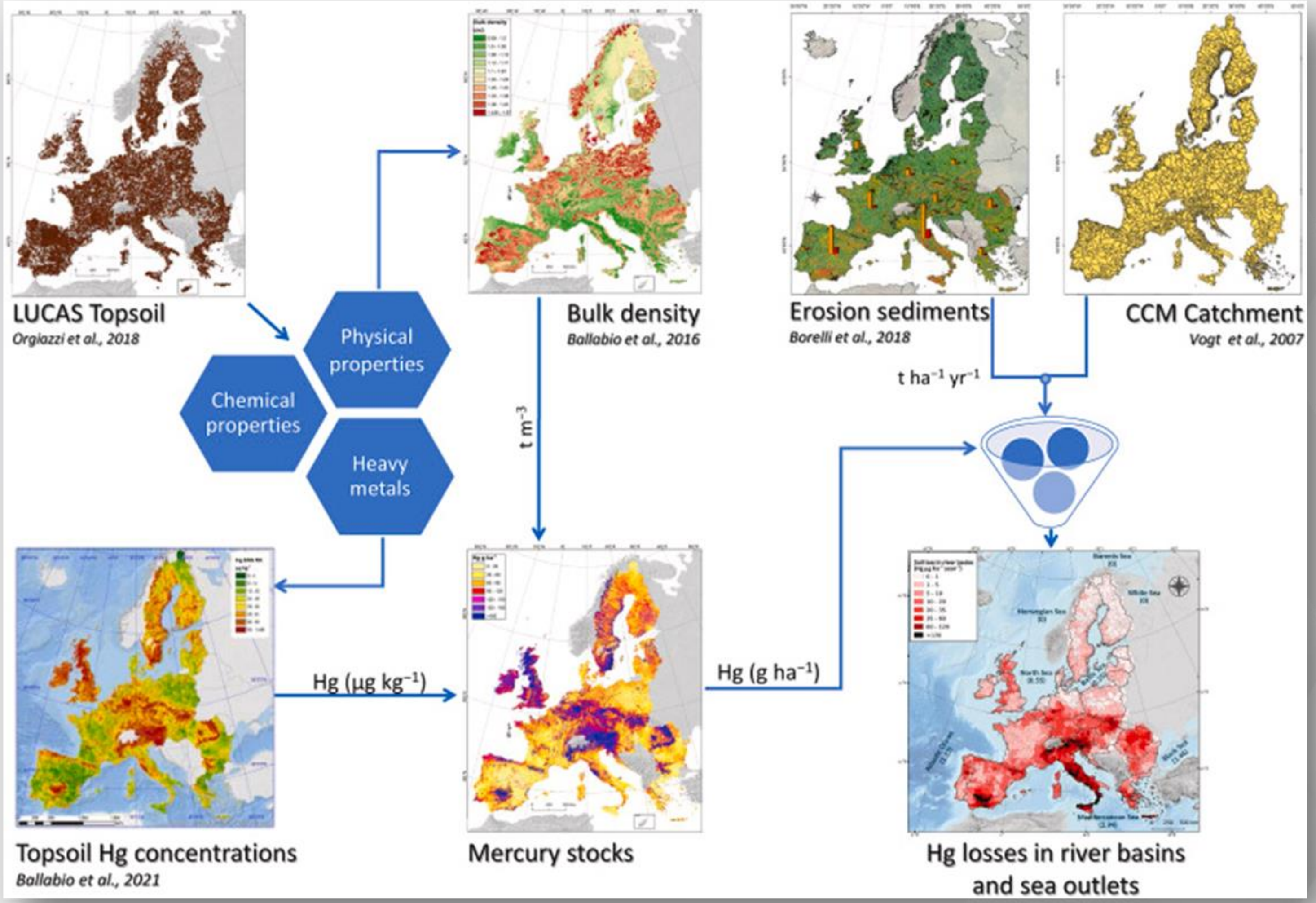
Ballabio et al. (2018) https://doi.org/10.1016/j.scitotenv.2018.04.268.



Panagos et al.(2018) https://doi.org/10.3390/su10072380

Mercury in European topsoils: Anthropogenic sources, stocks and fluxes

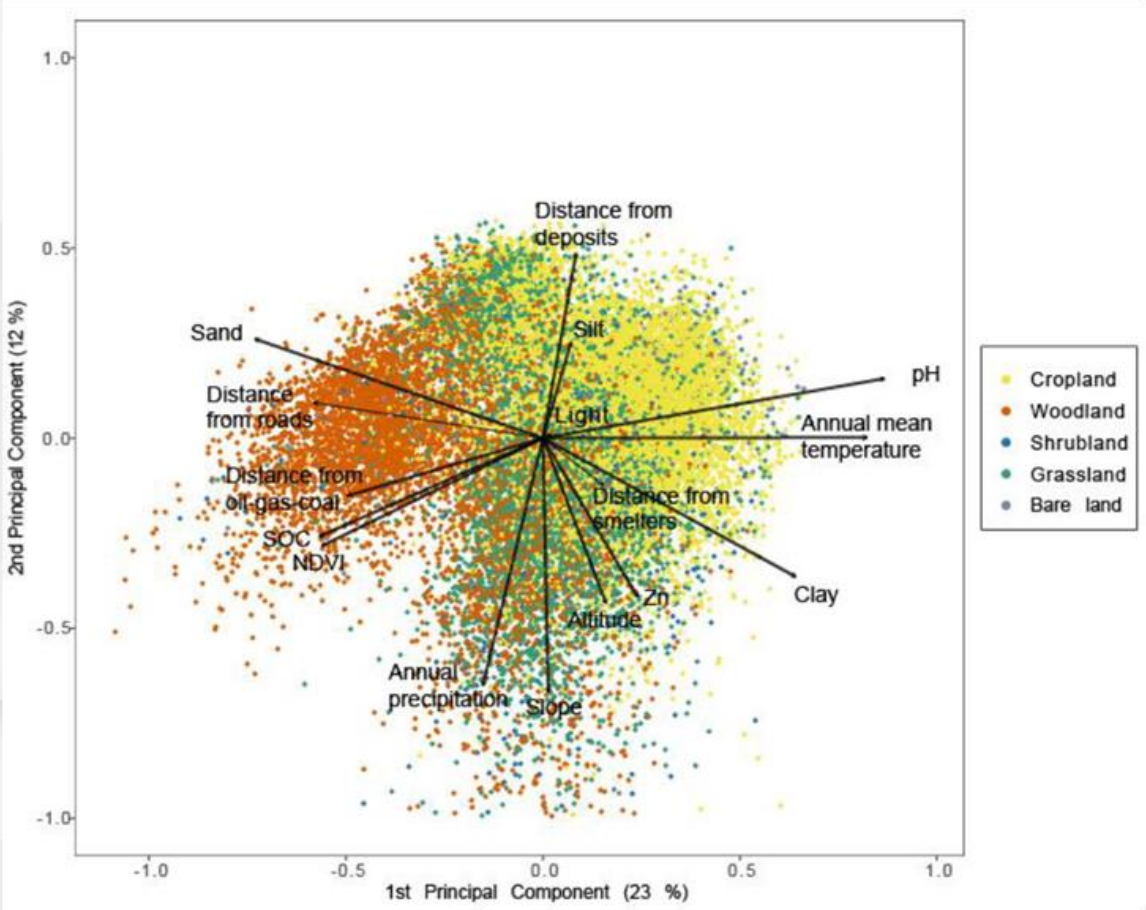
In the European Union and UK, about 43 Mg Hg yr^{-1} are displaced by water erosion and 6 Mg Hg yr^{-1} are transferred to river basins and to coastal Oceans. Panagos et al.(2021)., Environmental Research, 201, (111556) 0013-9351

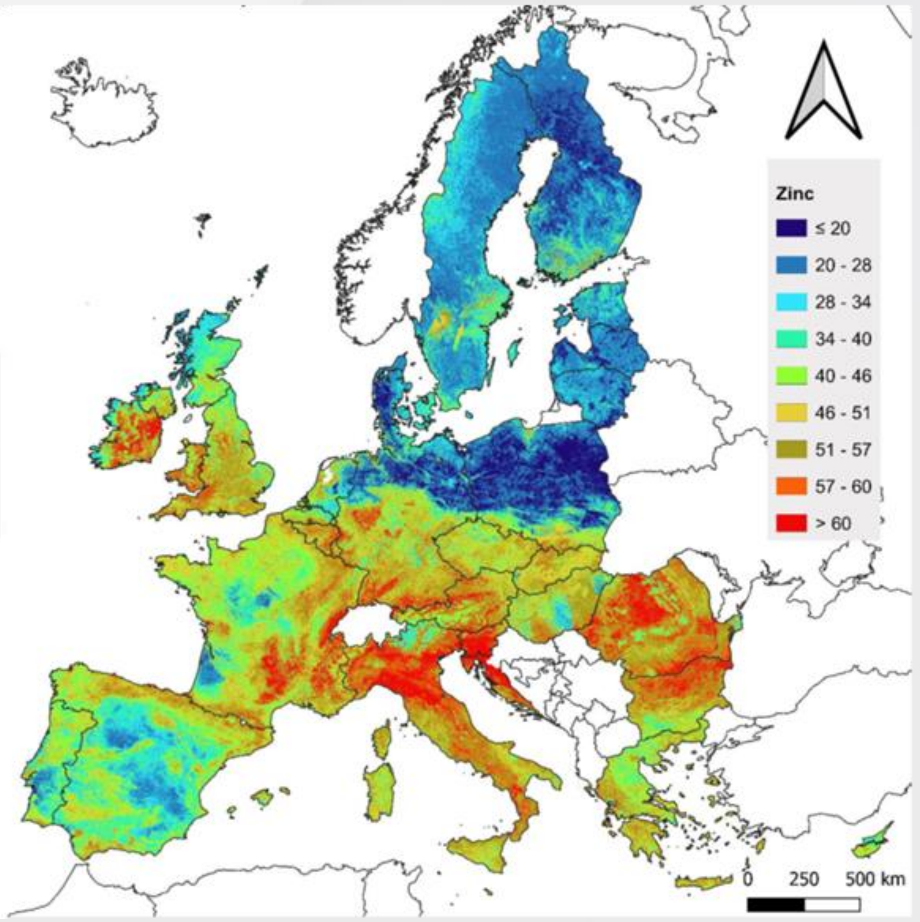


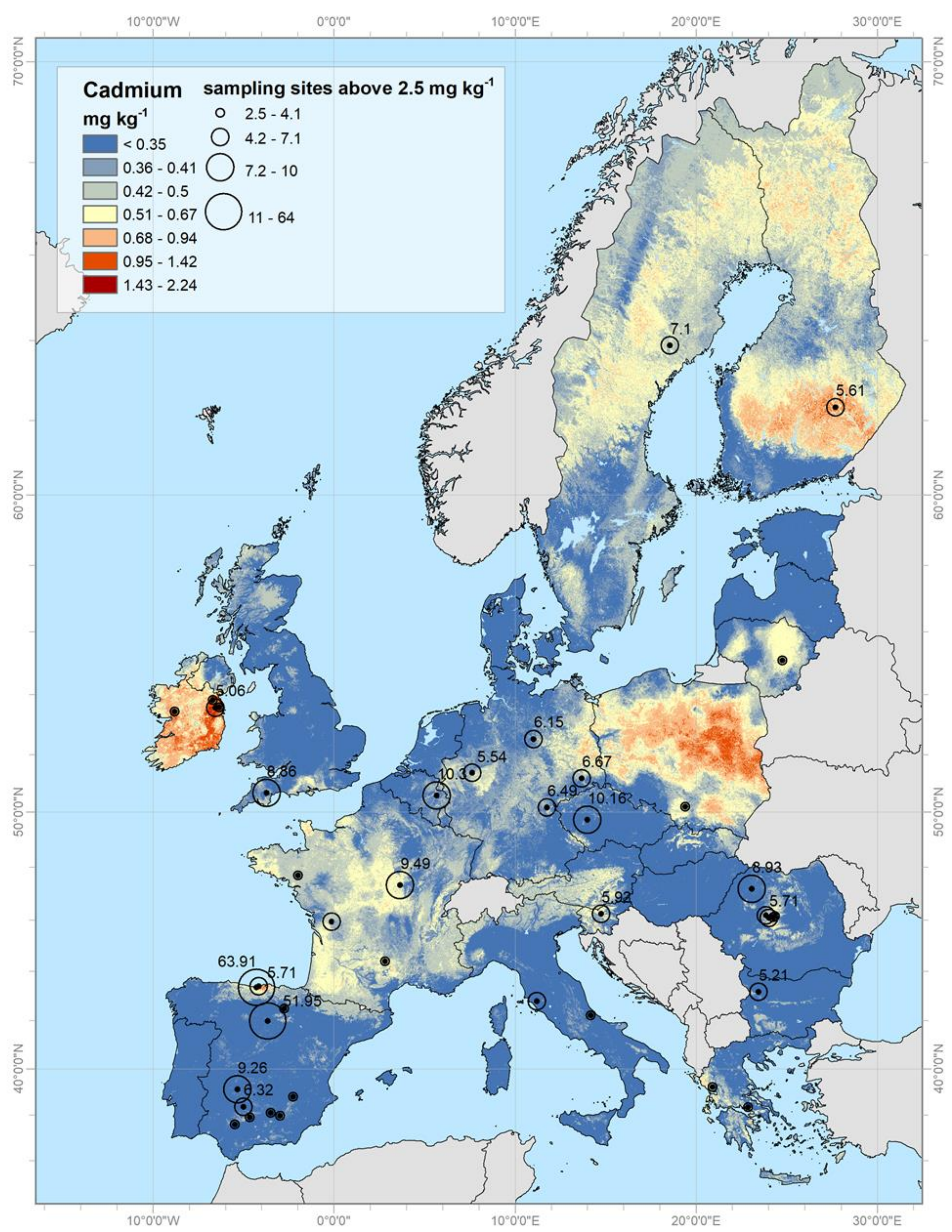
4 th MOPT CONFERENCE

Spatial assessment of topsoil zinc concentrations in Europe

Based on LUCAS topsoil database, the mean Zn concentration in Europe is 47 mg kg-1 and median Zn concentration is 40 mg kg-1. Ninety nine percent of all samples have concentrations below 167 mg kg-1. Soil texture and pH are most important drivers for the variation in topsoil Zn High Zn concentrations are found near Zn deposits, and in grasslands





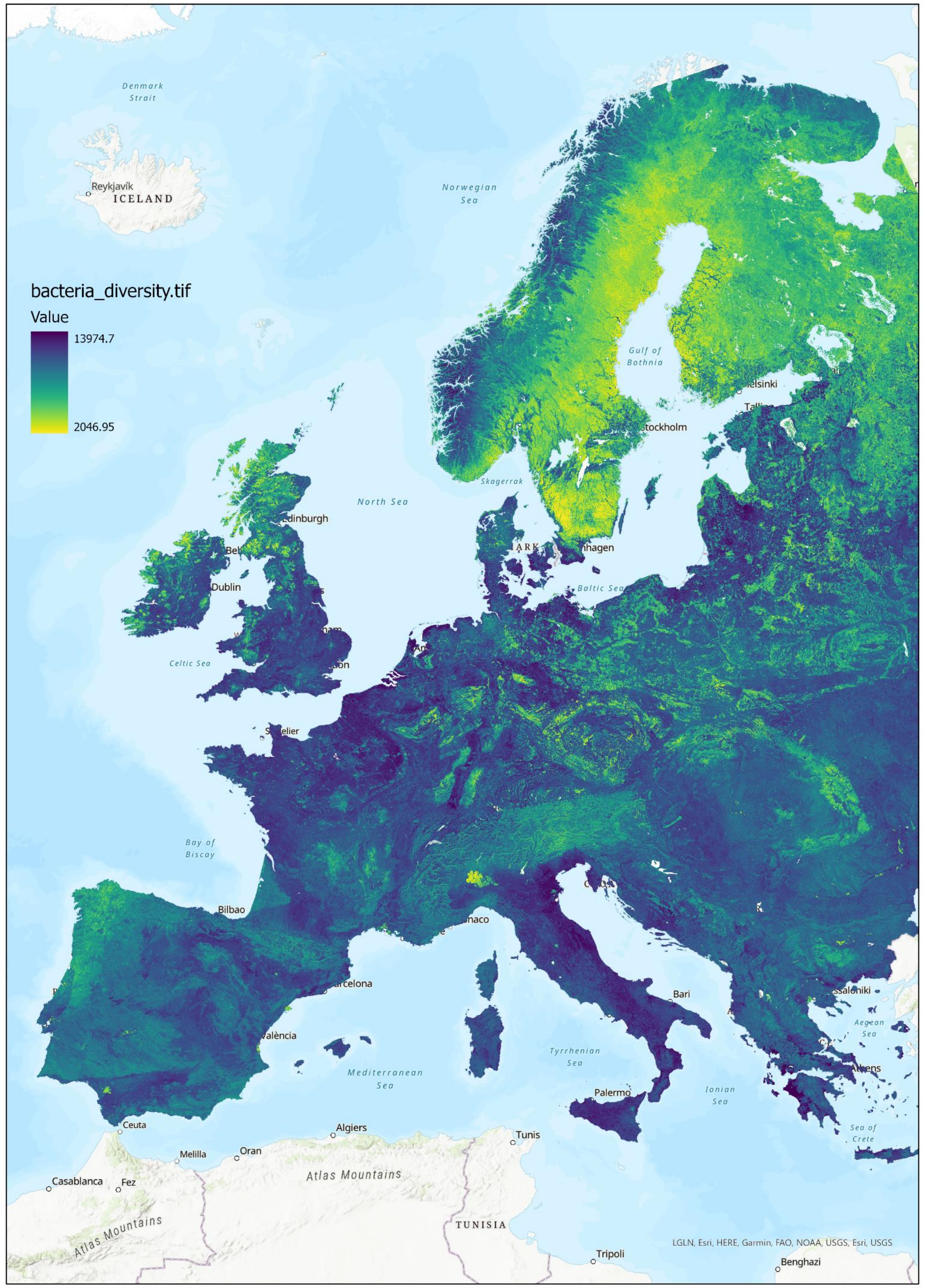


4th MOPT CONFERENCE

Study of the distribution of soil biodiversity

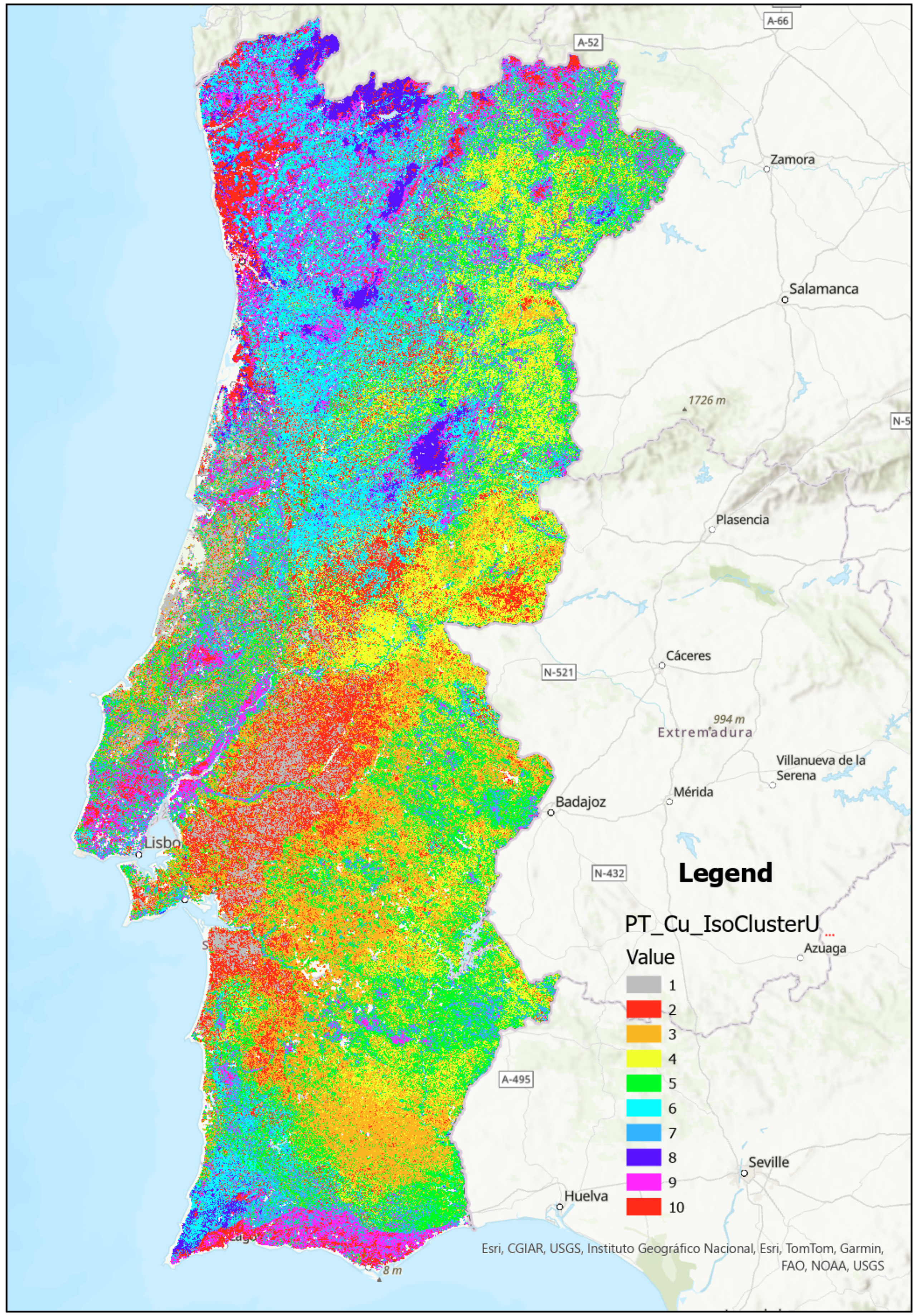
Using EO data (Sentinel 1 and 2, Landsat, etc.), climate (Worldclim) and other data, map the diversity of bacteria and fungi across the EU

High spatial resolution (100m) – reasonable accuracy (0.4 R2)???



Cluster ISODATA 10 clusters

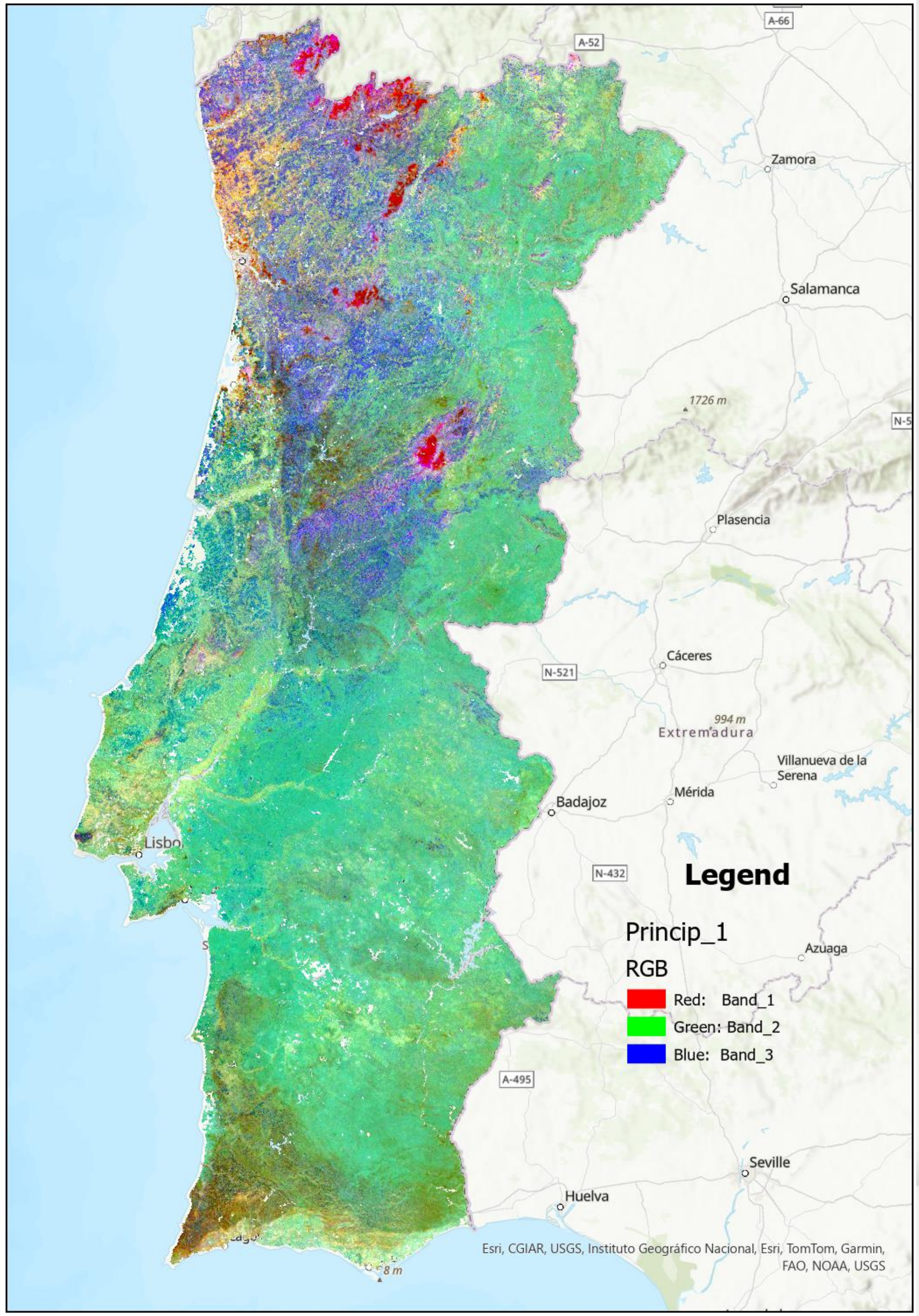
File	Description
PT_BD010.tif	Bulk density in the 0-10 cm layer
PT_BD1020.tif	Bulk density in the 10-20 cm layer
PT_BD2030.tif	Bulk density in the 20-30 cm layer
PT_CEC.tif	Cation exchange capacity
PT_CF.tif	Coarse fragments
PT_CLC.tif	Corine Land Cover class
PT_N.tif	Nitrogen content
PT_NUTS0.tif	NUTS0 region
PT_NUTS1.tif	NUTS1 region
PT_NUTS2.tif	NUTS2 region
PT_OC.tif	Organic carbon content
PT_pH.tif	Soil pH
PT_P.tif	Phosphorus content
PT_TXT.tif	Soil texture
PT_RUSLE.tif	RUSLE soil erosion map
PT_Cu.tif	Soil copper concentration
PT_Cd.tif	Soil cadmium concentration
PT_Hg.tif	Soil mercury concentration
PT_Zn.tif	Soil zinc concentration
PT_CaCO3.tif	Soil calcium carbonate



4 th MOPT CONFERENCE

Principal Component Analysis

Component	Eighenvalue (%)	Eighenvalue
1	83.7351	83.7351
2	10.5124	94.2475
3	5.7525	100.0000



4th MOPT CONFERENCE

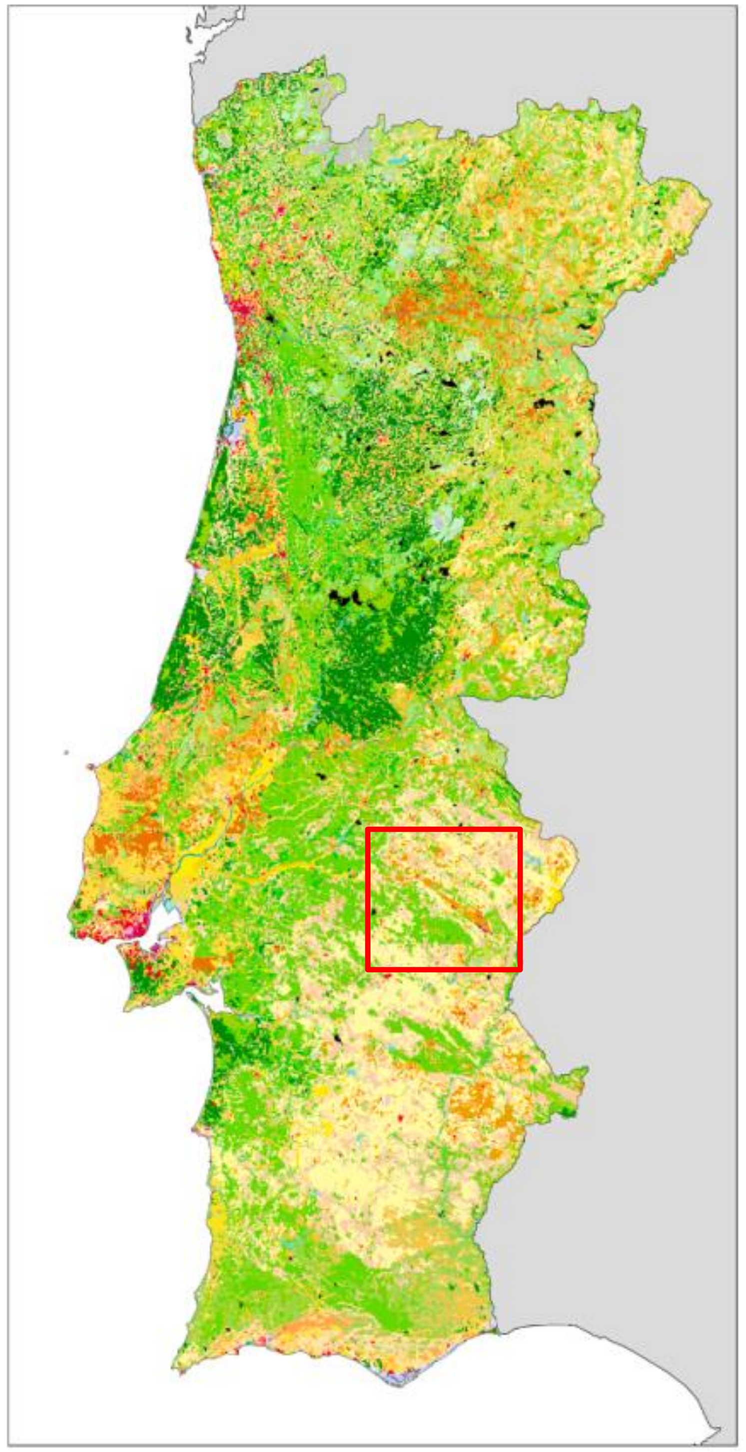
Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

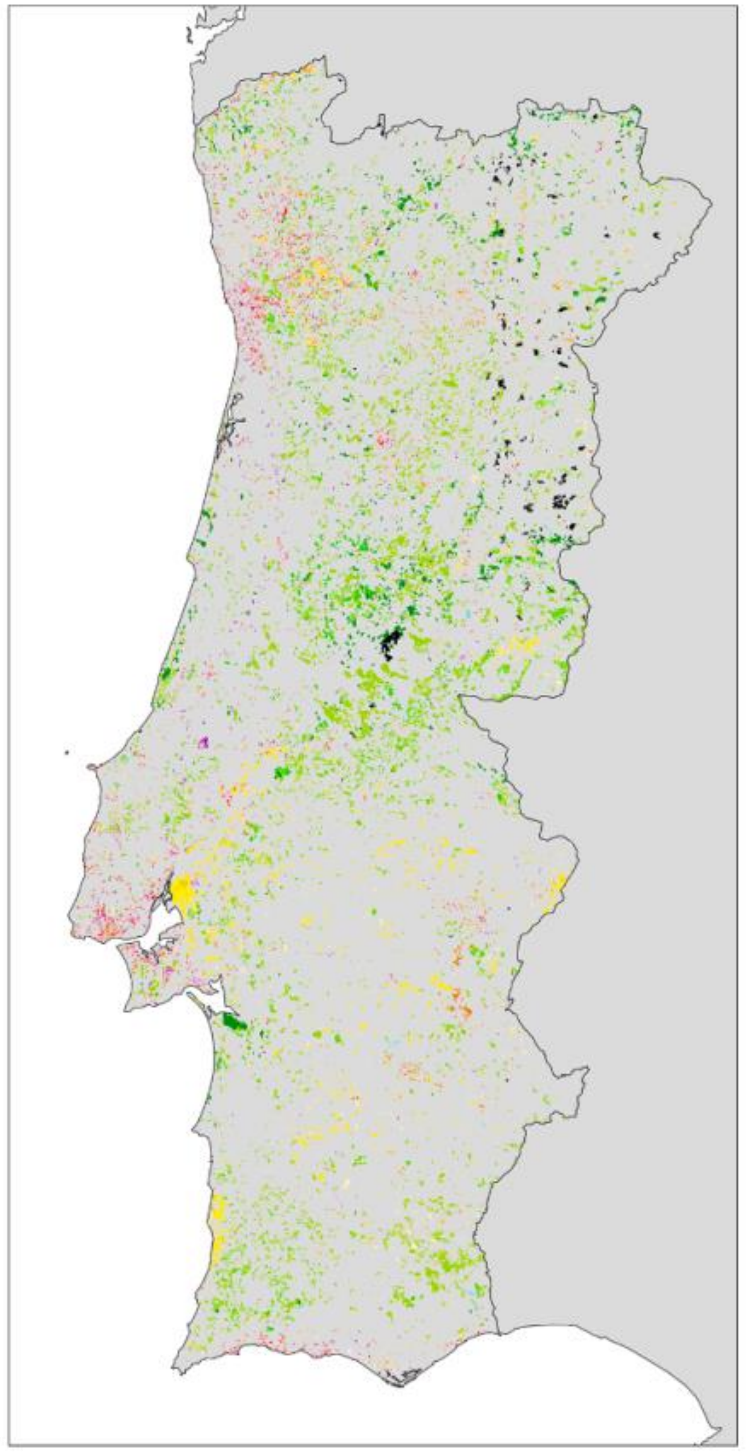
journal homepage: www.elsevier.com/locate/jag

Dealing with the uncertainty of technical changes in the CORINE Land Cover dataset: The Portuguese approach

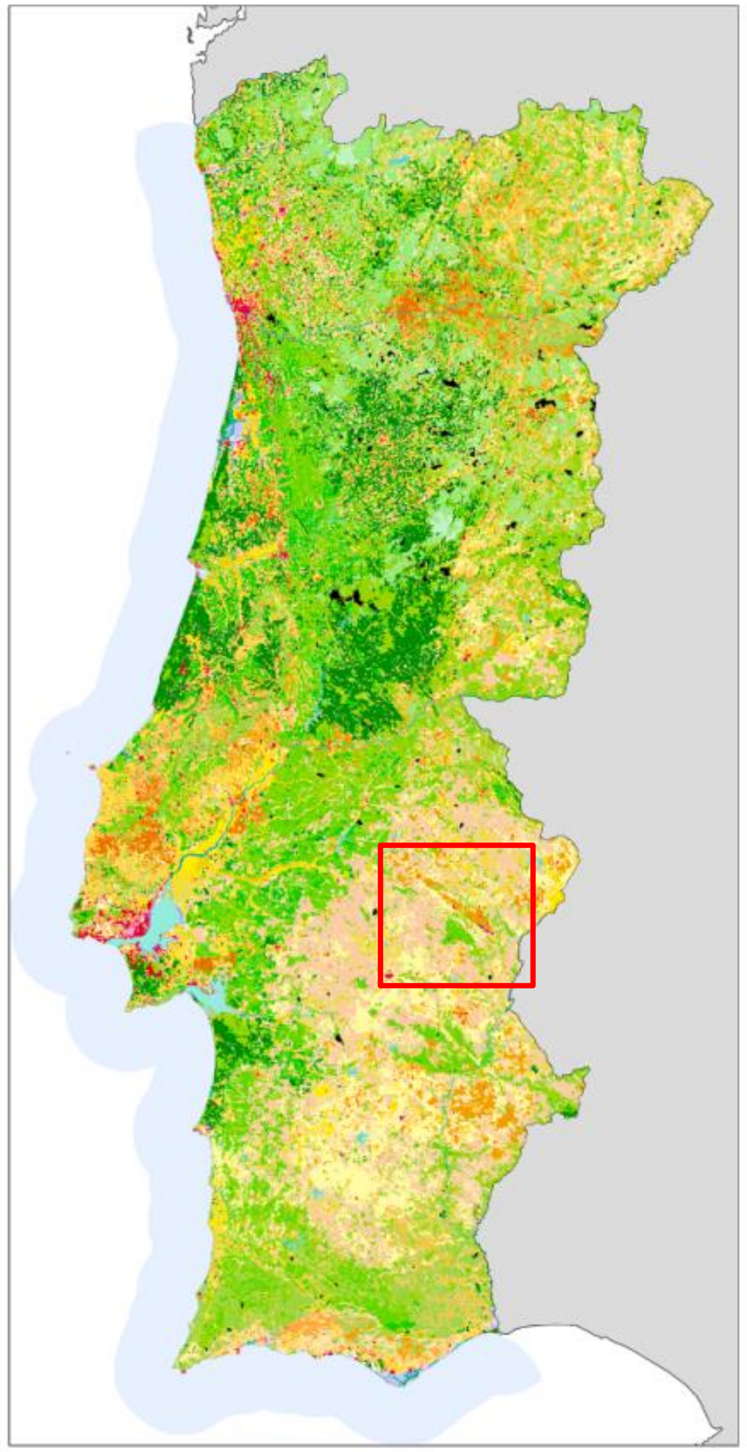
David García-Álvarez [a,*], Cláudia M. Viana [b,c], Eduardo Gomes [b,c], Filipe Marcelino [d], Mário Caetano [d,e], Jorge Rocha [b,c]



CORINE Status Layer (CSL) - 1990
Provided by
Copernicus Land Monitoring Service

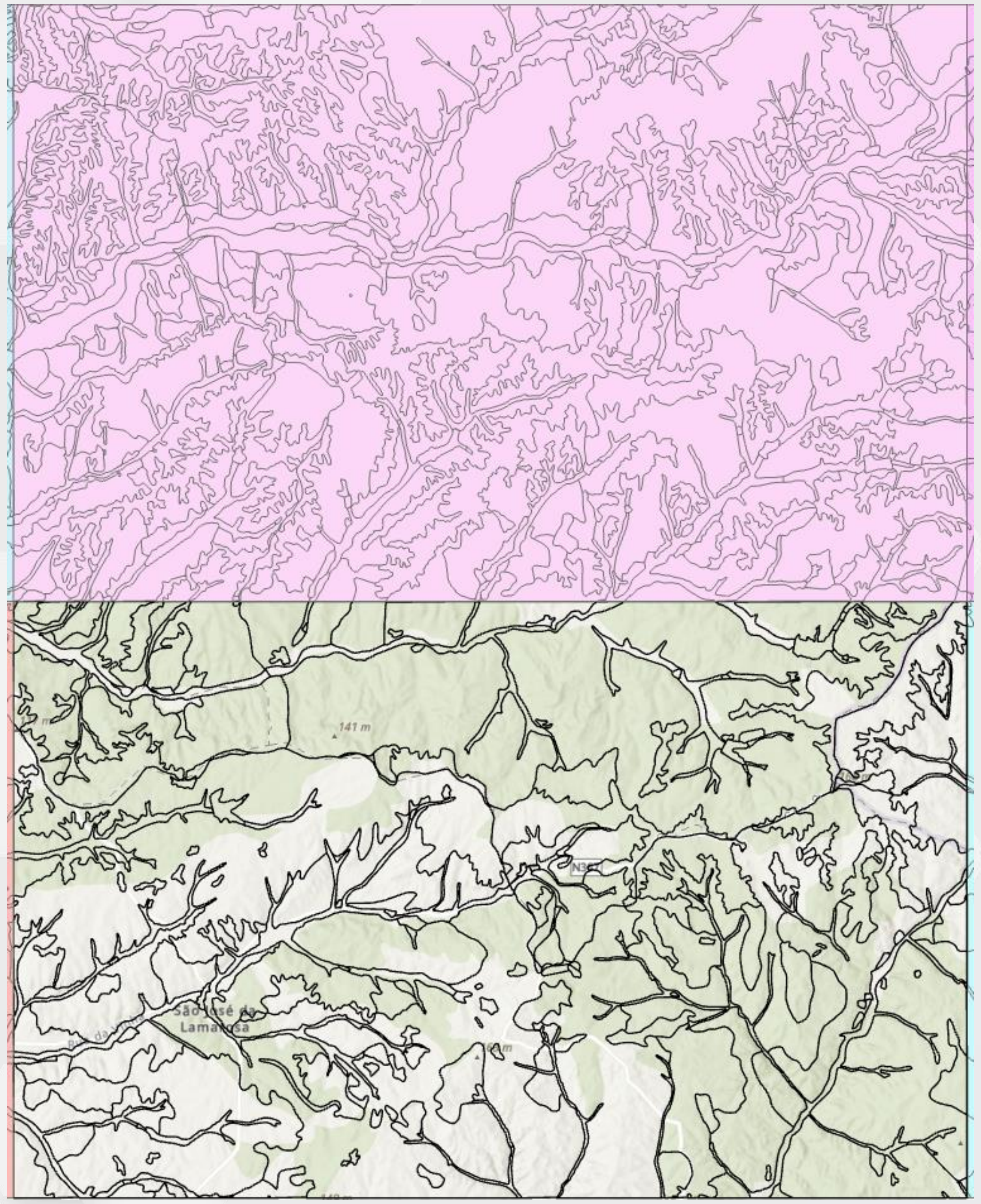


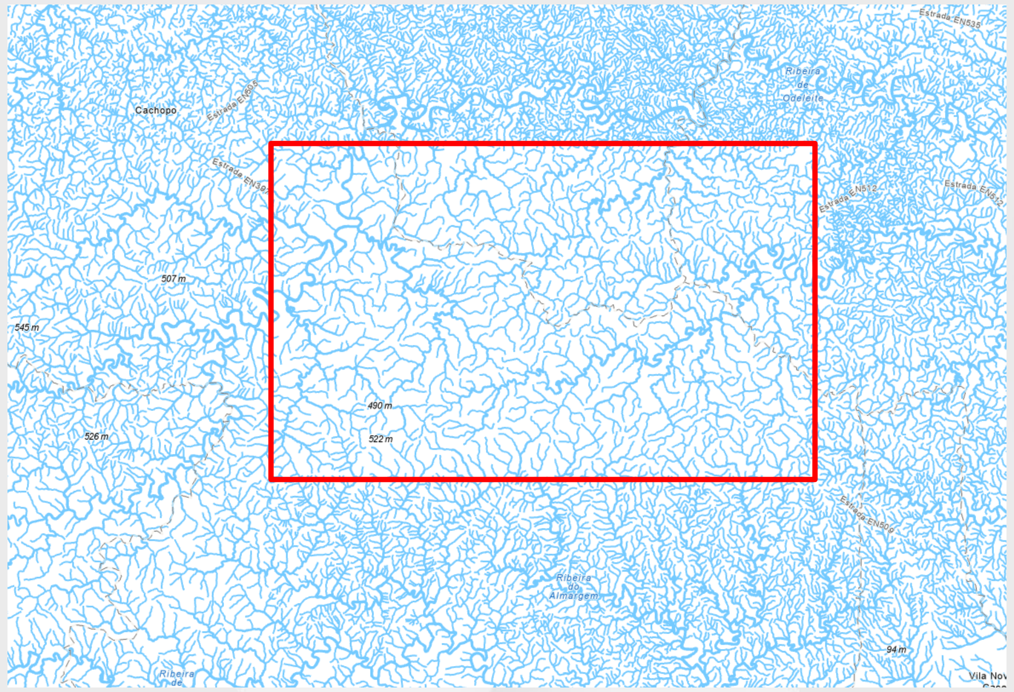
CORINE Layer of changes (CHA)
1990/2000

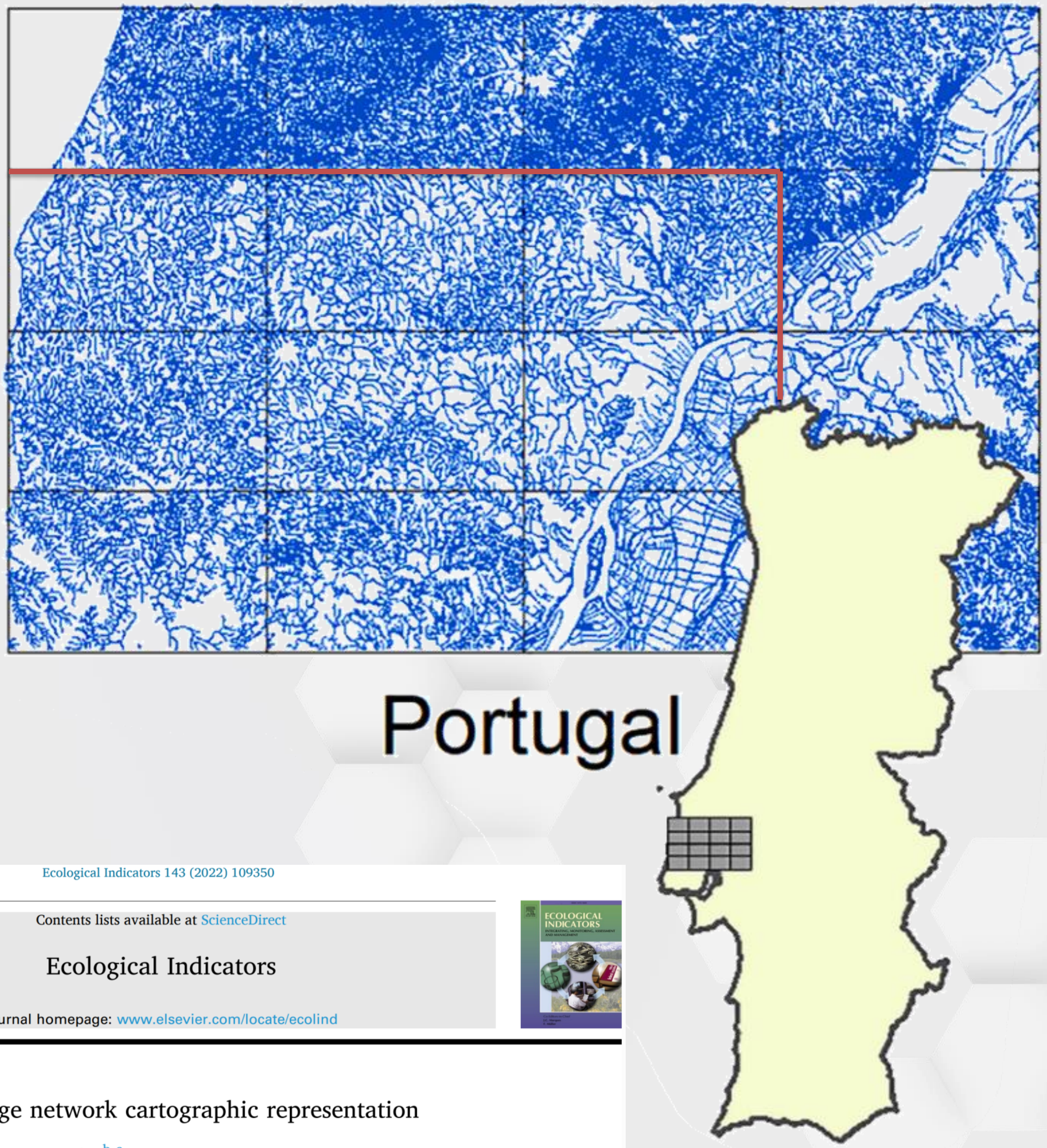


CORINE Status Layer (CSL) - 1990
Provided by
Direção-Geral do Território (DGT)

4 th MOPT CONFERENCE







Ecological Indicators 143 (2022) 109350

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

ELSEVIER

Original Articles

On the quality of the drainage network cartographic representation

Tony Vinicius Moreira Sampaio [a,*], Jorge Rocha [b,c]

To delineate Soil Districts:

- **Clustering Algorithms** (Tested):
  - Random Forest with MDS and Fuzzy k-means
  - Agglomerative Hierarchical Clustering
  - Fuzzy c-means
  - DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

To generate random sampling points:

- **Bethel Algorithm** developed with JRC to choose sampling locations

Variables:

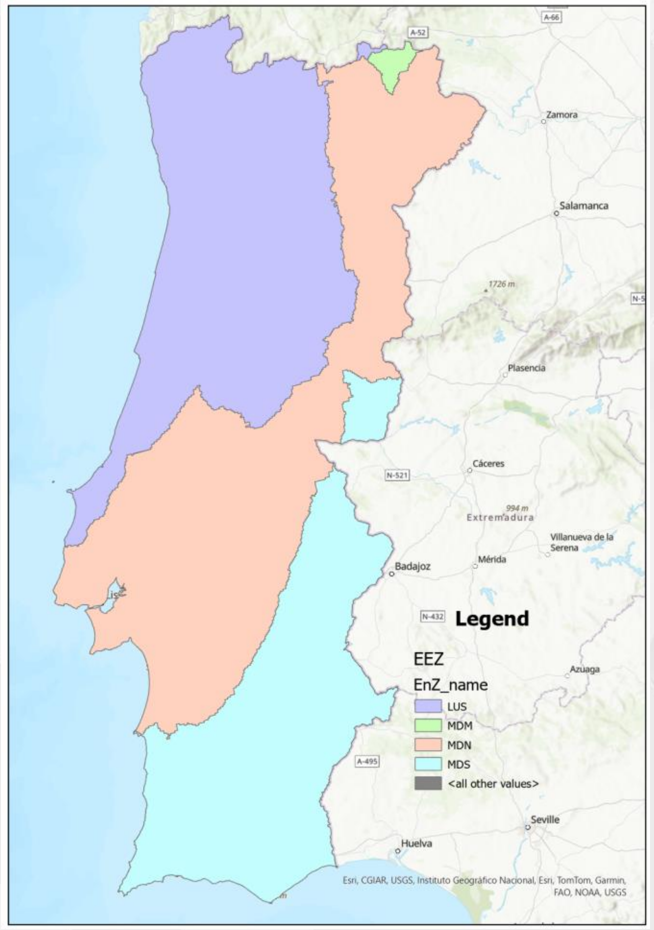
- Bulk Density at three dephts: 0-10 cm, 10-20 cm and 20-30 cm;
- Soil Copper Concentration;
- Nitrogen Content;
- Organic Carbon Content;
- Phosphorus Content;
- Soil pH;
- Soil Texture at three dephts: 0-10 cm, 10-20 cm and 20-30 cm;
- Positive Precipitation;
- Maximum Temperature;
- Minimum Temperature.

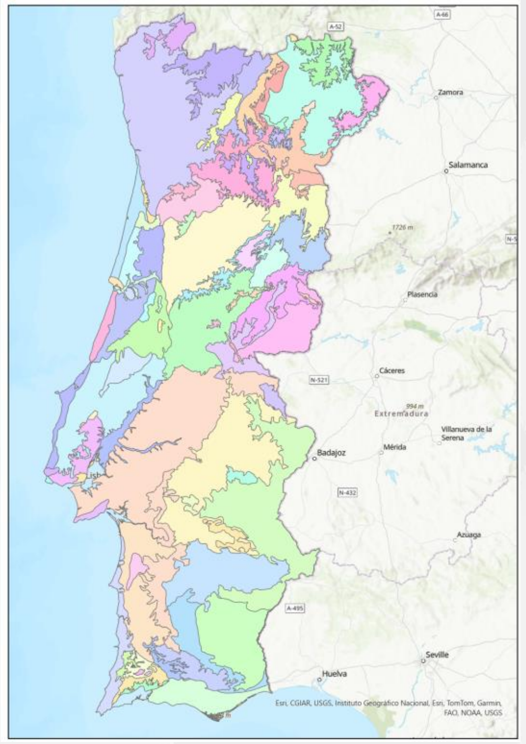
Measured statistics wich were used as input to the cluster algorithms:

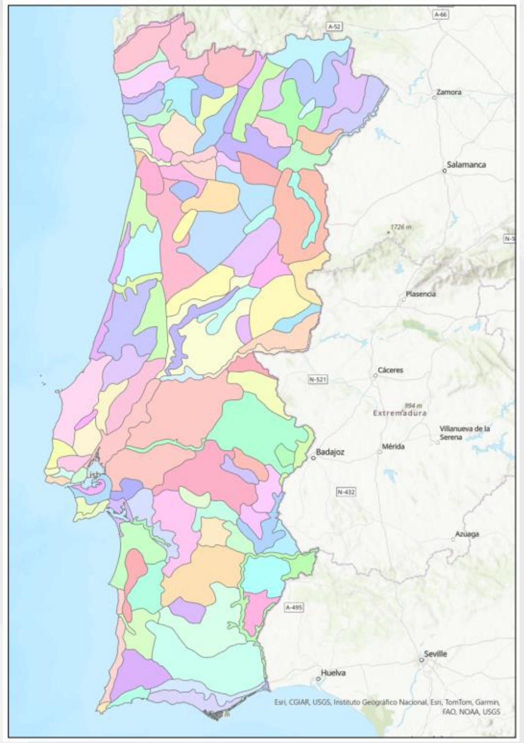
- Minimum
- Maximum
- Mean
- Standard Deviation

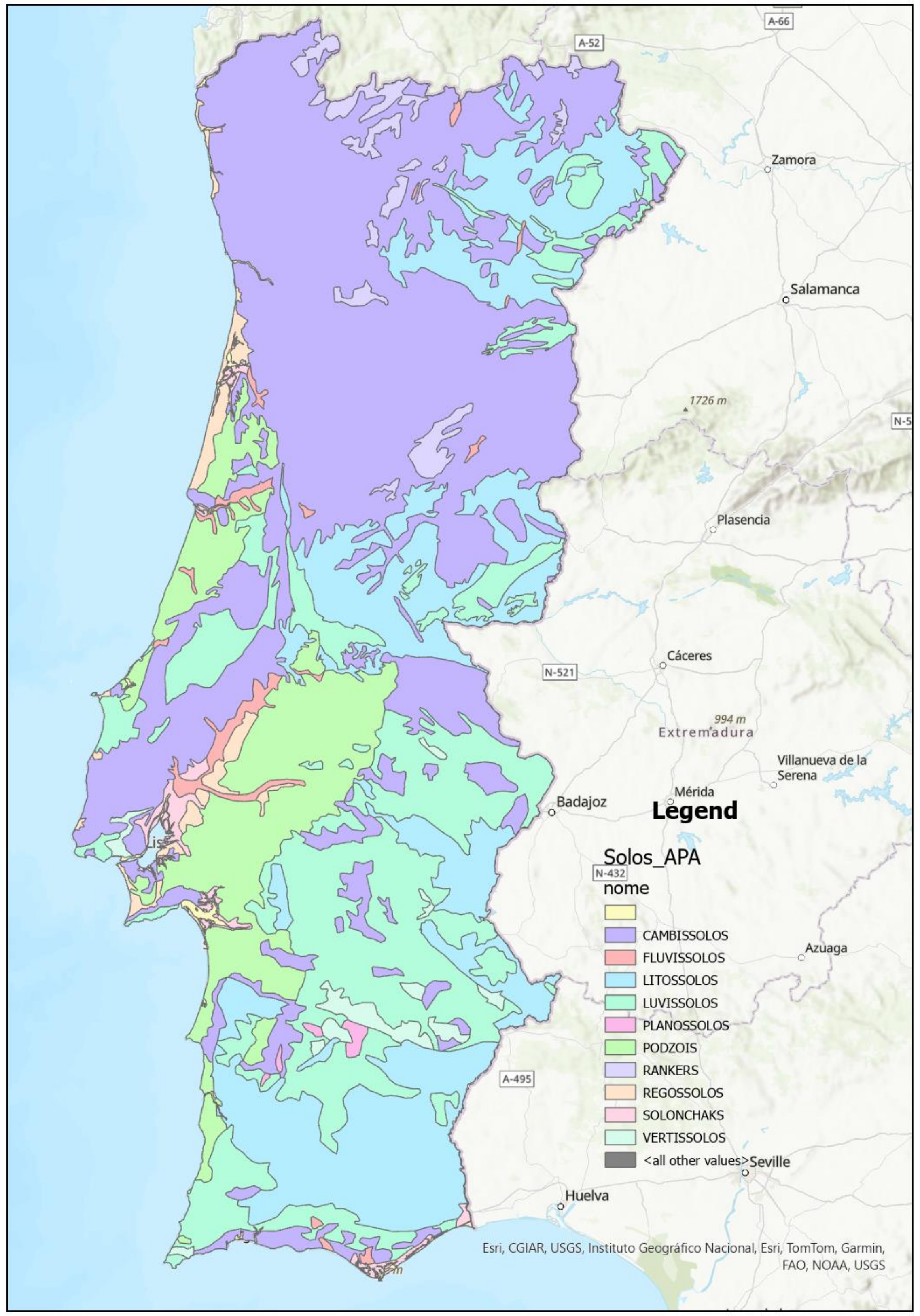
4 th MOPT CONFERENCE

Districts:



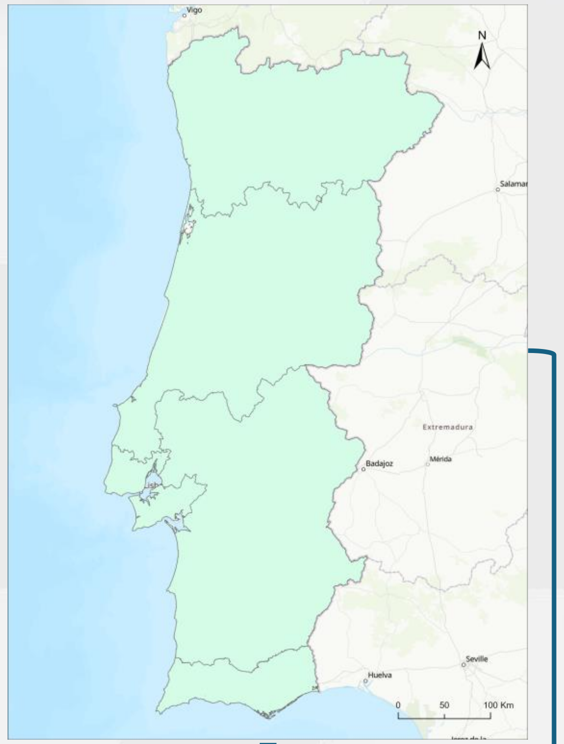




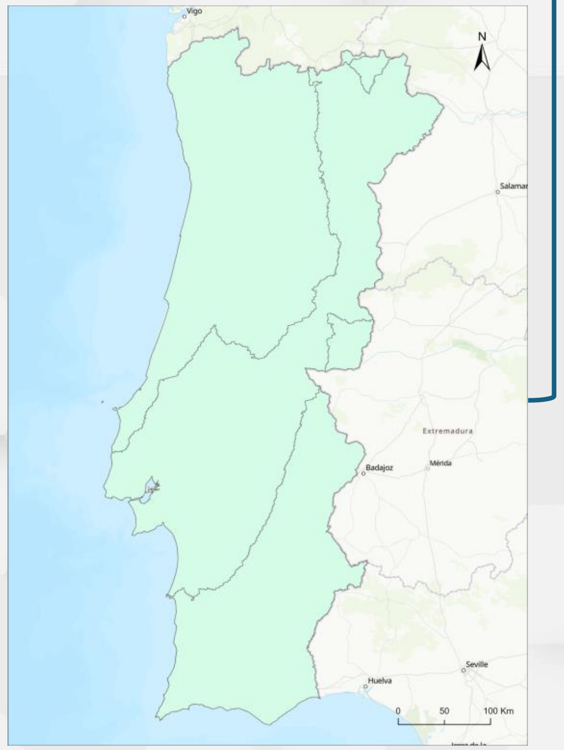


4th MOPT CONFERENCE

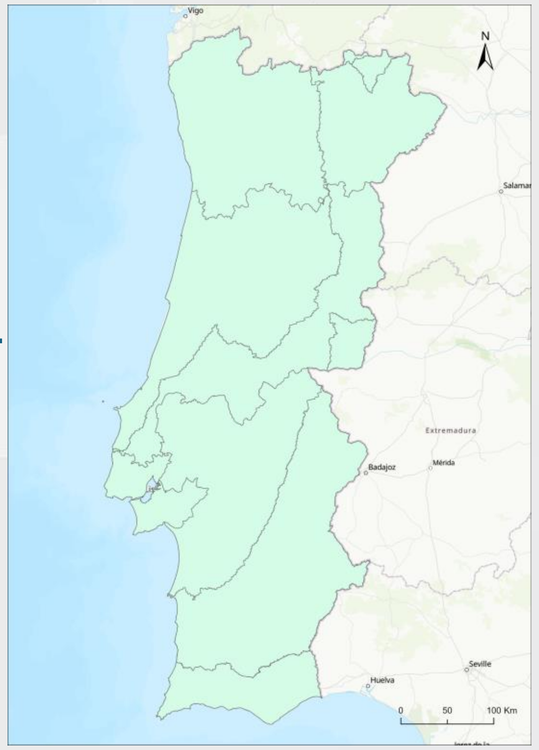
Nuts II



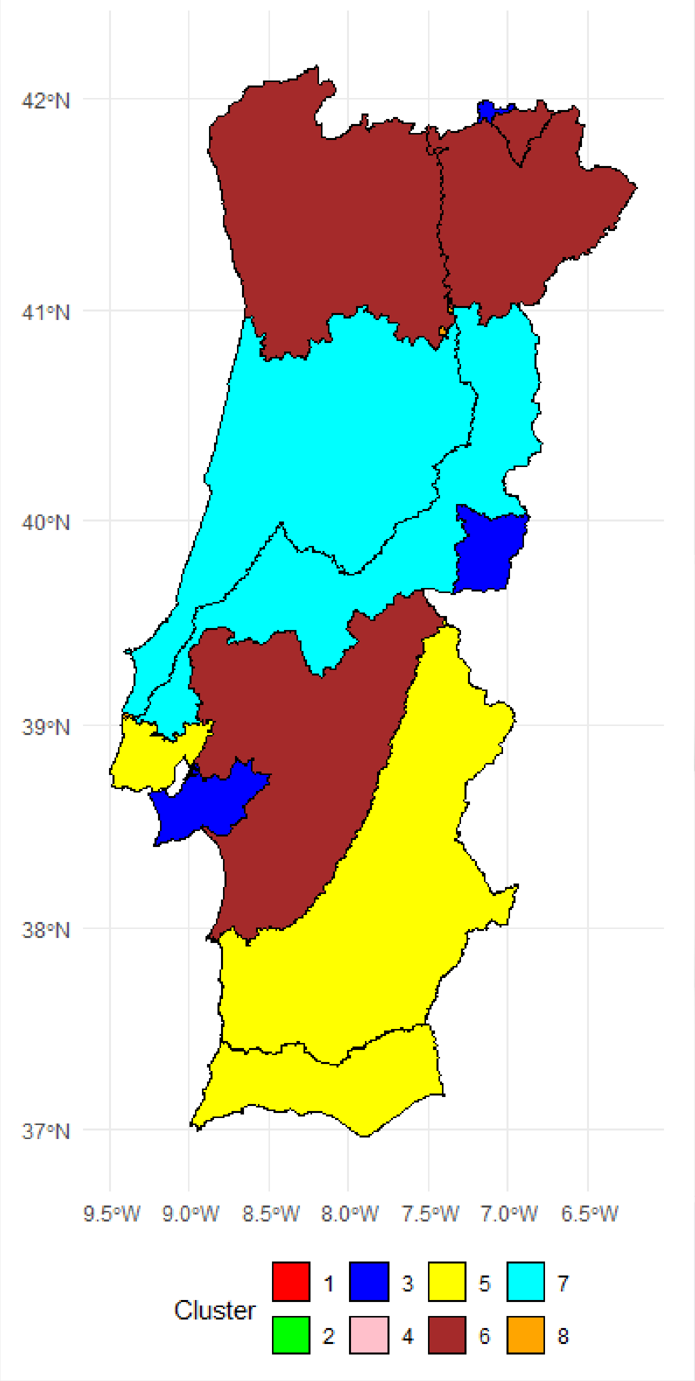
Environmental Zones



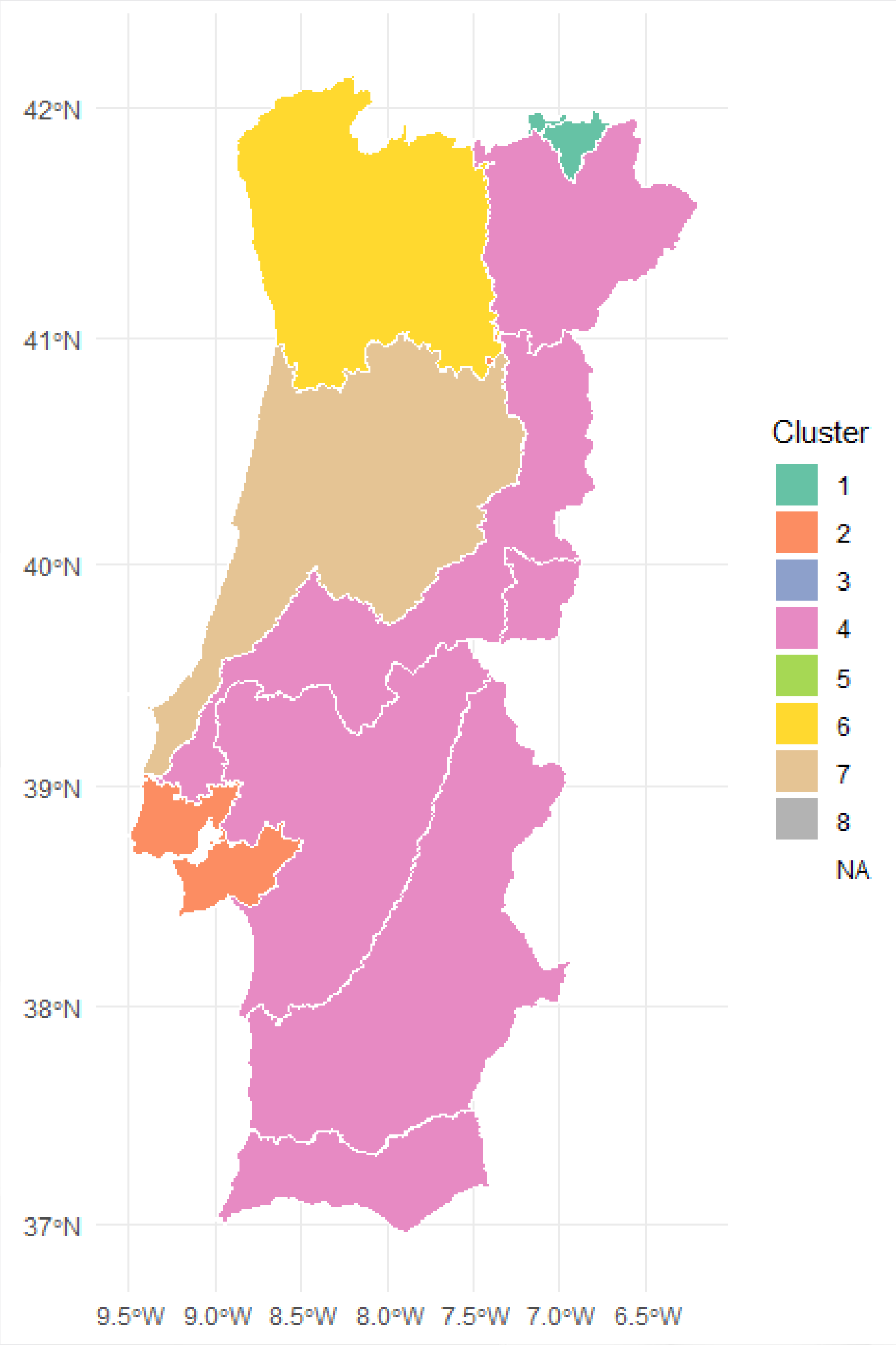
Nuts II intersected with Environmental Zones



Random Forest

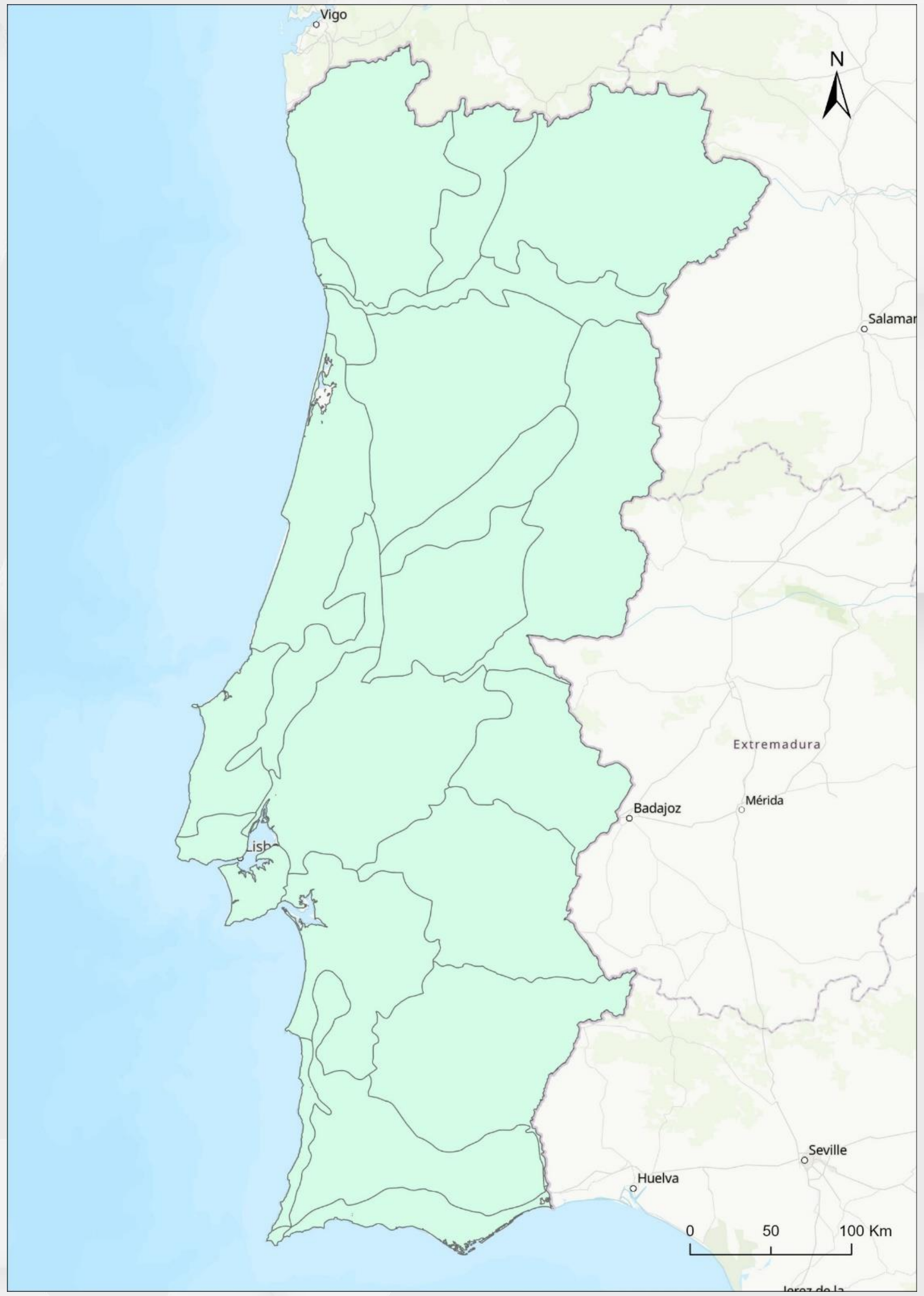


Agglomerative Hierarchical Clustering

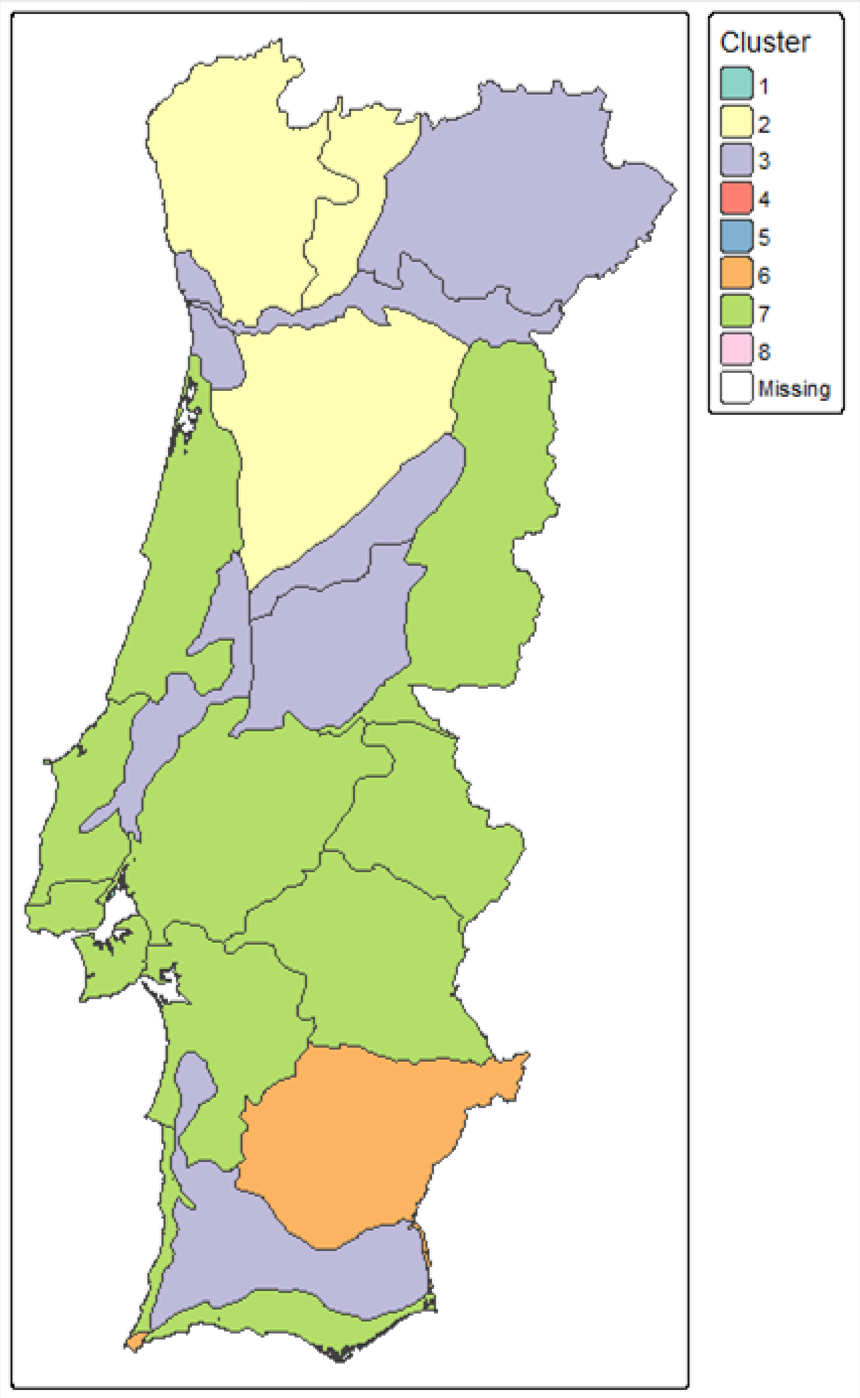


4th MOPT CONFERENCE

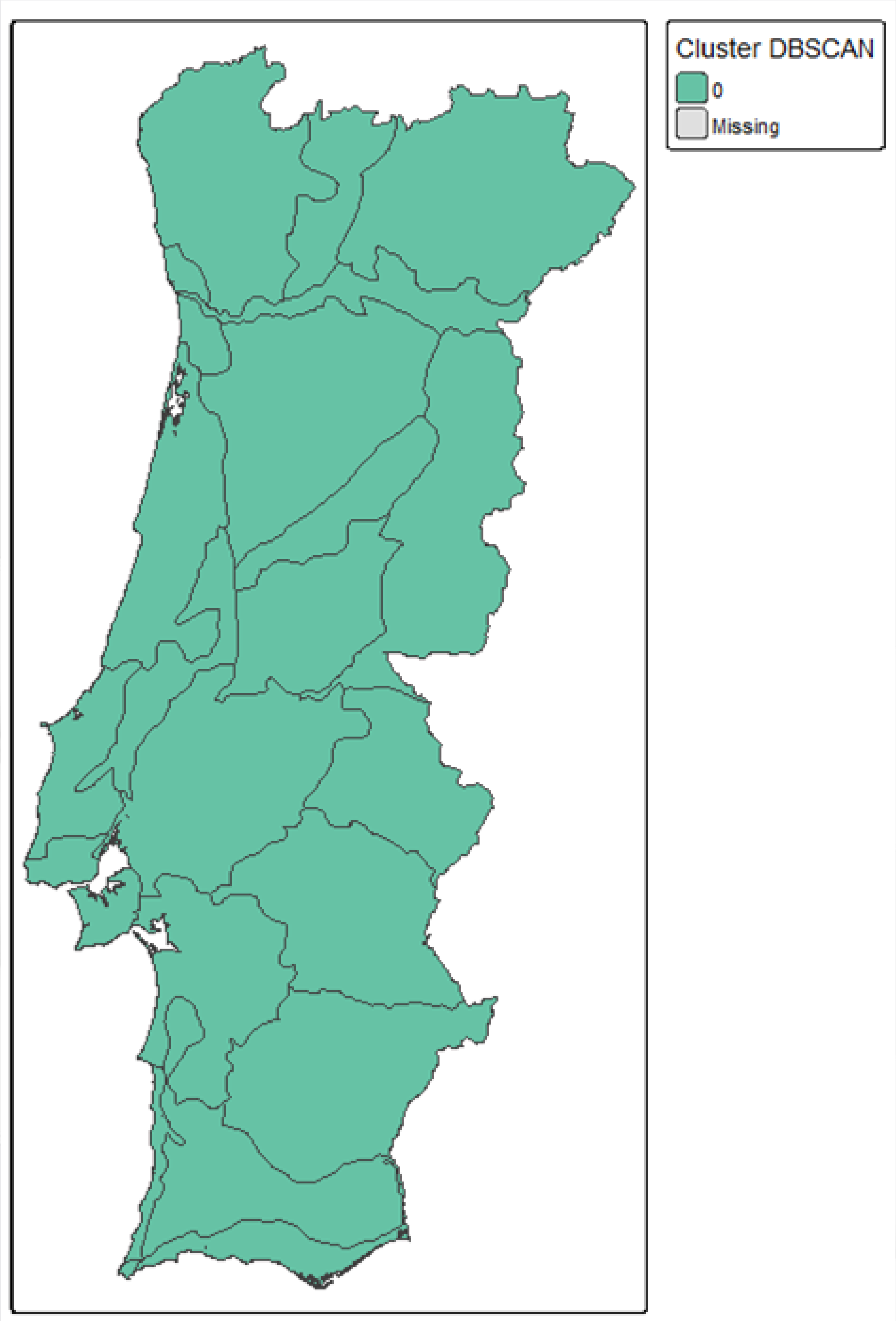
Lanscape Units



Fuzzy c-means

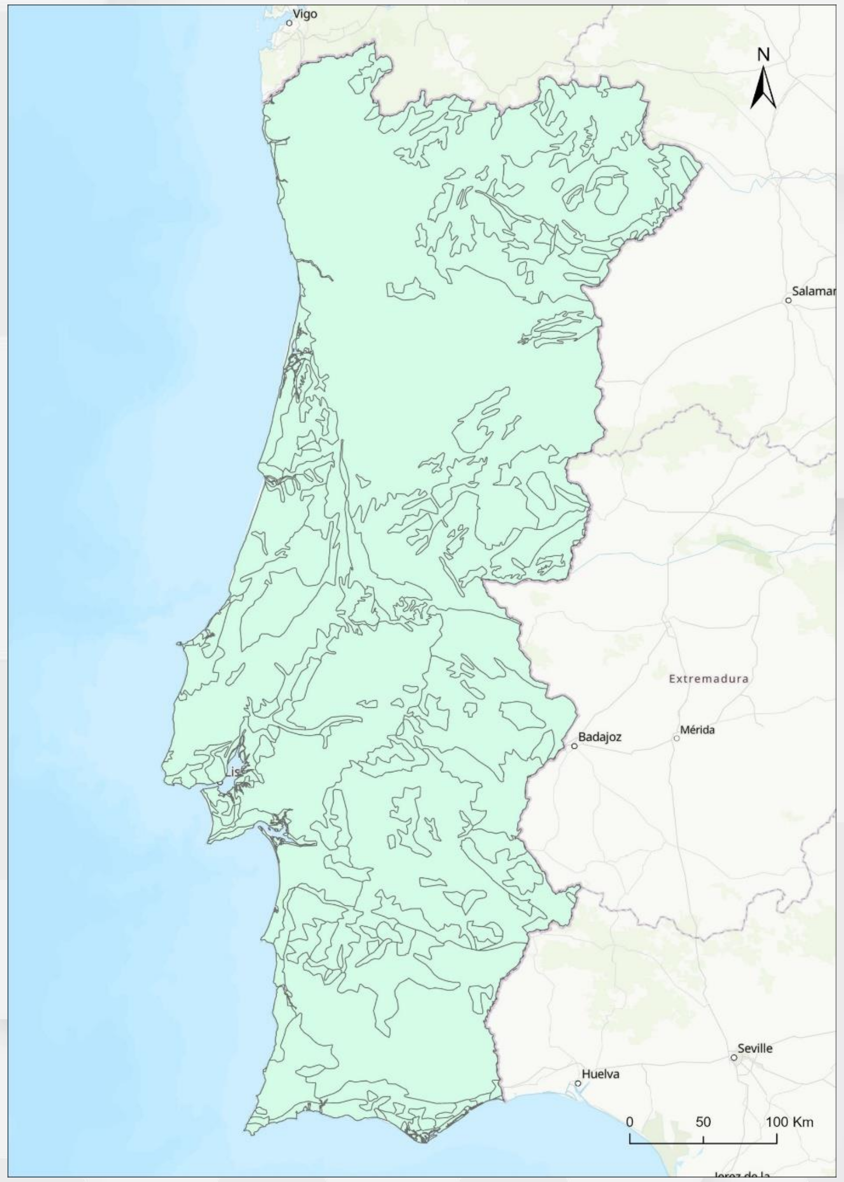


DBSCAN

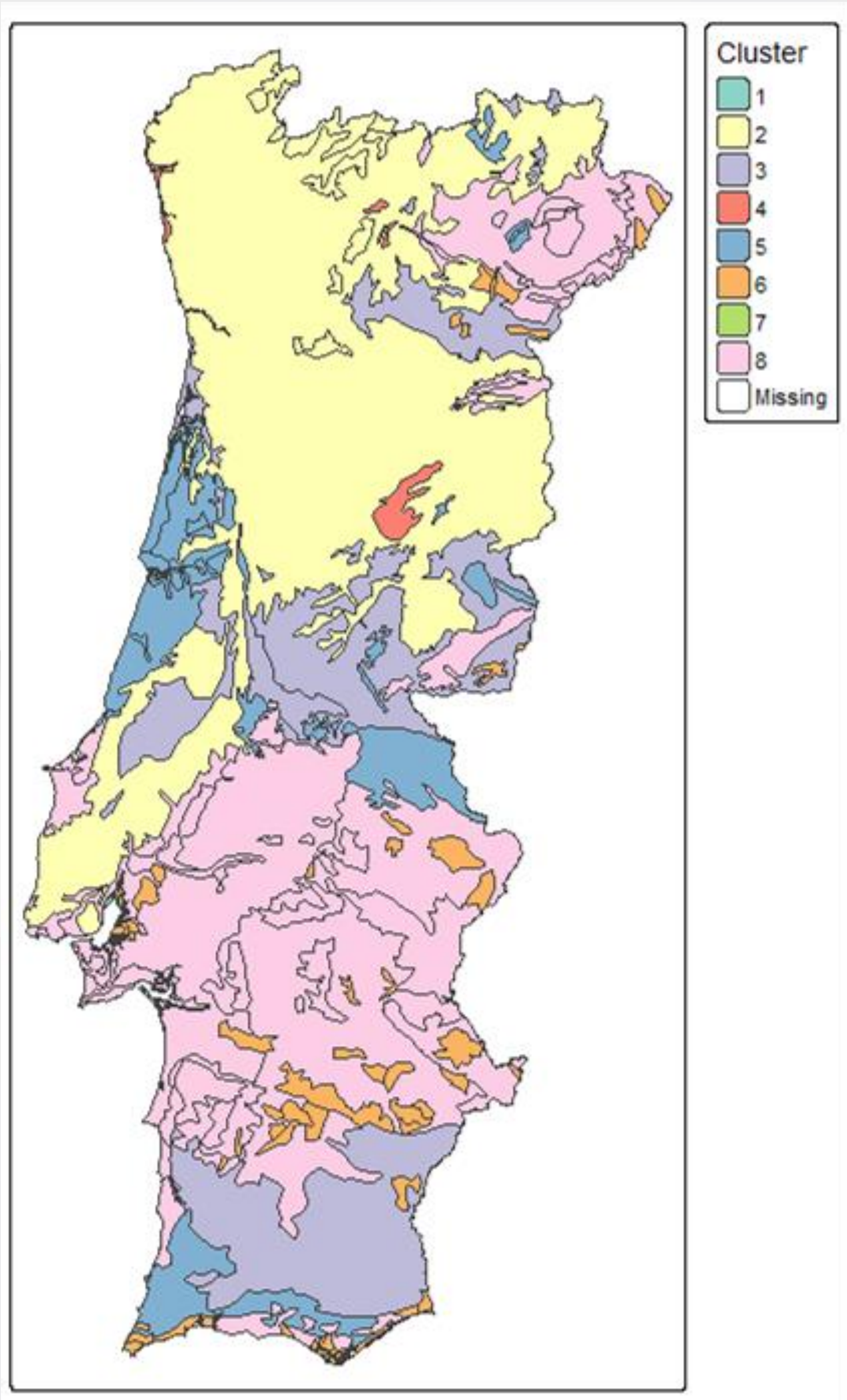


4 th MOPT CONFERENCE

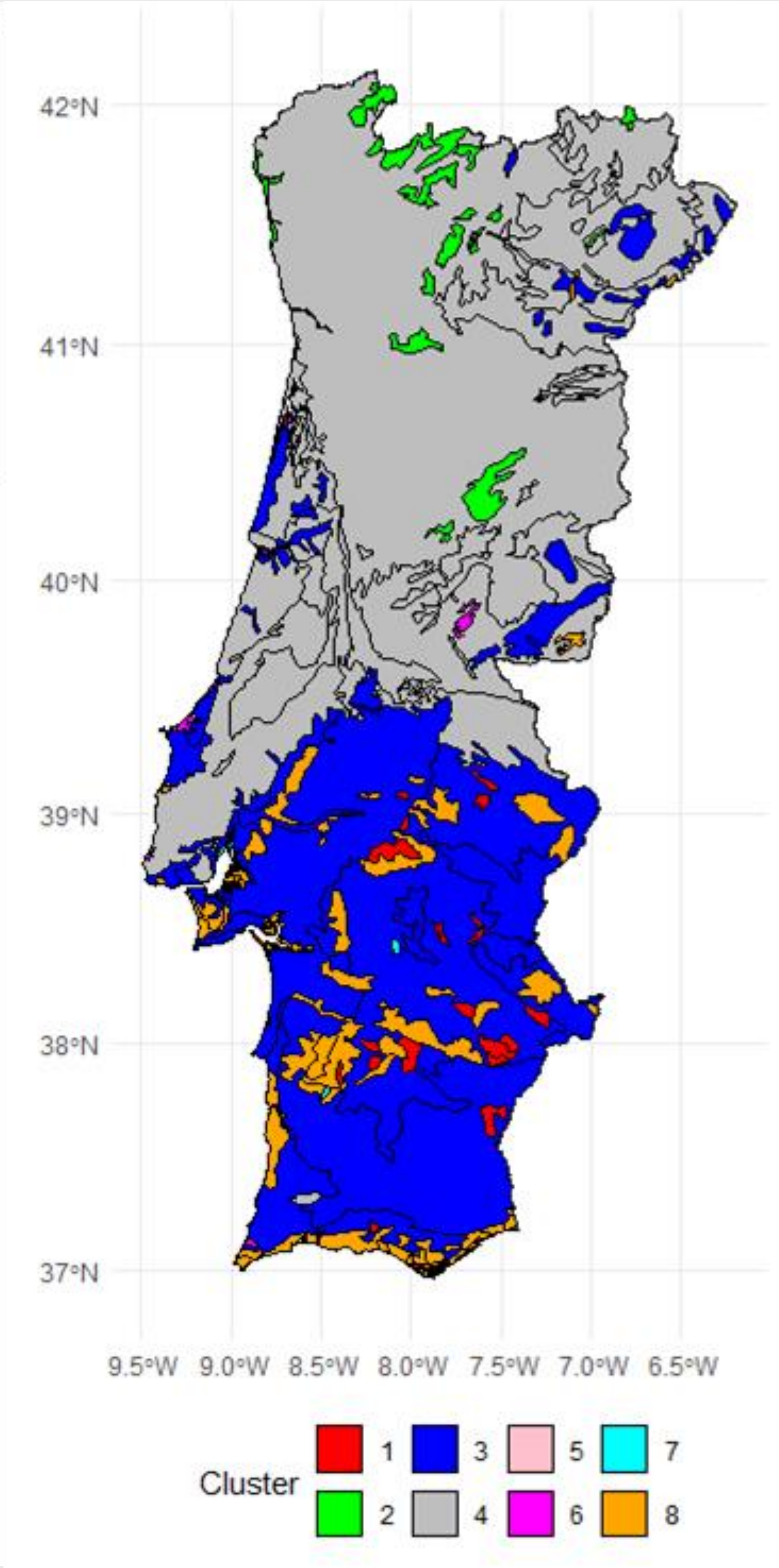
APA Soil Map



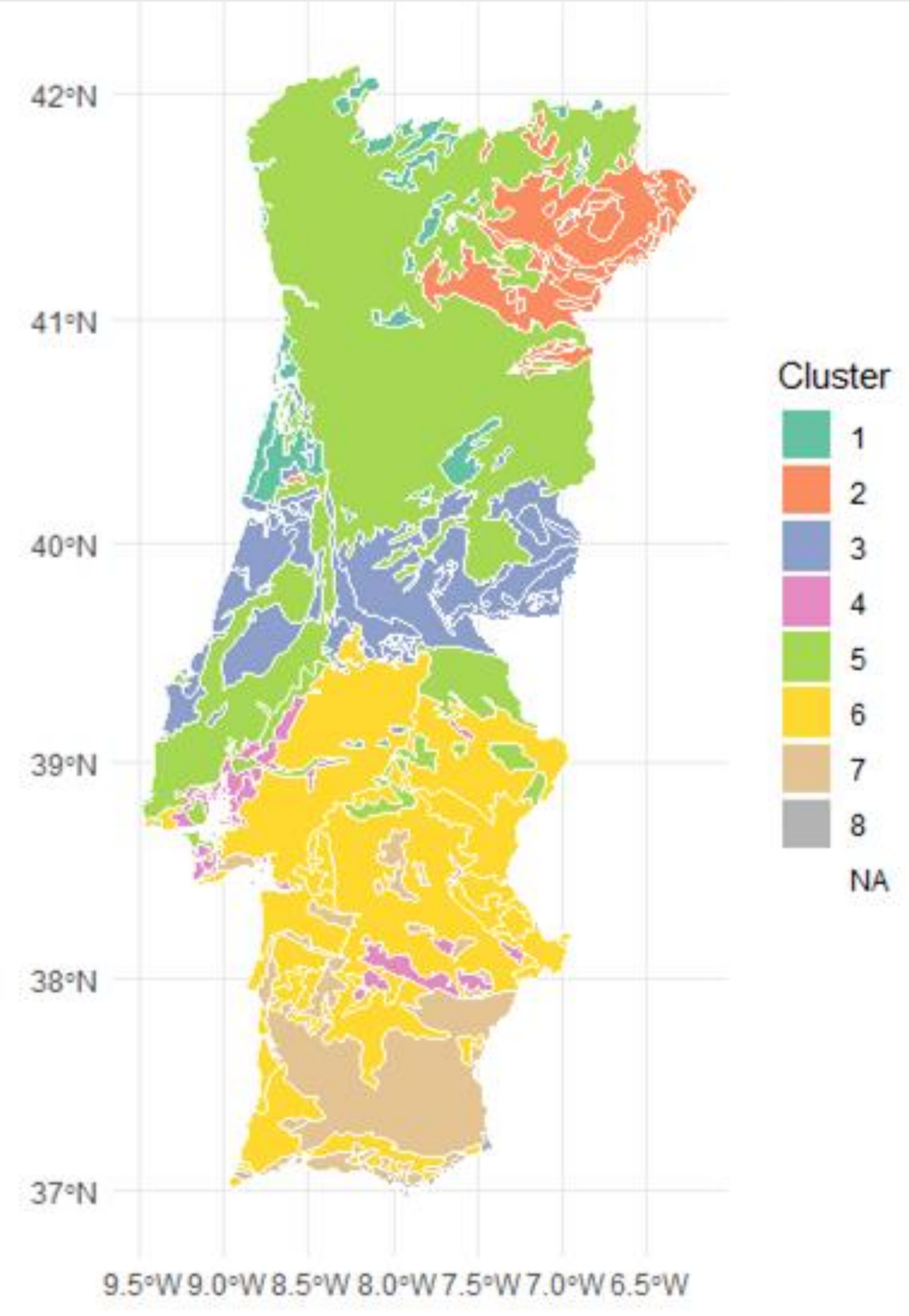
Fuzzy c-means



Random Forest

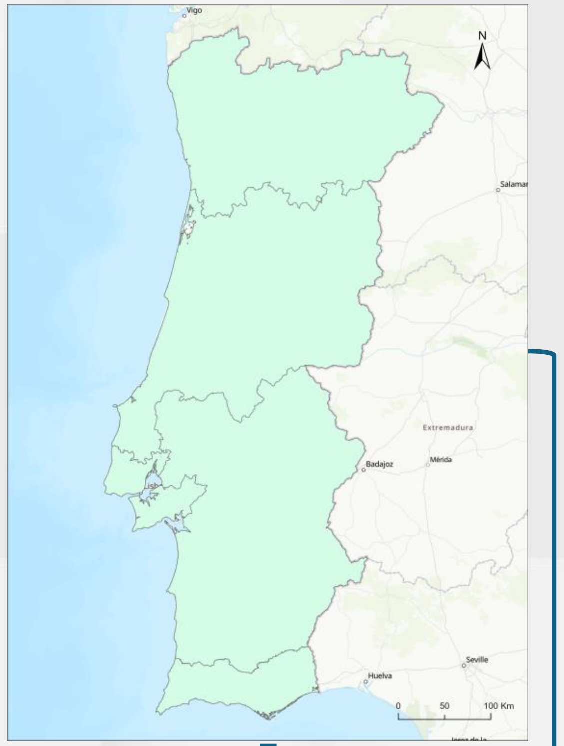


Agglomerative
Hierarchical Clustering



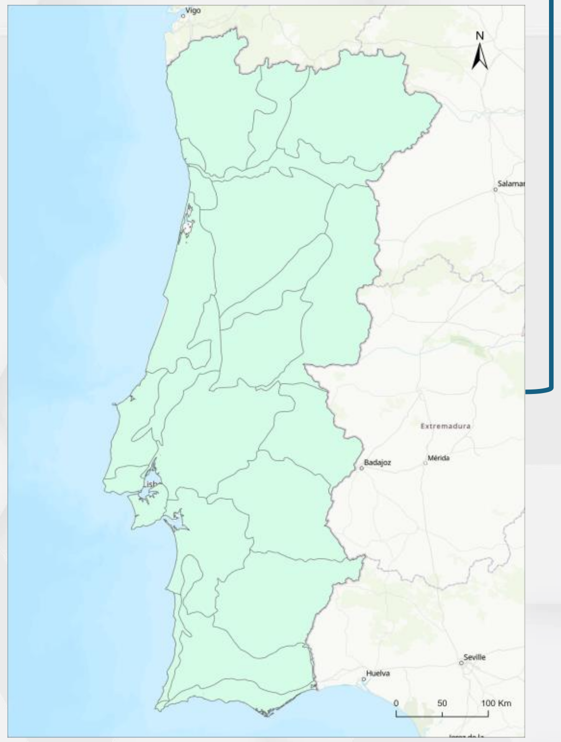
4th MOPT CONFERENCE

Nuts II

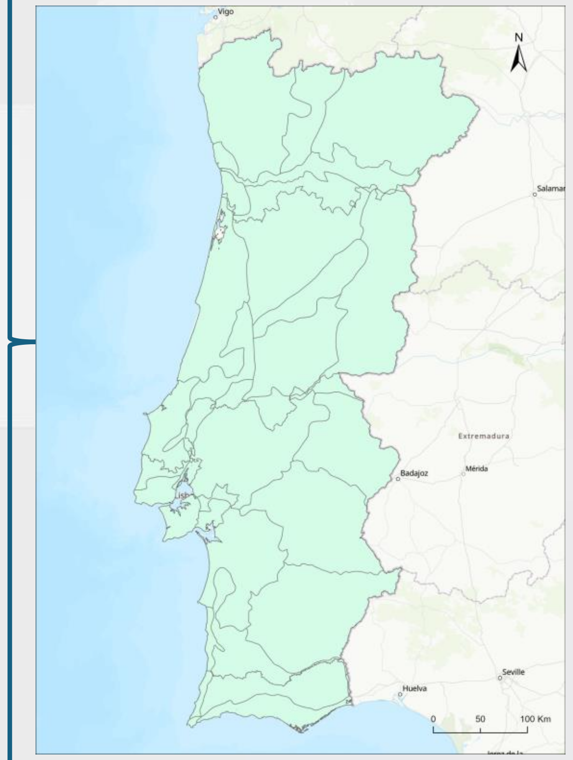


✚

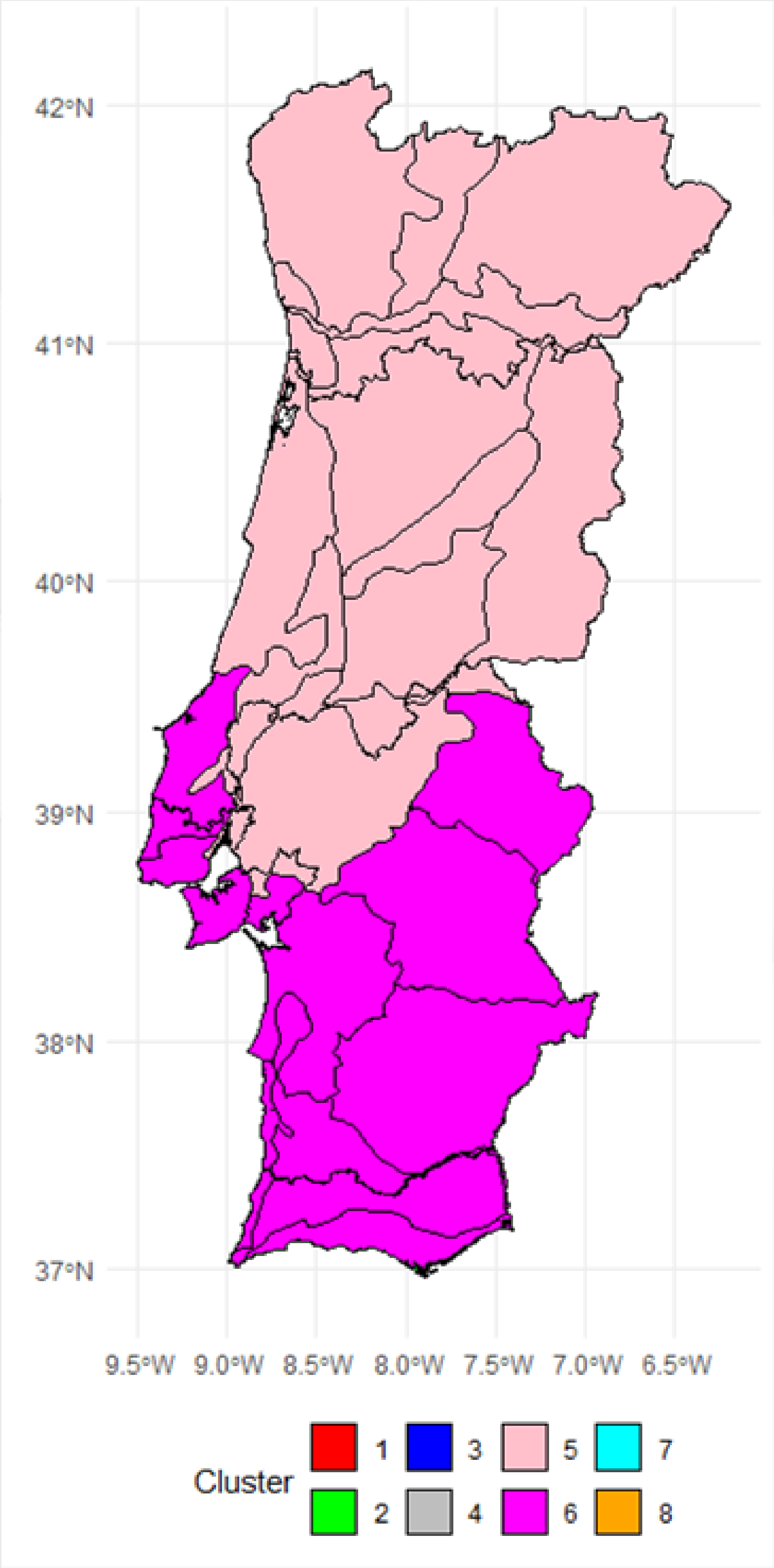
Landscape Units



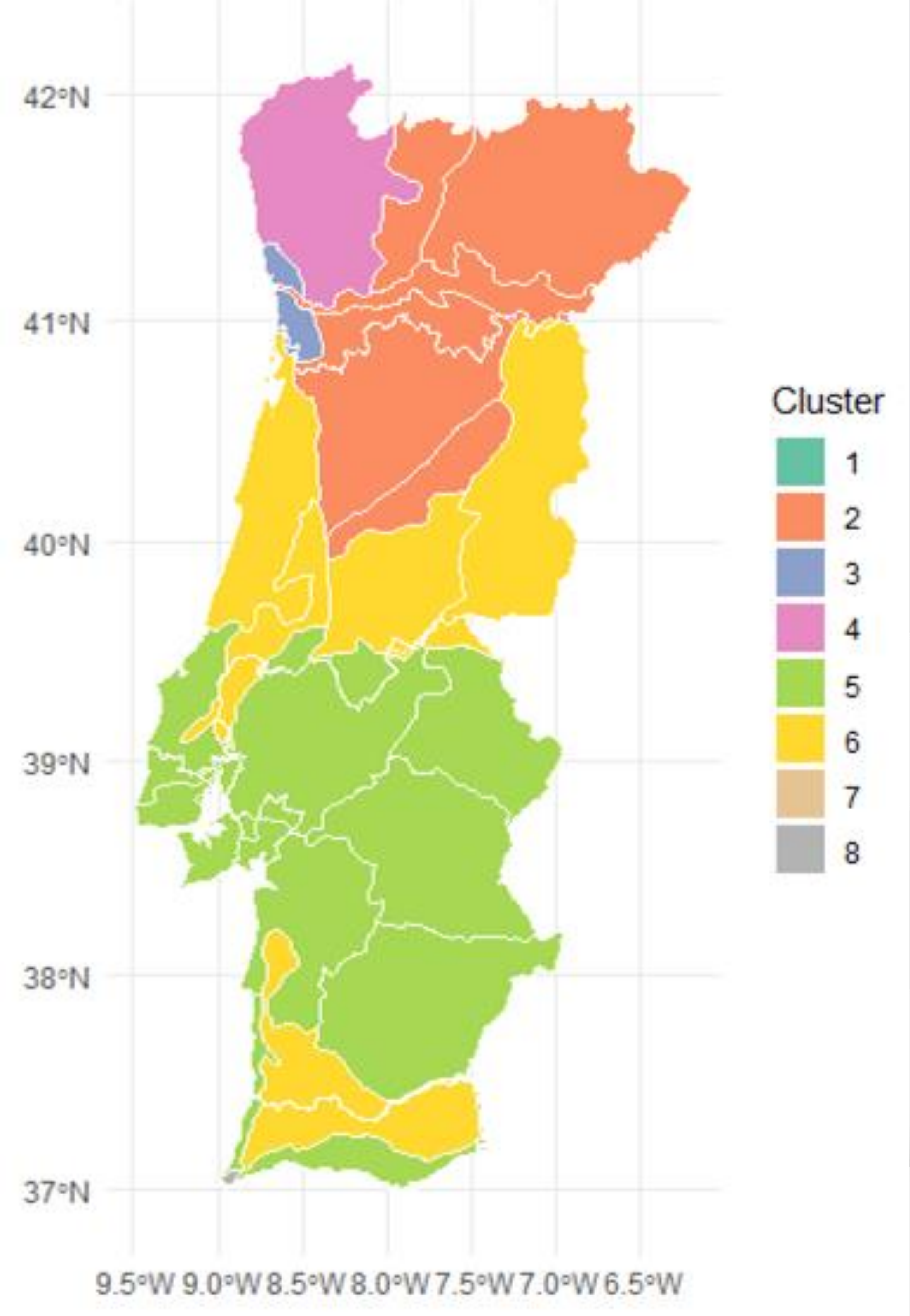
Nuts II intersected with Landscape Units



Random Forest

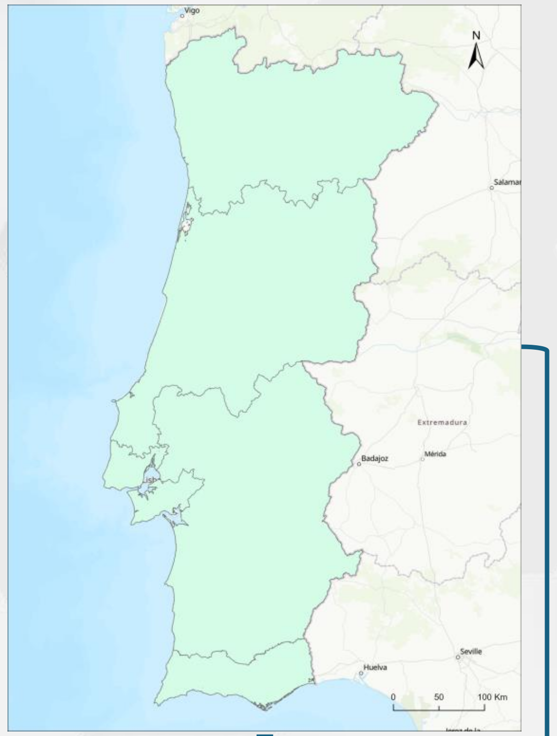


Agglomerative Hierarchical Clustering

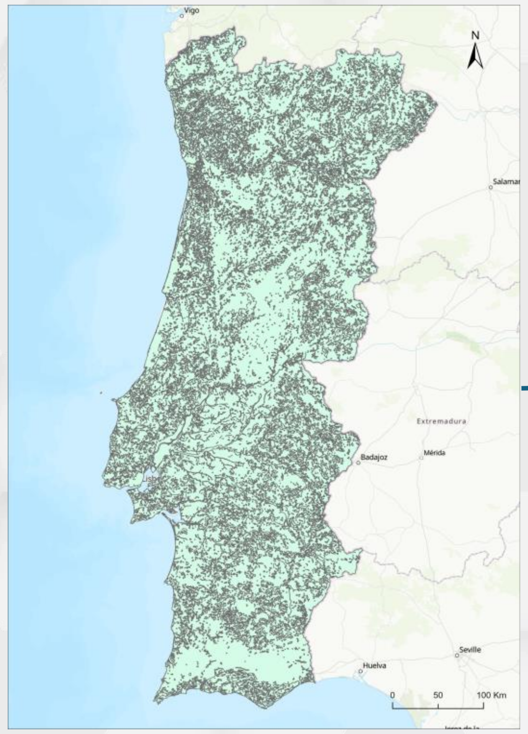


4 th MOPT CONFERENCE

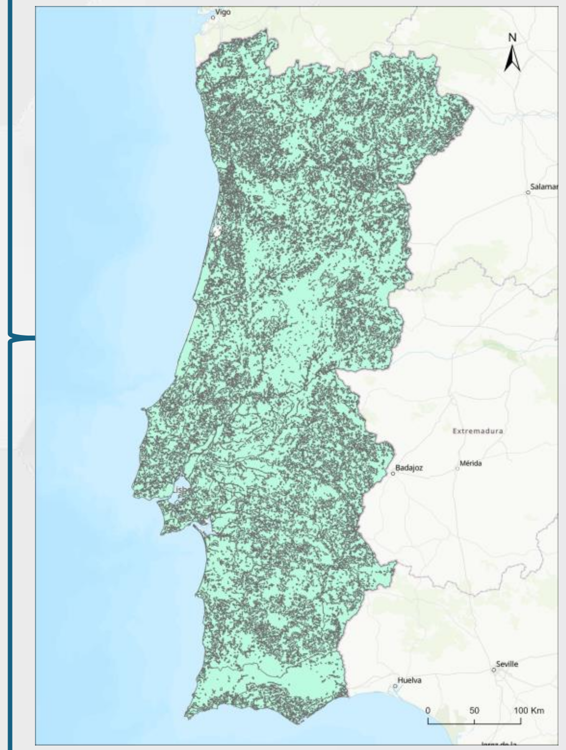
Nuts II



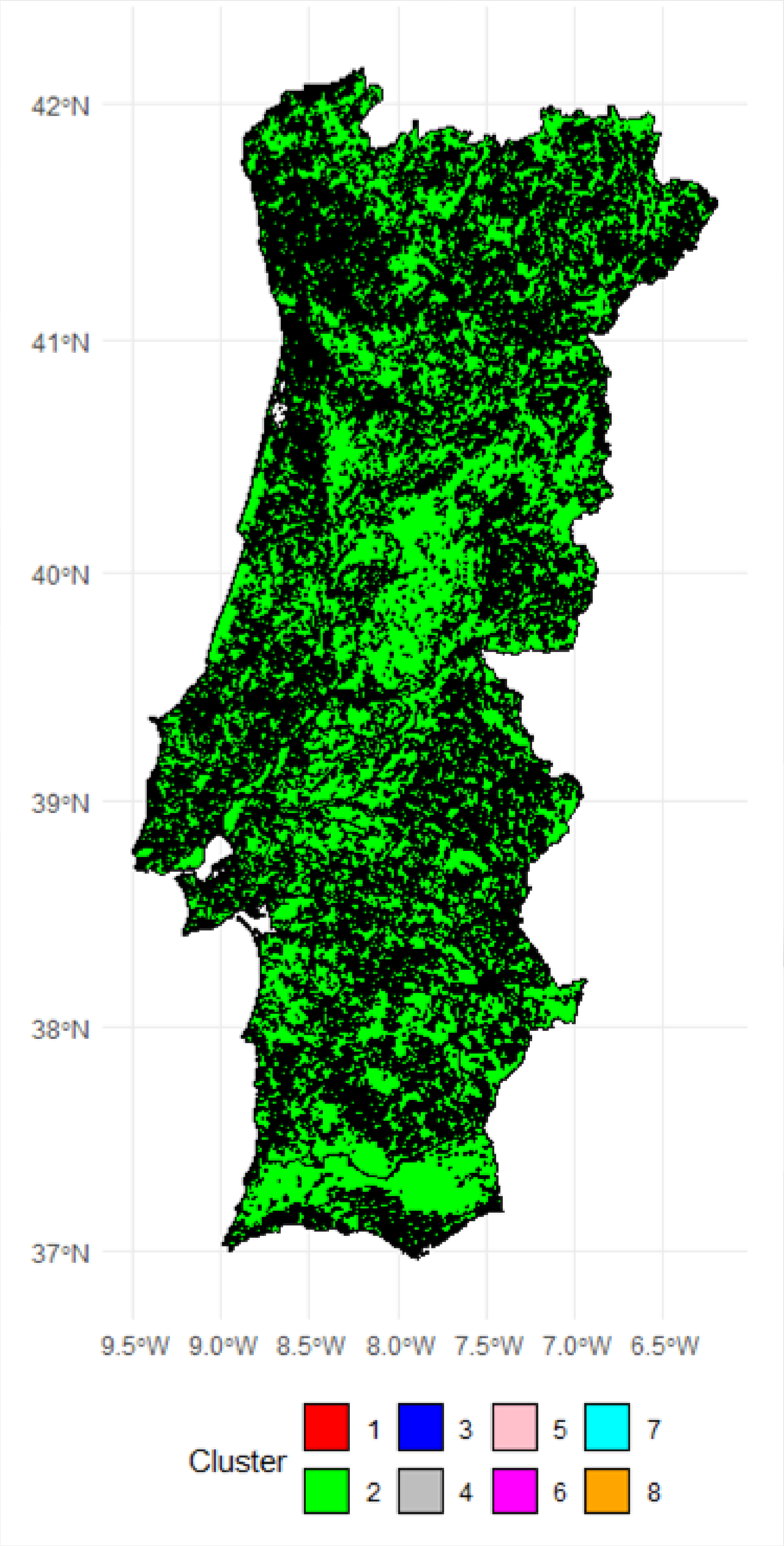
Corine Land Cover



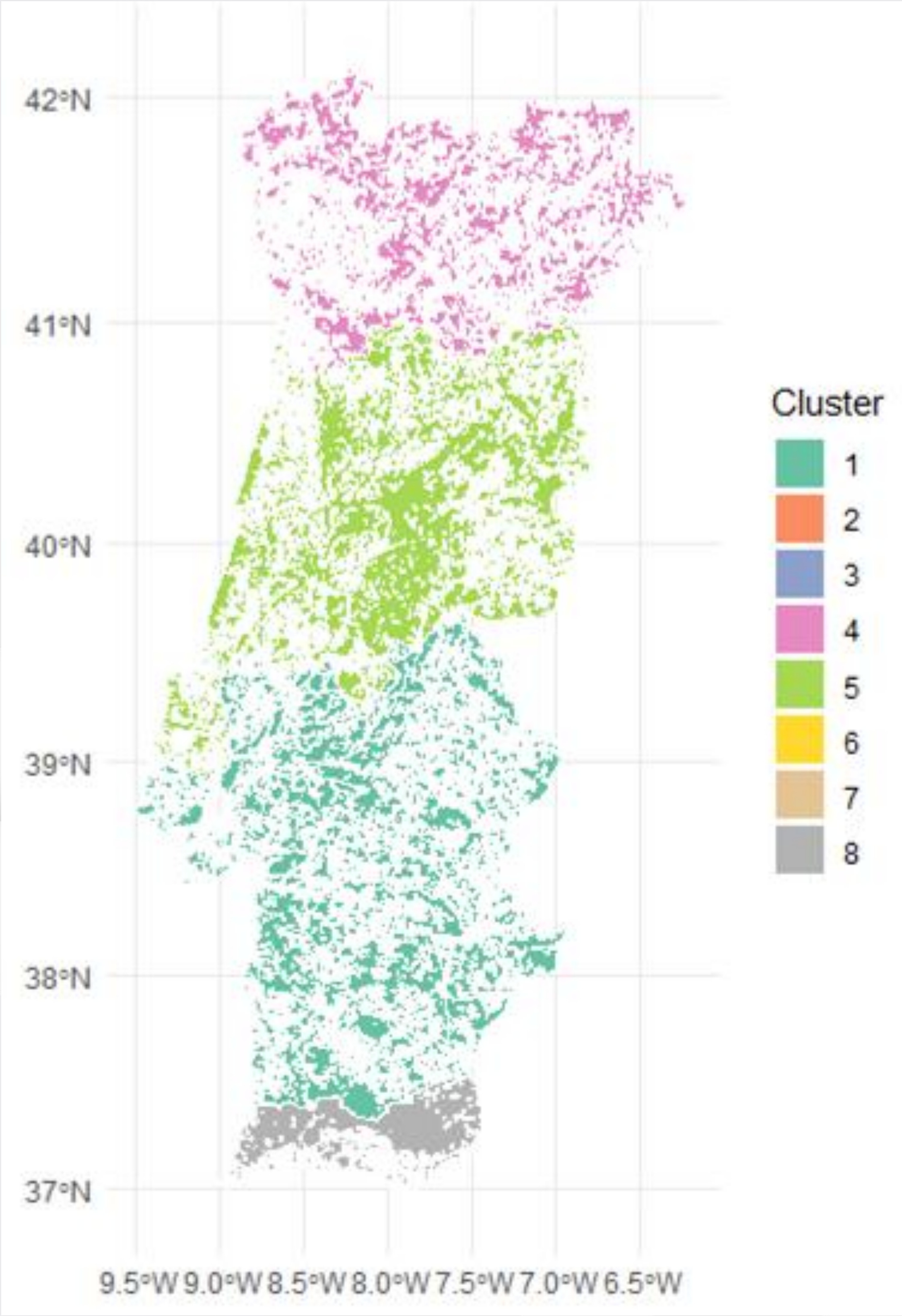
Nuts II intersected with Corine Land Cover



Random Forest

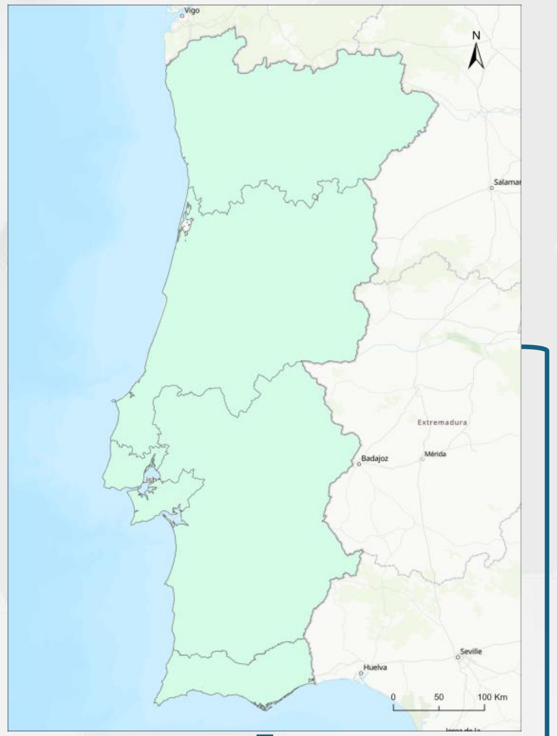


Agglomerative Hierarchical Clustering

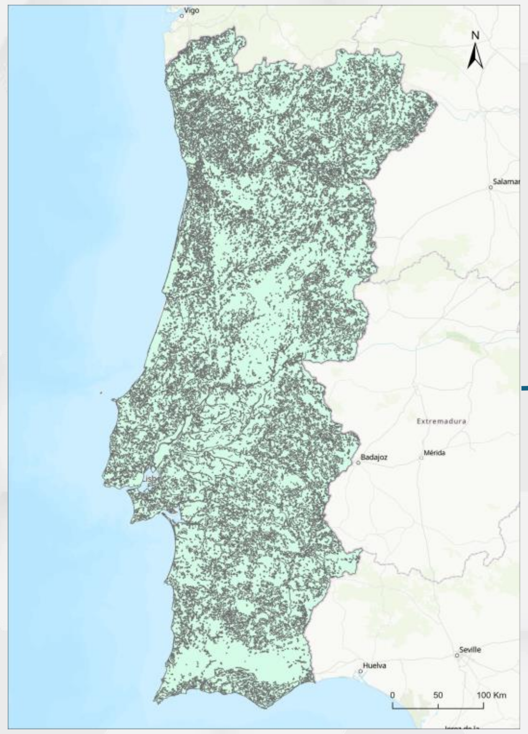


4 th MOPT CONFERENCE

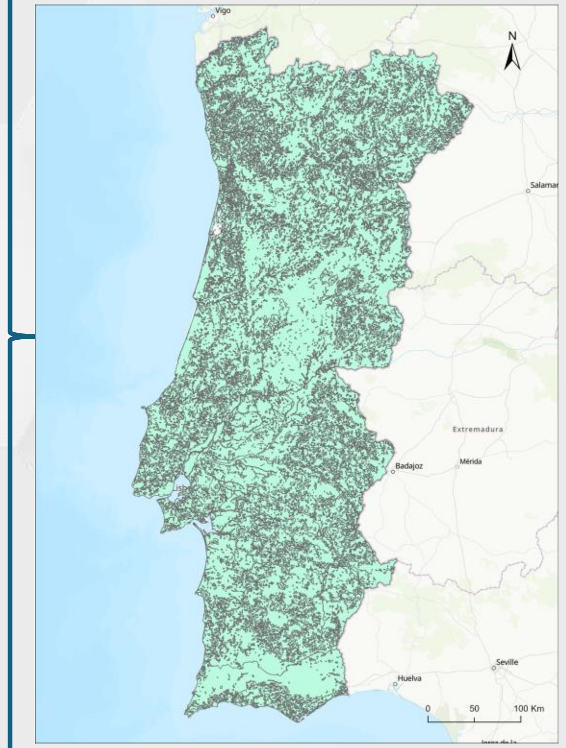
Nuts II



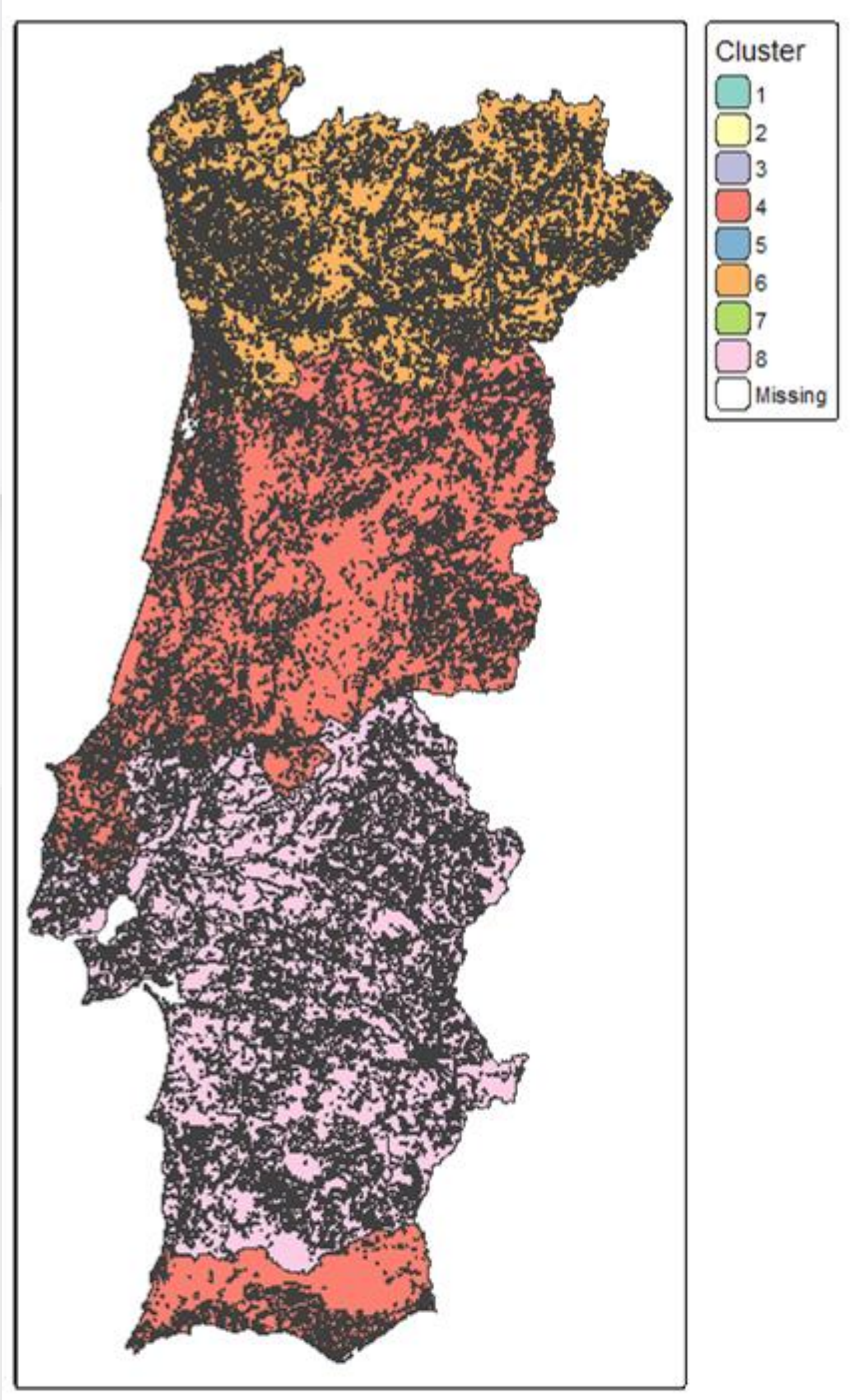
Corine Land Cover



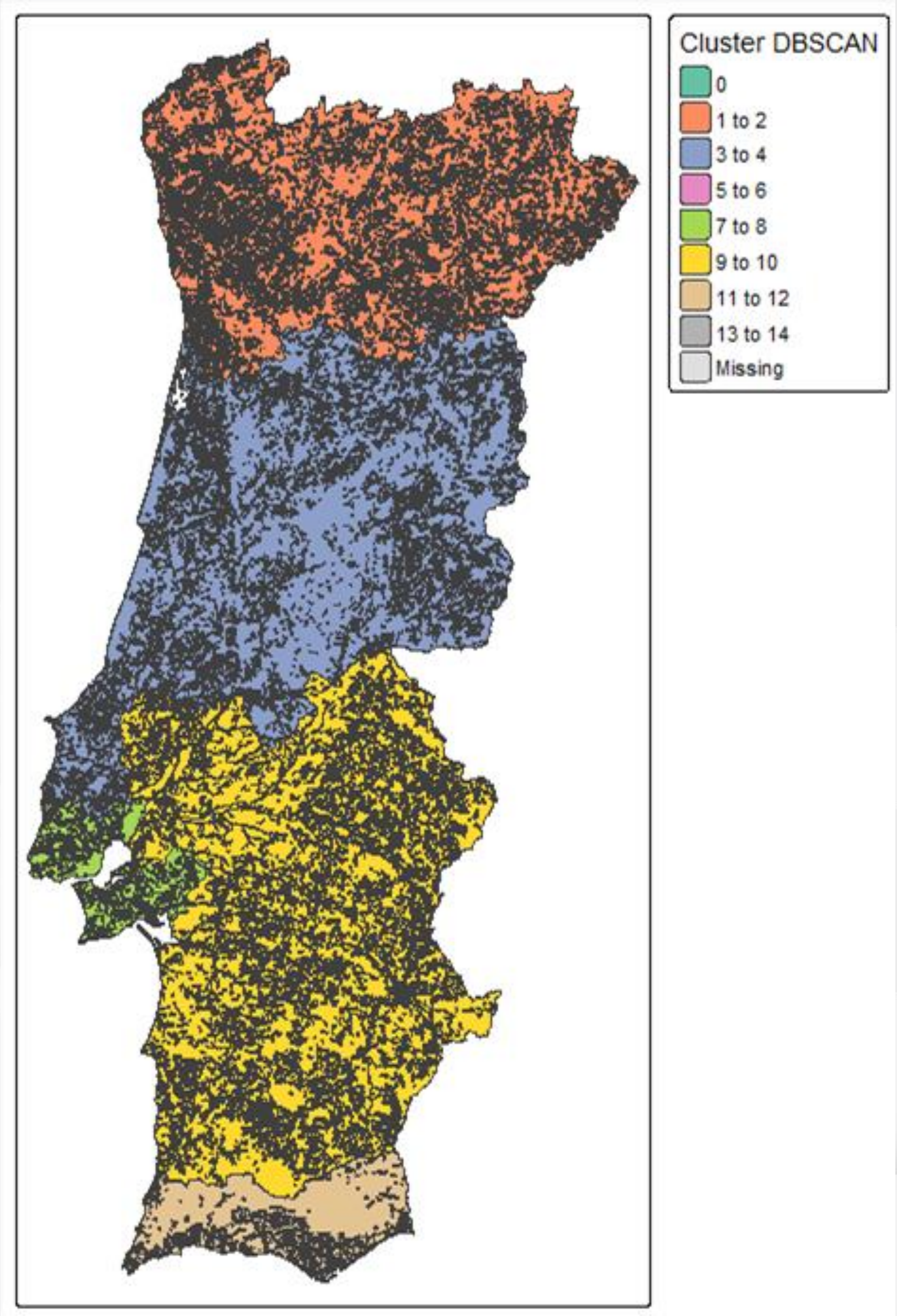
Nuts II intersected with Corine Land Cover



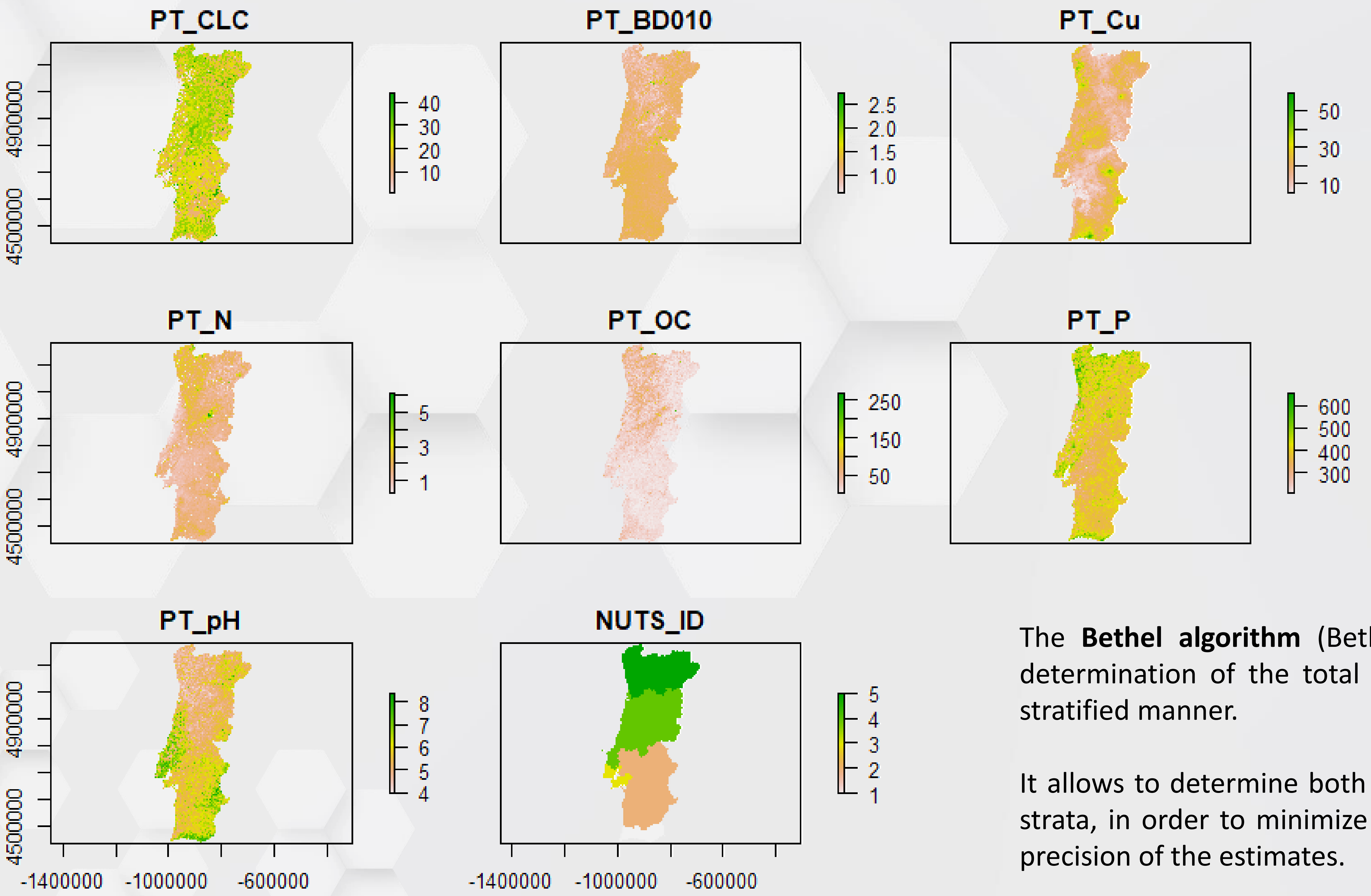
Fuzzy c-means



DBSCAN



4 th MOPT CONFERENCE



1. Select (few) soil properties
2. Obtain LUCAS soil property derived maps using ML and EO as proxy for the true soil property distribution
3. Prototype strata inputs based on land cover types, NUTS regions and climate.
4. Identify a reasonable number of candidate points by random sampling
5. Selecting a manageable subsample
6. The number of samples with a target of 0.05 VC on the selected soil properties

The **Bethel algorithm** (Bethel, 1989) constitutes an approach for the ideal determination of the total sample size and allocation of sampling units in a stratified manner.

It allows to determine both the total sample size and the allocation of units in strata, in order to minimize costs due to restrictions arising from the levels of precision of the estimates.

4 th MOPT CONFERENCE

File	Description
PT_BD010.tif	Bulk density in the 0-10 cm layer
PT_BD1020.tif	Bulk density in the 10-20 cm layer
PT_BD2030.tif	Bulk density in the 20-30 cm layer
PT_CEC.tif	Cation exchange capacity
PT_CF.tif	Coarse fragments
PT_CLC.tif	Corine Land Cover class
PT_N.tif	Nitrogen content
PT_NUTS0.tif	NUTS0 region
PT_NUTS1.tif	NUTS1 region
PT_NUTS2.tif	NUTS2 region
PT_OC.tif	Organic carbon content
PT_pH.tif	Soil pH
PT_P.tif	Phosphorus content
PT_TXT.tif	Soil texture
PT_RUSLE.tif	RUSLE soil erosion map
PT_Cu.tif	Soil copper concentration
PT_Cd.tif	Soil cadmium concentration
PT_Hg.tif	Soil mercury concentration
PT_Zn.tif	Soil zinc concentration
PT_CaCO3.tif	Soil calcium carbonate

Input variables

Administrative/Geographic divisions	Soil properties and Climatic variables	N° of sampling points generated	Nº of Soil Districts
Carta de Solos (APA)	B010, Cu, N, OC, P, pH	395	33
	CLC, B010, Cu, N, OC, P, pH	431	29
Unidades de Paisagem (DGT)	B010, Cu, N, OC, P, pH	877	71
	CLC, B010, Cu, N, OC, P, pH	938	57
NUTS II e Zonas Ambientais	B010, Cu, N, OC, P, pH	458	39
NUTS II	B010, Cu, N, OC, P, pH	286	16
	CLC, B010, Cu, N, OC, P, pH	236	14
	BD010, Cu, N, OC, P, pH, PP, TMAX, TXT010	557	13
	COS2018, BD010, Cu, N, OC, P, pH, PP, TMAX, TXT010	519	14
	BD1020, Cu, N, OC, P, pH, PP, TMAX, TXT1020	382	12
	COS2018, BD1020, Cu, N, OC, P, pH, PP, TMAX, TXT1020	375	11
	BD2030, Cu, N, OC, P, pH, PP, TMAX, TXT2030	377	12
	COS2018, BD2030, Cu, N, OC, P, pH, PP, TMAX, TXT2030	418	11

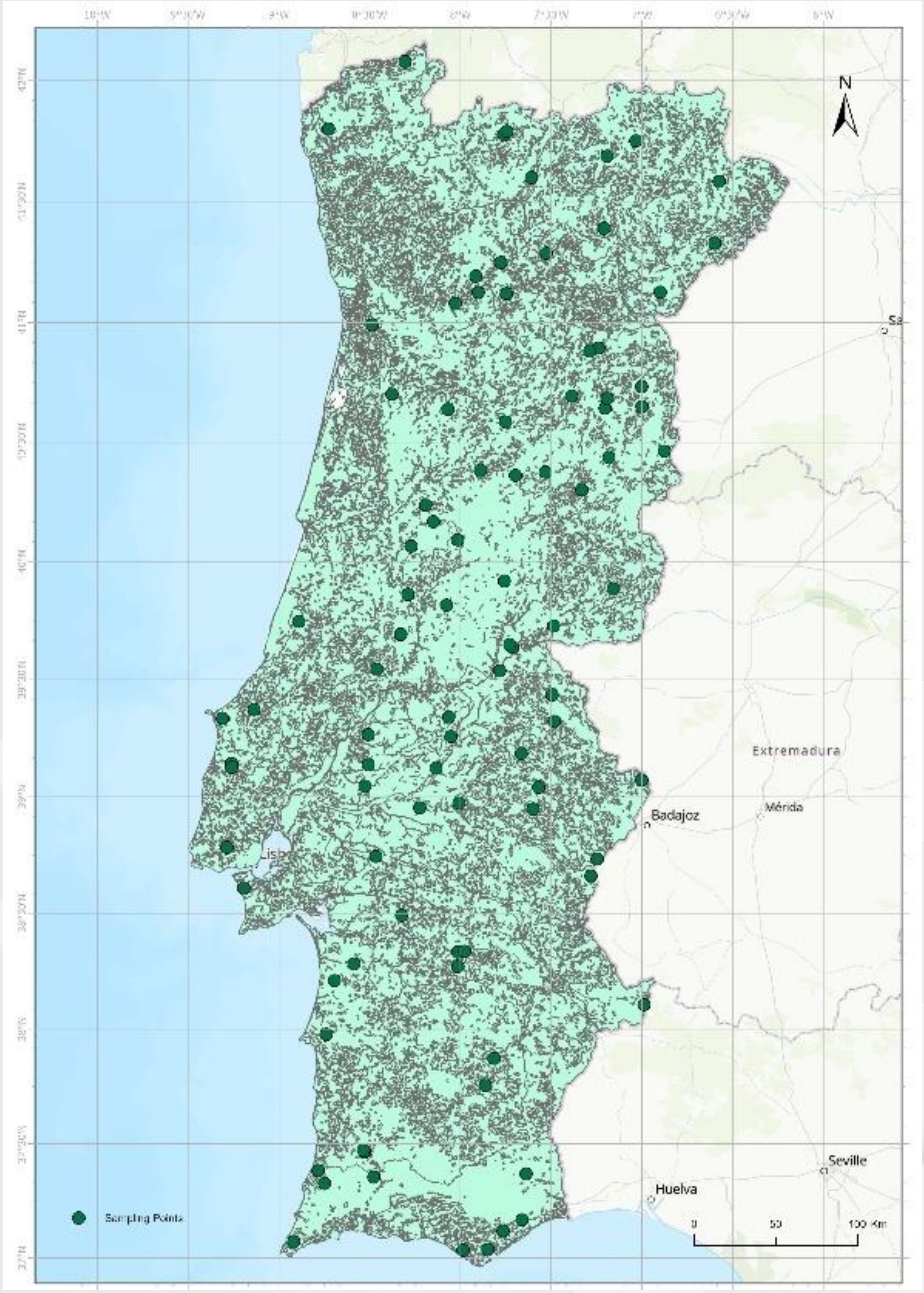
Best Result: CLC, B010, Cu, N, OC, P, pH — 236 — 14

4th MOPT CONFERENCE

Portugal are those where the number of sampling points is lower and the number of resulting soil units allows one to verify that there is neither excessive stratification of the territory nor excessive generalization, due to the heterogeneity of soils in mainland Portugal.

The best combination obtained in this study was Test 7, which used as a domain the Nuts II regions combined with the simplified Corine Land Cover map, as well as using few variables regarding soil properties and climatic variables, which nevertheless proved to be quite important in the algorithm. This was also the most favorable result in financial terms, since soil samples are expensive and the European Union only covers 20% of that cost.

It is concluded from this study that when using this type of algorithm, the use of too many variables does not show such favorable results, and overly detailed variables also end up impairing the algorithm's results. Variables with too many classes can likewise negatively affect the algorithm, since the classes may be disproportionate in size and intra-class values. The use of such variables leads to certain results showing a high number of sampling points that are disproportionate to the reality of the country's soils.



THANK YOU

JORGE.ROCHA@EDU.ULISBOA.PT