Delimitation of Soil Districts in mainland Portugal

M. Conceição, G. Batista, R. A. C. Garcia, S. C. Oliveira, A. Trindade, C. M. Viana, J. Rocha

This work has been developed under the Science4Policy 2024 (S4P-24), an annual Science for Policy Project call, an initiative promoted by Centre for Planning and Evaluation of Public Policies in partnership with the Foundation for Science and Technology (FCT), financed by Portugal Recovery and Resilience Plan. ML-SOIL - Project Co-participative Modelling of Soil Districts based on Machine Learning (2024.00178.S4P24)

Table of contents

01 Introduction

- Proposal for a directive on Soil Monitoring and Resilience (Soil Monitoring Law) by the European Union
- Delineation of Soil Districts – Its concept and importance
- The use of GIS tools – Machine learning algorithms, Statistical approaches and Cluster analysis
- Harmonization of monitoring systems at the national level (Portugal)

02 Methodology

To generate random sampling points:

- **Bethel Algorithm** provided by the European Union to choose sampling locations

To delineate Soil Districts:

- **Clustering Algorithms** (Tested):
  - Random Forest with MDS and Fuzzy k-means
  - Agglomerative Hierarchical Clustering
  - Fuzzy c-means
  - DBSCAN

Variables: Administrative and Geographic Divisions

Bethel Algorithm:

- NUTS II
- NUTS II combined with environmental zones
- APA soil map
- Landscape Units

Used as combination with all of the above:

- Corine Land Cover
- COS 2018

Clustering Algorithms:

- NUTS II
- NUTS II combined with Environmental Zones
- APA soil map
- Landscape Units
- NUTS II combined with Landscape Units
- NUTS II combined with Corine Land Cover

Variables: Soil Properties and Climatic Variables

- Bulk Density at three depths: 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 depths: 0–10 cm, 10–20 cm and 20–30 cm;
- Positive Precipitation;
- Maximum Temperature;
- Minimum Temperature.

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

- Minimum
- Maximum
- Mean
- Standard Deviation

03 Results

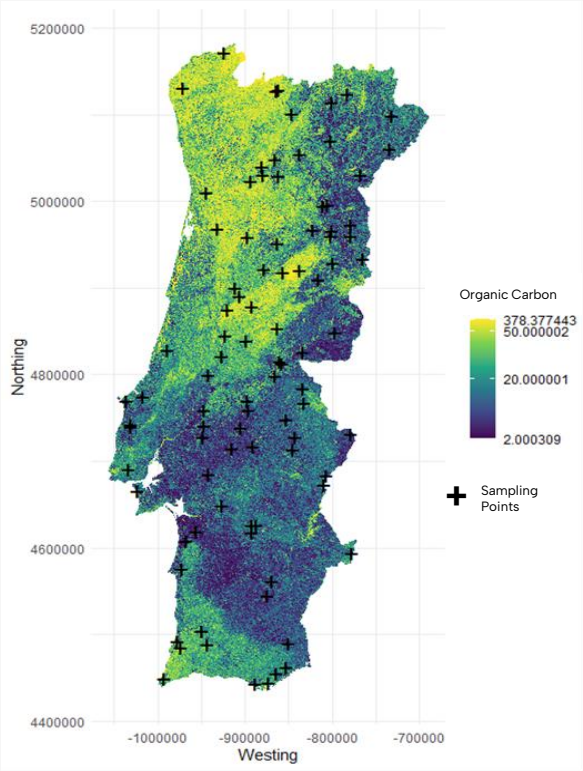
Bethel Algorithm

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

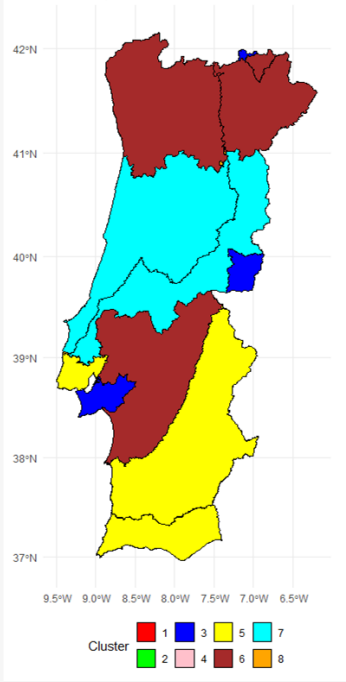
Best Result

Bethel Algorithm



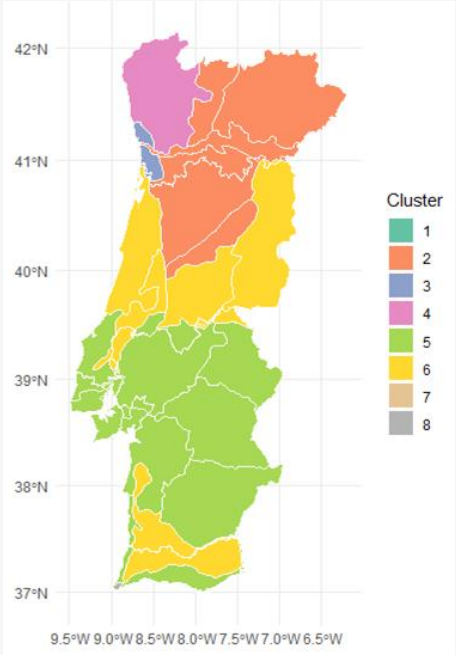
Clustering Algorithms

Random Forest



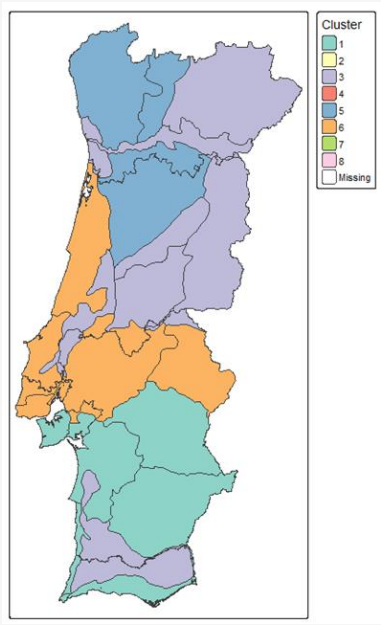
Nuts II intersected with Environmental Zones

Agglomerative Hierarchical Clustering



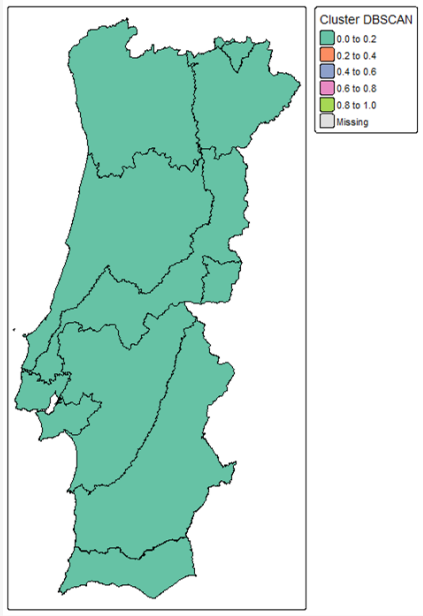
Nuts II intersected with Landscape Units

Fuzzy c-means



Nuts II intersected with Landscape Units

DBSCAN



Nuts II intersected with Environmental Zones

04 References

- Direção-Geral do Ambiente. (2023). Proposal for a Directive on Soil Monitoring and Resilience. https://environment.ec.europa.eu/publications/proposal-directive-soil-monitoring-andresilience_en
- European Comission. (2023). *Proposal for a Directive of the European Parliament and of the Council on Soil Monitoring and Resilience (Soil Monitoring Law)*. https://doi.org/10.2777/821504
- Kerry, R., Ingram, B., & Oliver, M. (2021). Sampling needs to establish effective management zones for plant nutrients in precision agriculture. Precision Agriculture'21, 653–660. https://doi.org/10.3920/978-90-8686-916-9_78
- Lawrence, P. G., Roper, W., Morris, T. F., & Guillard, K. (2020). Guiding soil sampling strategies using classical and spatial statistics: A review. Agronomy Journal, 112(1), 493–510. https://doi.org/https://doi.org/10.1002/agj2.20048
- Taghizadeh-Mehrjardi, R., Nabiollahi, K., Rasoli, L., Kerry, R., & Scholten, T. (2020). Land Suitability Assessment and Agricultural Production Sustainability Using Machine Learning Models. Agronomy, 10(4). https://doi.org/10.3390/agronomy10040573
- Weninger, T., Ramler, D., Bondi, G., Asins, S., O'Sullivan, L., Assennato, F., Astover, A., Bispo, A., Borůvka, L., Buttafuoco, G., Calzolari, C., Castanheira, N., Cousin, I., van den Elsen, E., Foldal, C., Hessel, R., Kadžiulienė, Ž., Kukk, L., Molina, M. J., … Klimkowicz-Pawlas, A. (2024). Do we speak one language on the way to sustainable soil management in Europe? A terminology check via an EU-wide survey. European Journal of Soil Science, 75(2), e13476. https://doi.org/https://doi.org/10.1111/ejss.13476
- Zeraatpisheh, Mojtaba, Bakhshandeh, E., Emadi, M., Li, T., & Xu, M. (2020). Integration of PCA and Fuzzy Clustering for Delineation of Soil Management Zones and Cost-Efficiency Analysis in a Citrus Plantation. Sustainability, 12(14). https://doi.org/10.3390/su12145809.