Using Artificial Neural Networks for Digital Soil Mapping – a comparison of MLP and SOM approaches

Sérgio Freire
e-GEO, FCSH, Univ. Nova de Lisboa
Av. de Berna 26-C
Lisboa, Portugal
sfreire@fcsh.unl.pt

Inês Fonseca
Centro de Estudos Geográficos - IGOT
Alameda da Universidade
Lisboa, Portugal
i.fonseca@campus.ul.pt

Ricardo Brasil
Centro de Estudos Geográficos - IGOT
Alameda da Universidade
Lisboa, Portugal
rmsb@campus.ul.pt

Jorge Rocha
Centro de Estudos Geográficos - IGOT
Alameda da Universidade
Lisboa, Portugal
jorge.rocha@campus.ul.pt

José A. Tenedório
e-GEO, FCSH, Univ. Nova de Lisboa
Av. de Berna 26-C
Lisboa, Portugal
ja.tenedorio @fcsh.unl.pt

Abstract

Portuguese soil map coverage remains incomplete, while the existing cartography has some shortcomings. Artificial Neural Networks (ANN) are advanced computer-based techniques which are being used for Digital Soil Mapping (DSM). These techniques allow mapping soil classes in a cheaper, more consistent and flexible way, using surrogate landscape data. This work compares the performance of two ANN approaches, Multi-layer Perceptron (MLP) and Self-Organizing Map (SOM), for DSM. The tests were carried out in IDRISI Taiga for three catchments in Portugal and one in Spain, using different sampling designs to obtain the training sets. Results reveal that best ANN performance is obtained with a MLP model rather than a SOM model, independently of data transformation and sampling method. However, MLP is also the most sensitive method to the data used to develop the models.
Keywords: Digital Soil Mapping, MLP, SOM, IDRISI Taiga, AutoMAPticS, Portugal.

1 Introduction

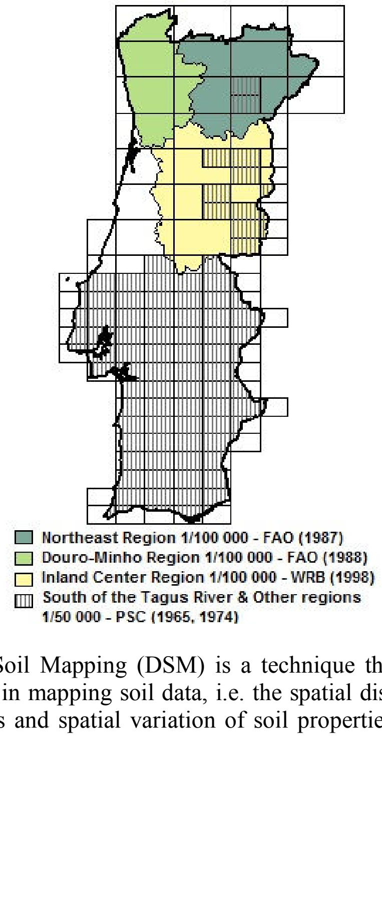
Soils play a fundamental role in sustainable land use by supporting valuable services, such as biodiversity, food production, and pollution buffering. Vital human activities depend on this important non-renewable resource [1].

The absence of soil information contributes to the uncertainties of predicting food production, and lack of reliable and harmonized soil data has considerably limited environmental impact assessments, land degradation studies and adapted sustainable land management interventions [2].

Despite soil surveys having been carried out in many countries, the scale and spatial coverage of many conventional soil maps are not ideal for planning applications at national level [3]. There is also a lack of consistency across countries concerning soil classifications and legends, which hampers the necessary integration of soil datasets, even in Europe [4]

Portugal, like most European Union countries, has only part of its territory covered with soil maps at semi-detailed or reconnaissance scales [5]. While 55% of continental Portugal has soil maps at 1:50 000 produced by traditional methods of soil survey before the 1970s, only about 40% of the territory has more recent soil map coverage at 1:100 000 with some overlap (Figure 1). Therefore, in addition to the published coverage remaining incomplete, significant problems remain regarding the existing cartography, namely concerning cartographic uniformity and taxonomic systems adopted [6].

Figure 1: Scale and legends of regional soil maps of continental Portugal.



Digital Soil Mapping (DSM) is a technique that has been successful in mapping soil data, i.e. the spatial distribution of soil classes and spatial variation of soil properties [3]. DSM may combine Geographical Information Systems (GIS) with advanced techniques such as Artificial Neural Networks (ANN), which have enabled mapping the spatial distribution of soils in a cheaper, more accurate, reproducible, and flexible way in terms of data storage and visualization, using surrogate landscape data easy to obtain. Thus, ANN modelling is able to provide the means to predict soil types at locations where there are no current soil maps, by combining soil map data from other areas with landscape features known to be responsible for the spatial variation of soils [5]. ANNs are therefore sophisticated computer programs able to model complex functional relationships, which applied to the soil mapping problem use a set of variables related to soil forming factors and the respective soil type as training data in order to constructs rules [7] that can be extended to the unmapped areas.

Although the usefulness of ANNs for DSM has been demonstrated (e.g., [8, 9, 10]) more needs to be known about their relative advantages and limitations. Specifically, there remains a lack of research on the performance of different ANN architectures and related aspects for soil spatial modelling at different scales and in different conditions. While MLP and SOM are probably the most frequently-used ANNs for DSM, these networks utilise different classification approaches. Therefore it is important to investigate their capabilities and limitations for DSM using different experimental settings.

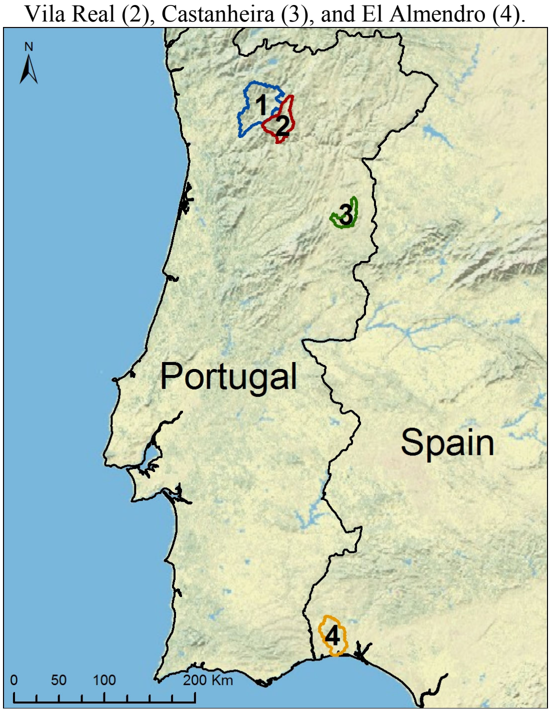
Sarmento et al. [11] compared three ANNs and a decision tree approach for mapping soil classes at a detailed scale, concluding that the latter method and a Multi-Layer Perceptron (MLP) network had similar and the best performances, obtaining the highest overall accuracies. Albuquerque et al. [12] have tested MLP and Kohonen's Self-Organizing Map (SOM) models in the context of image analysis, obtaining better results with MLP.

The present work reports the results of using different supervised ANN methods and experimental setups for DSM at regional scale in selected study areas in Portugal and Spain.

2 Study areas

The methodology has so far been applied in four medium-sized catchments: three in Portugal (Mondim de Basto, Vila Real, Castanheira) and one in Spain (El Almendro) (Figure 2).

Figure 2: Location of the study areas: Mondim de Basto (1), Vila Real (2), Castanheira (3), and El Almendro (4).



These catchments were selected for two main reasons: (a) their location in areas where soil maps at 1:100 000 are available, and (b) the fact that they display varied geomorphological and ecological features and include soil types representative of those occurring in the respective regions. Their main geographic characteristics are summarized in Table 1.

While Mondim de Basto is the largest catchment, its soil map (1:100 000) comprises only four classes, whereas Castanheira includes seven, despite being the smallest.

Table 1: Main characteristics of the selected catchments.

Catchment	Region	River	Area (km^2)	Min. elev. (m)	Max. elev. (m)	No. of soil classes
Mondim de Basto	Douro-Minho	Tâmega	911	56	1298	4
Vila Real	Northeast	Corgo	468	67	1405	4
Castanheira	Inland Center	Rib. das Cabras	227	558	969	7
El Almendro	Andalucía (Spain)	Rio Piedros	517	1	361	10

3 Data and methods

3.1 Datasets

The ANNs were trained using independent variables, which included both continuous terrain data and categorical (thematic) geoinformation. The terrain surrogate data were derived from the Shuttle Radar Topography Mission (SRTM) digital elevation data (www2.jpl.nasa.gov/srtm) with a 90 m resolution and selected after multicollinearity tests showed little data redundancy. Seven morphometric variables, which are frequently used in DSM, were extracted from the terrain data: slope steepness, plan and profile curvatures, upslope catchment area, dispersal area, wetness index and potential solar radiation. These continuous variables were rescaled to a 0-255 value range. In addition to altitude, land use from Corine Land Cover 2006 (CLC2006) and geological data were also included, as well as existing digital soil data at 1:100 000 (Mondim de Basto, Vila Real, e Castanheira), and 1:400 000 (El Almendro). All layers were clipped to the study areas and converted to a raster structure with a 90-m cell size.

In order to test for the possible effects of spatial autocorrelation which is commonly high in datasets derived from altitude data, two sets of independent variables were used for network training and classification. One set included the coordinates (latitude and longitude), in addition to the ten input variables indicated above.

3.2 Methodology

To model the spatial distribution of soil classes in each catchment, two types of ANN architectures were employed: the Multi-Layer Perceptron (MLP) and Kohonen's Self-Organizing Map (SOM). These were run in hard classification mode, using IDRISI Taiga software (Clark Labs).

MLP is a widely used supervised method based on the back-propagation learning algorithm [13]. Such a network typically comprises an input layer, one output layer, and at least one intermediary hidden layer. Independent variables are nodes in the input layer, while final classes result as neurons in the output layer. The experimental setup for each training set in MLP shifted from the default specifications presented in Table 2 used as initial values.

Table 2: Default characteristics and parameters of the ANN MLP in IDRISI Taiga.

Group	Parameter	Default value
Input specifications	Avrg. training pixels per class	200 / 250
	Avrg. testing pixels per class	200 / 250
Network topology	Hidden layers	1
Training parameters	Automatic training	No
	Dynamic learning rate	Yes
	Learning rate	0.01
	End learning rate	0.001
	Momentum factor	0.5
	Sigmoid constant "a"	1
Stopping criteria	RMS	0.01
	Iterations	(Variable)
	Accuracy rate	100%

The SOM model in IDRISI Taiga is based on Kohonen's Self-Organizing Map [14, 15]. The basic architecture includes a layer of input neurons connected by synaptic weights to neurons in an output layer arranged in a two-dimensional (usually square) array. The process begins with a coarse tuning phase that is in effect a form of unsupervised classification. In a subsequent fine tuning stage, intra-class decision boundaries are refined using a Learning Vector Quantization (LVQ) procedure. The default parameters of SOM are presented in Table 3.

Table 3: Default characteristics and parameters of the ANN SOM in IDRISI Taiga.

Group	Parameter	Default value
Sampling in band images	Interval in column	3
	Interval in row	7
Network parameters	Output layer neuron	15x15= 225
	Initial neighborhood radius	22.21
	Min learning rate	0.5
	Max learning rate	1
Fine tune parameters	Min gain term	0.0001
	Max gain term	0.0005
	Fine tuning rule	LVQ2
	Fine tuning epochs	50
Classification Specification	Output hard classification map	Y
	Display feature map	Y
	Algorithm for unknown pixels	Min Mean Distance

Regarding sampling of soil classes for training the networks, the potential impact due to sampling scheme was not yet well addressed in the existing literature. Considering the significance of spatial autocorrelation present in both the distribution of environmental variables and resulting soil classes, different sampling schemes were used, in addition to testing the inclusion of latitude and longitude in the variable set for each watershed. Therefore the ANNs were trained by presenting them with a number of different examples of the same soil type drawn (i) randomly (RS), or (ii) in a stratified fashion (SS). For the latter, training pixel vectors were located by choosing (a) random coordinates within soil types strata (SRS), (b) random coordinates within soil types and chosen evenly in the frequency space (SRPS), (c) nearest coordinates within soil types and chosen evenly in the frequency space (SNPS), and (d) farthest coordinates within soil types and chosen evenly in the frequency space (SFPS) [16, 17].

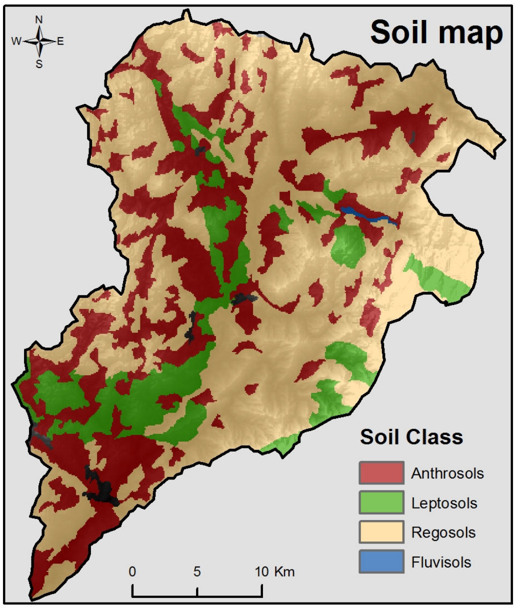
In MLP, an average of 250 pixels per class were used for training and testing in the Vila Real and El Almendro catchments, while 200 were used for Mondim and Castanheira, due to constraints in the total area covered by some soil classes. Some of the network parameters were progressively changed and the network performance monitored, namely: number of layer 1 nodes, use of automatic training, use of dynamic learning rate, and number of iterations (maximum of 100 000). Training ended when one of the stopping criteria was achieved: either a RMSE ≤ 0.01, an accuracy of 100%, or the defined maximum number of iterations. Therefore the default neural network included 10 or 12 input layer nodes, and one hidden layer with 7 nodes (see Table 2).

In a given study area, for a specific combination of sampling method and parameters, results of different runs can vary due to different seeding of training pixels. Thus, five model runs were initially performed for each combination, in order to average their accuracies and derive the best parameters.

The quality of estimated maps was assessed in Map Comparison Kit (MCK) v3.2.2. software (Geonamica), using the conventional soil maps as reference information. For each study area and model run, overall accuracy was computed from the contingency table (error matrix), as the percentage of agreement between the classified map and the reference map.

Figure 3 shows the 1:100 000 conventional soil map for Mondim de Bastos used to extract training samples and for accuracy assessment.

Figure 3: Soil map for Mondim de Bastos catchment.



4 Results and discussion

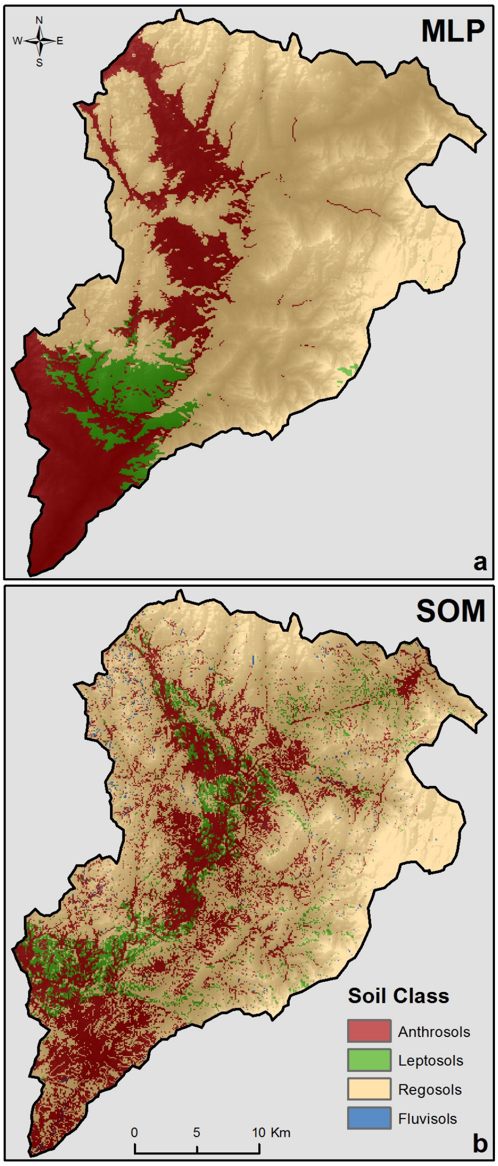
The results obtained using ANNs for soil mapping in the four study areas are presented in Table 4. For each catchment and for each of the ANN methods employed, the sampling strategy which obtained the highest accuracy, using 12 and 10 variables, is shown. This is shown only for the combination of ANN model and variable set obtaining the highest overall accuracy, as computed in MCK.

Table 4: Impact of ANN method on the performance of ANN models.

Catchment	ANN Method	Sampling	No. of variables	Accuracy (%)
Mondim de Basto	MLP	RS	12	67.9
	SOM	RS	10	64.8
Vila Real	MLP	RS	12	74
	SOM	RS	10	72.6
Castanheira	MLP	RS	12	60.9
	SOM	RS	12	55.2
El Almendro	MLP	RS	12	74.3
	SOM	RS	12	72.2

In general, accuracies are rather high, with best simulations obtained for El Almendro´s catchment, which also has the largest number of soil classes. The lowest accuracy values were obtained in Castanheira, the smallest catchment. It is possible that some differences are due to (1) scale of soils maps being rather different (1:100 000 vs 1:400 000, for Portuguese and Spanish datasets respectively) and (2) the number of soil classes. For instance, Castanheira catchment has fewer soil classes which are more likely to include different types of soils, whilst in El Almendro the opposite happens: higher number of classes which are also pre-defined as a combination of soil types. The latter approach to soil classification may result in soil patches that reflect better the differences in the landscape, i.e. higher purity of soil mapped patches, thus resulting in better ability of the ANN to classify correctly.

Figure 4: Modelled soil class distribution in the Mondim de Basto catchment, using MLP and SOM.



Even though specific accuracy values differ for each catchment, results show that MLP consistently outperformed SOM in all study areas, although the level of accuracy varied only between 1.4% (Vila Real) and 5.7% (Castanheira). Additionally, in MLP higher accuracies were obtained when location variables were included, while SOM performed generally better if latitude and longitude were not included as independent variables. In all catchments, best results were obtained using random sampling for selecting training sites.

Inasmuch as MLP emerges as a robust method for this purpose, it sometimes suffers from the problem of falling into a local minimum, resulting in a poor prediction when the whole study area is analyzed. Therefore it is important to average accuracies while deriving the best parameters and monitor the classification process closely. The algorithm used by SOM model has the benefits of being faster and easier to use.

The spatial distribution of soil classes obtained with each of the ANN architectures, in the Mondim de Basto catchment, is illustrated in Figure 4. The figure shows that MLP (a) maps soil classes in larger patches, while SOM's prediction (b) is more detailed and 'pixelated'. It is possible that the higher degree of agreement between the MLP map and the reference map may result from the scale-dependent cartographic generalization of the latter, and/or that it is necessary to use a set of indices that reflect not only the degree of overlap, but also the size and location of patches, as well as the number of classes correctly predicted. Due to the resulting SOM maps being very fragmented (i.e., 'pixelated'), it may be advisable to re-asses after conducting a spatial generalization. Future experiments using both soil data and environmental variables at different resolutions and a set of evaluation indices will allow testing these hypotheses.

In this work, ANN-modelled soil maps were compared to conventional soil maps, despite the inexistence of accuracy or confidence measures for these maps. Future work will use available soil survey data for a true validation of the models.

5 Conclusions

Conventional mapping of soils is based on labour-intensive field surveys, which are very costly. While the demand is rising for high-resolution spatial soil information for modelling and environmental planning, Portugal still lacks complete soil-map coverage. Artificial Neural Networks are powerful techniques for Digital Soil Mapping that can model high-quality digital soil maps in a fast and cost-effective way. However, there exist different ANN model architectures and there is a shortage of studies testing and comparing their performance for various problems, including the spatial prediction of soil classes at a regional scale.

This research tested the modelling of soil classes using two ANN approaches in four catchments in Portugal and Spain and compared the results. Noteworthy conclusions are that 1) the modelling accuracy of the ANN models in supervised mode is highly dependent on the sampling method used to select training sites and, 2) the best ANN performance was obtained with a MLP model rather than a SOM model, independently of sampling method. However, MLP is also the most sensitive method to the data used to develop the models and the soil class patterns it predicted produced more compact and fewer patches than the original soil maps.

Ongoing work involves classifying soils using ANNs at different spatial resolutions, for testing sensitivity to mapping scales. Subsequent research will explore the application of Fuzzy Logic for the evaluation of accuracy levels, as well as for producing hybrid ANNs, and results obtained using both methodologies will be compared and validated using existing maps and soil profile data. The best model will be used to map soil classes across areas which are currently lacking spatial soil data, ultimately enabling the completion of the Portuguese soil map coverage at 1:100 000.

Acknowledgements

The authors thank DRAEDM for providing digital soil data. This work was produced in the framework of AutoMAPticS (Automatic Mapping of Soils), a project supported by a grant from FCT - Portugal (PTDC/CS-GEO/111929/2009), whose PI is employed under the FCT Science 2008 Programme.

References

[1] J. Potocnik and S. Dimas. Preface In *Soil Atlas of Europe*. European Soil Bureau Network. European Commission, Office for Official Publications of the European Communities, L-2995 Luxemburg, 2005.

[2] A.J. Muller and S. Nilsson. Foreword to FAO/IIASA/ISRIC/ISS-CAS/JRC, Harmonized World Soil Database (version 1.1). FAO, Rome, Italy and IIASA, Laxenburg, Austria, 2009.

[3] E. Dobos, F. Carré, T. Hengl, H.I. Reuter, and G. Tóth. *Digital Soil Mapping as a Support to Production of Functional Maps*, EUR 22123 EN. Office for Official Publications of the European Communities, Luxemburg, 2006.

[4] ESBN (European Soil Bureau Network). *Soil Atlas of Europe*. European Commission, Office for Official Publications of the European Communities, L-2995 Luxemburg, 2005.

[5] A.B. Mcbratney, M.L. Mendonça Santos, and B. Minasny. On digital soil mapping. *Geoderma*, 117:3-52, 2003.

[6] I. Fonseca, R. Brasil, J. Rocha, S. Freire, and J. A. Tenedório. Soil maps – turning old into new. In *Proceedings of the 7th European Congress on Regional Geoscientific Cartography and Information Systems*, vol. 1:146-147, 2012.

[7] B. Tso and P.M. Mather. Classification Methods for Remotely Sensed Data. Taylor and Francis, London, 2001.

[8] A.-X. Zhu. Mapping soil landscape as spatial continua: The neural network approach. *Water Resources Research*, 36(3): 663-677, 2000.

[9] T. Behrens, H. Forster, T. Scholten, U. Steinrucken, E. Spies, and M. Goldschmitt. Digital soil mapping using artificial neural networks, *J. Plant. Nutr. Soil Sci*. 168:1-13, 2005.

[10] W. Carvalho Junior et al. Digital soilscape mapping of tropical hillslope areas by neural networks. *Sci. Agric. (Piracicaba, Braz.)*, 68(6):691-696, 2011.

[11] E. C. Sarmento, E. Giasson, E. Weber, C. A. Flores, and H. Hasenack. Comparison of four machine learning algorithms for digital soil mapping in the Vale dos Vinhedos, RS, Brasil. In International Workshop on Digital Soil Mapping, 4. Anais. Roma: CRA-RPS, 2010.

[12] V.H. Albuquerque, A.R. Alexandria, P.C. Cortez, and J.M. Tavares. Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images, NDT & E International, 42(7):644-651, 2009.

[13] S. Haykin. Neural Networks – a Comprehensive Foundation. Prentice Hall, New Jersey, 1999.

[14] T. Kohonen. The Self-Organizing Map. In Proceedings of the IEEE, 78:1464-80, 1990.

[15] T. Kohonen. Self-Organizing Maps, first ed. Springer, Berlin–Heidelberg, 1995.

[16] R. Brasil, I. Fonseca, S. Freire, J. Rocha, and J. A. Tenedório. Digital soil mapping using artificial neural networks – sampling issues. In *Proceedings of the 7th European Congress on Regional Geoscientific Cartography and Information Systems*, vol. 1:140-141, 2012.

[17] S. Freire, I. Fonseca, R. Brasil, J. Rocha, and J. A. Tenedório. The importance of sampling for the efficiency of artificial neural networks in digital soil mapping. In *Proceedings of the XIII Iberian Colloquium of Geography*, Santiago de Compostela, 2012.