

**Discrimination of soils and assessment of soil fertility using
information from an ISE array and artificial neural networks**

Aitor Mimendia¹, Juan M. Gutiérrez², Josep M. Alcañiz³, Manel del Valle^{1,*}

¹ *Sensors and Biosensors Group, Department of Chemistry, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain*

² *Bioelectronics Section, Department of Electrical Engineering, CINVESTAV, 07360 Mexico D.F., Mexico*

³ *CREAF (Centre for Ecological Research and Forestry Applications) and Department of Animal and Plant Biology and Ecology, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain*

* corresponding author: e-mail: Manel.delValle@uab.cat

Tel: +34-93-5813235

Fax: +34-93-5812477

Running title: Discrimination of soils and assessment of soil fertility

Keywords: soil fertility, soil sensing, organic matter, ion selective electrodes, neural networks

This is the pre-peer reviewed version of the following article: Mimendia, A. et al. "Discrimination of soils and assessment of soil fertility using information from an ion selective electrodes array and artificial neural networks" in Clean, vol. 42, issue 12 (Dec. 2014), p. 1808-1815, which has been published in final form at DOI 10.102/clen.201300923. This article may be used for non-commercial purposes in accordance with Wiley terms and conditions for self-archiving

Abstract

Multichannel sensor measurements combined with advanced treatment is the departure point for a new concept in sensorics, the electronic tongue. Our proposal worked with an array of 20 potentiometric sensors plus an Artificial Neural Network (ANN) used as a pattern recognition method and applied to soil analysis. With this design, we got a versatile tool which was able to perform qualitative and quantitative determinations. As first application, the qualitative discrimination between 6 distinct soil types based on their extractable components was attempted. The procedure was simplified to a single extraction step before measurements. Water, a BaCl_2 saline solution and an acetic acid extract were evaluated as extracting agents. The best performance was reached with the acid extraction method with a correct classification rate of 94%, sensitivity of 94% and a specificity of 100%. In addition, a quantitative determination of several physicochemical properties of agricultural interest, such as organic carbon content and selected cations (like K^+ or Mg^{2+}) and anions (like NO_3^- or Cl^-) is also demonstrated, showing satisfactory agreement with the reference methods.

1. Introduction

Soil analysis is a crucial step in many areas, for example in soil fertility studies, where detailed soil information is important to make decisions in precision-agriculture management; also for soil mapping, land planning, and for civil engineers who need to choose excavation zones. Precision, speed, cost and the possibility to process many samples are important features expected of modern techniques for soil analysis. In addition, reduced sample size, simplicity of sample preparation and multidetermination capabilities are some other desirable characteristics.

Many analytical techniques and methods are used in classification or assessment of soils. These analyses have to deal with two common difficulties related to environmental analysis: the strong variation observed between samples of the same category or class but from different environmental compartments due to natural heterogeneity of the sample, and the common appearance of some parameters between different sample groups or classes [1]. To overcome these difficulties, classical soil classification methods require a significant analytical effort in order to determine the large number of physical and chemical properties necessary to fulfil the classification; usually, this ultimate step depends on an evaluation by an expert. For these reasons it is advantageous to develop quicker and more efficient classification methods. For major soil components, use of direct spectroscopic methods as near-infrared (NIR) [2] is already employed; for availability of elements however, a chemical extraction method is currently required, whereas their extracts are subsequently analyzed by a variety of techniques.

In the recent years, a new strategy in analytical science has appeared: the electronic tongue (ET). It entails using an array of non-specific sensors in concert with advanced chemometric data processing tools able to interpret and extract meaningful data from the complex signals, and to relate them with their associated chemical nature [3]. The idea behind this concept is to use an appropriate sensor array with some cross-

sensitivity features, which allows the simultaneous determination of a large number of species, while the chemometric treatment of the data permits the resolution of the interferences, drifts or non-linearities originated by the sensors [4]. This work follows this approach by conjugating an array formed by potentiometric sensors with generic cross-response to several ions, along with artificial neural networks (ANNs) being used as pattern recognition method. Although the use of potentiometric sensors in soil analysis is not new [5-7], and pattern recognition methods [8-10] or ANNs [11] have been applied to the field, the use of electronic tongues is complete novelty. In fact, specialized pattern recognition methods of different types have been coupled with conventional analytical techniques to perform soil discrimination and estimation of some parameters [12-14], but never with the use of sensors.

Research exploiting ETs has been an active and dynamic research field in the late years. Several techniques, mainly electrochemical (voltammetry, electrochemical impedance, potentiometry) have been used in a wide number of applications [15], also incorporating biosensors based on enzymes. The pattern recognition counterpart offers a variety of options too, from the partial least squares (PLS) to the k-nearest neighbour (KNN) techniques, including ANNs or support vector machines [16].

As can be appreciated from the literature, the multidisciplinary character of the ET technique promotes a continuous evolution and adaptation of advances from mathematics, electronics, chemistry or biology. Apart from the qualitative applications regarding the identification of sample varieties using a particular sensors' signature, ETs can be also employed for quantitative purposes; taking advantage from the cross-response generated by the sensor array, the ANN may solve interference problems without requiring special procedures to mask or eliminate the interfering species. Another benefit is the possibility of modelling certain parameters, like the organic matter content, for which no specialized sensors is available. Several examples of this situation have already been described: the prediction of natural orange juice content in orange-based drinks [17], the bitterness or the alcoholic strength in beers [18] or the

ageing of wines [19]. Analogue applications related to fertilizing ions have been described also in the agricultural field [20, 21]. This could be accomplished with the help of the abilities of ANNs to generate models correlating sensor responses with parameters indirectly linked with measured species [7].

The aim of this work, then, is to present the tool known as the electronic tongue, and to illustrate its capability to carry out the discrimination of soil types and the quantitative estimation of selected soil fertility parameters. In our case, in order to assess ions bioavailability, the followed methodology compares three extraction procedures: a soil water-extract that simulates leaching processes occurring in the field, a saline solution extract able to mobilize exchangeable elements and the more aggressive conditions of an acidic extract. In all cases, analytical data has been compared against reference analytical procedures in order to contrast the obtained results. The use of such a tool might be of great help in diagnosis of soil status, management of crop, or exhausted soil remediation [22, 23].

2. Materials and Methods

2.1. Soil samples and extraction method

A selection of six different soils mainly based on their texture, pH, organic matter and lime content were selected for the study, which is an initial stage of the electronic tongue application to soil analysis. All came from Catalonia (NE Spain), from zones of different climate and parent material. Table 1 shows the main composition of these soils and their classification. Studied samples belonged to surface horizons (A), except for histosol VI that was taken at a depth of 40-80cm. Soil texture range from sandy to clay loam; pH from very acidic to ca. 8 in those dominated by carbonates; organic carbon ranged from 1% in soils of xeric moisture regime to 16% in the histosol.

A total of 6 samples from each soil and equivalent location were air dried and sieved to 2mm. A very simple sample preparation procedure that consisted on a unique

1
2
3 1:5 (w:v) liquid extraction, e.g 10 g of soil and 50 mL of extractant, using a plastic
4 sterile container was carried out. Three different extracting agents were compared for
5 the study: doubly distilled water, acetic acid (0.4 mM) and barium chloride (0.01 M) in
6 order to later evaluate their performance. The extraction process consisted in 1 h mixing
7 step by an orbital shaker and 30 additional min required for the sedimentation of the
8 soil. After this sedimentation, potentiometric measurement with electronic tongue were
9 done without any filtering stage, 5 min after dipping the sensor array in the sample,
10 enough time for the potential values to stabilise. All this extraction and measuring
11 process were replicated 6 times, each time using a new soil sample, in order to check the
12 robustness of the method in front of any heterogeneity in the soil.
13
14
15
16
17
18
19
20
21
22
23
24

25 2.2. Chemicals

26
27 Instead of purchasing them, the ion-selective poly(vinyl chloride) (PVC) membranes
28 (Table 2) were prepared in the laboratory for a larger variety. These were formulated
29 using high-molecular weight PVC (Fluka) and bis(1-butylpentyl)adipate (BPA), o-
30 nitrophenyloctyl ether (o-NPOE), dioctylsebacate (DOS), dibutylsebacate (DBS),
31 dibutyl butyl phosphonate (DBBP), dioctylphenylphosphate (DOPP) and dibutyl
32 phtalate (DBP) as plasticizers (all from Fluka). The recognition elements employed to
33 formulate the potentiometric membranes were the ionophores nonactin (ammonium
34 ionophore I, Fluka), valinomycin (potassium ionophore I, Fluka), bis[(12-crown-
35 4)methyl]-2-dodecyl-2-methylmalonate (CMDMM, Dojindo Laboratories, Kumamoto,
36 Japan), 1,3-[bis(3-phenylthioureidomethyl)benzene (Sulphate Ionophore I, Fluka), tri-
37 N-dodecylamine (TDDA, hydrogen ionophore I, Fluka), the sodium salt of the
38 antibiotic tetronasin (provided by the University of Cambridge [24]), ETH1001 (Fluka),
39 bis(bis(4-1,1,3,3-tetramethylbutyl)phenyl)phosphate calcium salt (BBTP) (Fluka),
40 ETH4030 (Fluka), monensin sodium salt (Acros), N,N,N0,N0-Tetrabutyl-3,6-
41 dioxaoctanedi(thioamide) (Cadmium Ionophore I), o-Xylylenebis(N,N-
42 diisobutylidithiocarbamate) (Copper(II) Ionophore I), S,S'-Methylenebis(N,N-
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

diisobutyldithiocarbamate) (Lead Ionophore II), tert-Butylcalix[4]arene- tetrakis(N,N-dimethylthioacetamide) (Lead Ionophore IV), Tetrabutylthiuram disulfide (Zinc Ionophore I), 3,7,12,17-Tetramethyl-8,13-divinyl-2,18-porphinedipropionic acid disodium salt (Zinc Ionophore II) and the charged carriers tridodecylmethylammonium chloride (TDMAC, Fluka), 1,3-[Bis(3-phenylthioureidomethyl)]benzene (Sulphate Ionophore I, Fluka) and tetraoctylammonium bromide (TOAB, Fluka). Additionally, two generic response recognition elements were used: dibenzo-18- crown-6 (Fluka) and lasalocid sodium salt (Fluka), both for cations. All these formulations were used and reported previously by our research group [25-27].

Materials used to prepare the solid electrical contact were epoxy resin Araldite M and Hardener HR (both from Fluka) and graphite powder (50 μm , Merck), as a conducting filler. All other reagents used were of analytical grade, pro-analysis or the equivalent.

2.3. *Sensor array*

The sensors were all-solid-state ion selective electrodes (ISEs) with a solid electrical contact made from conductive epoxy-graphite composite. They were constructed by following standard procedures from our laboratory [28]. The PVC membranes were formed by solvent-casting the sensor cocktail dissolved in THF over the solid contact. The specific formulation of the different membranes is detailed in Table 2. Once formed, membranes were conditioned in a 0.1 M solution of their primary ion for 24h. Arbitrarily, we used ammonium chloride solution for conditioning the generic electrodes. The sensor array was composed by 20 sensors: one for ammonium, one for sodium, one for potassium, two with generic response for cations and one more generic for anions, one for chloride, one for sulphate, one for barium, one for nitrate, one for magnesium, three different for calcium, one for copper, for lead and zinc two different for each one and finally one for hydrogen ion. The sensor configuration used guarantee a proper stable response for at least three months [21], but normally may be used for a

longer time with adequate performance. The sensor array contained electrodes with certain correlated response, as the different ionophores for calcium, or sensors with strong cross- response, like the generic ones; from our experience, this fact helps in increasing the accuracy of the generated response models.

2.4 Apparatus

Potentiometric measurements were done using a laboratory made data-acquisition system. It consisted of 32 input channels implemented with amplifier-follower circuits employing operational amplifiers (TL071, Texas Instruments), which adapt the impedances of each sensor. Measurements were unipolar, with the reference electrode connected to ground. Measurements were referred to an Orion 90-02-00 double junction Ag/AgCl reference electrode. Each channel was noise-shielded with its signal guard. The outputs of each amplifier were filtered using a passive low-pass filter and connected to an A/D conversion card (PC-Lab 813, Advantech, Taiwan) installed into a Pentium PC. The readings were done employing custom designed software programmed with QuickBASIC 4.5 (Microsoft).

2.5 Reference determinations of soil physicochemical properties

Several soil properties were determined in order to evaluate predicting capabilities of the electronic tongue. These include whole soil parameters as oxidizable organic carbon (by dichromate) [29], pH in water (1:2.5 w/v) [30], total carbonates (Bernard calcimeter) [31], and particle size analysis (sedimentation, pipette method) [32]. From each extract used for ET measurements, an aliquot was reserved for parallel reference determinations of electrical conductivity [33], the content of several cations as ammonium, potassium, sodium, magnesium and calcium (determined in parallel by atomic spectroscopy), and also some anions as chloride, sulphate and nitrate (by

capillary electrophoresis). For determinations of conductivity and ions, extracts were filtered using a Whatman grade GF/F glass fibre filter for better reproducibility.

2.6 Software and data processing

The obtained data of electronic tongue measurements for each soil extract were processed and evaluated independently. The potentiometric measurements were done after reaching steady state responses (wait time 5 min), while 5 measurements were collected for each sensor in intervals of 10 s. Mean value of these measurements was calculated, obtaining one value per sensor and sample for further data analysis process. Sensor responses were autoscaled before applying multivariate statistical methods.

In the case of the soil classification application, the potentiometric measurements were firstly processed using principal component analysis (PCA). This is a multivariable analysis tool, which permits to compress the amount of data volume into a few variables (principal components) preserving the major part of the variability originally present. The visualization of the samples in the new coordinates is useful to verify any clustering, that may occur because of their similarity. Next, the PCA score values fed a feedforward backpropagation artificial neural network (ANN) used as the classifier, which had a three layer structure and used Bayesian regularization (BR) as training algorithm. This and the ANNs used to model the combined response of the sensor array were developed in MATLAB 7.1 (Mathworks, USA), with help of its Neural Network Toolbox 4.0.6 (Mathworks, USA).

3. Results and discussion

Both qualitative and quantitative applications were performed with a set of extracts from the six different types of soils selected for the initial stages of this research (see Table 1). Each soil belonged to a different taxonomic group [34, 35] and had

noticeable differences in their chemical composition or properties, as its discrimination would be expected. Nevertheless, some parameters were similar in some soils, as pH in III, IV and V soils, lime content in IV and V, organic carbon, salinity and particle size distribution in III and IV. Organic carbon content, salinity and silt proportion were the more discriminative parameters to differentiate the histosol from the others.

3.1 Qualitative application

Table 3 shows the concentration of the more abundant soluble cations and anions (except for hydrogencarbonate) in soil water-extracts, as determined by the reference methods. Low ion concentration was found in soil I (arenosol) according to their sandy texture and acidic reaction. Chloride is the more abundant anion, followed by sulphate, being these anions abundant (and also Na^+) in the histosol from Ebro Delta River that has saline conditions. Nitrate was noticeable in soil V that was cultivated few years ago. As expected, Ca^{2+} and Na^+ were the more abundant cations in soil water-extracts. High ammonium concentration was found in histosol probably coming from surrounding fertilized rice fields of Ebro Delta.

Figure 1 depicts a principal component analysis (PCA) loadings plot of sensor responses; this was done in order to visualize if all the sensors provided significant information to the system. Proximity to the centre – point (0, 0) – on the PCA graph of loadings would identify a sensor as uninformative. This criterion might be used to discard non useful sensors; in our case, as no sensor was placed in this central region, all of them were kept for the study. A second type of information that may be derived is the co-linearity of some of the sensors, e.g. whether they appear in close proximity. In this study, only the three Ca^{2+} sensors, or the NO_3^- and anion generic sensor showed correlated response; nevertheless all of them were used for the experiments, given only significant information extracted by PCA was further used for the classification. Apart,

it is significant how anions appeared on the right, and heavy metals on the left, illustrating their opposite nature response.

When soil responses were examined, in this case through the PCA graph of scores (see Figure 2), even in the worst case (water extractant) around 90% of variance of sensor readings was described by the three first principal components (PCs), demonstrating the multivariate interpretation is possible. The clustering of the replicas for each extraction method is also visible in the 3-D score plots (Fig. 2). Comparing the cluster distribution with the soil physicochemical properties, it seems in all cases the PC1 and PC2 plane direction is related with the organic matter content, silt proportion and pH, while the PC3 direction is more correlated with the concentration of cations. In addition, it has to be remarked that the clusters formed with the acid acetic soil extract data are more sharply defined than the others. This better definition among the soil representations (shown as the proximity of the replicates) proves the good repeatability of the proposed methodology and how sample heterogeneity is under control. Also, the simplest water extract well discriminated all samples and put closely soils III, IV and V (the same distribution was observed for the acid extract) which had a similar mineral composition (carbonates) and related properties. Larger dispersion is obtained from barium chloride extract that represents exchangeable cations of soils, probably indicating longer extraction times would have been desirable.

After this initial representation of data, an ANN was used as classifier. The aim was to predict, for a given sample, its belonging to some of the classes previously trained, in a pattern recognition application. In contrast, the projection of samples into the established PCA space already performs visual classification, although of subjective nature. The first step to perform the ANN classification is selecting the topology of the network, in this case built as a binary predictor of the soil class. This requires a significant effort of optimization due to the difficulties in defining the optimum ANN configuration in advance. A trial-error process is needed, where several parameters (training algorithms, number of hidden layers, transfer functions, etc.) are fine-tuned in

order to find the best configuration to optimize the performance of the response model. After an exhaustive evaluation, the optimum configuration had 3 neurons (the first three PC scores previously calculated) in the input layer, 10 neurons in the hidden layer and 6 neurons (one for each different type of soil) in the output layer. The transfer functions used were *logarithmic sigmoidal* in the three layers. The learning strategy used was Bayesian regularization and the expected output error was programmed to reach a value of 0.001.

The data was subdivided into two subsets, randomly selected within each soil; one subset was used for training, the other for comparing performance (testing); 66% of the total data (24 measurements) were used for building the model, and the other third (12 measurements) was used to evaluate the model's predictive ability. In this case, as the algorithm used is Bayesian regularization, no internal validation subset is needed [36]. The obtained results for the three different extraction methods are summarized on Table 4. It presents the confusion matrix, where the expected class is compared with the predictions of the PCA-ANN classifier. The lower classification rate was obtained for soils III and IV that are quite similar.

In order to calculate the classification reliability we took account of false positive and negative identification of each class. A repeated random subsampling validation method was used in order to reduce the variability related with the data portioning (each indicated result corresponds to the average of 10 training/testing splits). Best performance was reached with the acetic acid extraction method, showing a percentage of correct classification of 94% and a sensitivity of 94%. It also has to be remarked the 100% specificity value found, which means that no false positive identifications were obtained. In the case of doubly distilled water extractant the correct classification rate was 89%, the sensitivity 93% and the specificity 96%. Finally for the barium chloride the results were 80% as correct classification rate, 83% for the sensitivity and a specificity of 97%.

3.2 Quantitative application

The second part of the study was focused on the application of the electronic tongue system for quantitative prediction of different soil physicochemical properties, in this occasion organic carbon, two cations (K^+ and Mg^{2+}) and two anions (NO_3^- and Cl^-) as the example. A quantitative response model to directly estimate these properties was developed from a properly trained ANN, with sensors' readings as the inputs and property values, being provided by reference methods, as the outputs [37]. Three ANN models were built, one for each of the extracting agents, in order to compare final performance.

At this point, a training step was required to construct the model; to this aim, a random subsampling validation method was used. The data was divided with 2/3rd (24 samples) for the training of the ANN and the rest (12 external test samples) for evaluating the model's predictive ability. Comparison of predicted vs. expected values for the different physicochemical properties involved was done to check the ANN performance. The best configuration was chosen regarding the Root Mean Squared Error (RMSE) of the fit, plus the intercept, slope, and correlation coefficient of the comparison graphs (ideal values equal to 0.0, 1.0, 0.0 and 1.0, respectively) – In all cases, corresponding to the external test subset. The number of neurons used in the hidden layer and the transfer functions (linear and sigmoidal) were optimized iteratively [36]. Several parameters, as the training algorithm (Bayesian Regularization, BR), learning rate ($\alpha=0.1$) and the momentum ($\beta=0.4$) were set based on the group previous experience [26].

For each of the three extraction methods, one independent model was generated. In the case of the acid acetic and water extracts the model had 5 neurons and the transfer functions were: *logarithm sigmoidal* for the hidden layer, and *purelin* for the output one. For the barium chloride extracts the best performance was obtained with 6 neurons in the hidden layer, and the *tansigmoidal* transfer function both in the hidden layer and output layers. In this case, also a random subsampling validation method was used;

average results for testing of ten different training-test splits are reported in Table 5. As it can be seen, obtained comparison lines were indistinguishable to the ideal values for the three extraction methods. Comparing the three procedures, the RMSE values were lower when the extracting agent used was acetic acid, next water and barium chloride, in this order. This arrangement is in agreement with the qualitative application reliance on the extraction method described in Sec 3.1.

4. Discussion

An electronic tongue (ET) system to carry out qualitative and quantitative analysis of soils was developed and tested for the first time in a selection of soil samples representing a broad range of taxonomic groups. The proposed tool implies a very simple procedure with just one liquid-extraction step and a sensor array measurement, requiring at most one day to obtain final results. This represents a great saving of time and laboratory efforts compared with traditional taxonomic classification methodologies and also with the individual techniques used to determine physicochemical properties. Three different extracting agents were evaluated and compared. The best performance was obtained with the acetic acid extraction (94% sensitivity), but good results were also obtained for the other variants assayed (sensitivities over 80%). Our research demonstrated that some soil properties could be also estimated quantitatively. Again, acetic acid showed to be the best extracting agent. It has to be remarked that all the properties were estimated with the same procedure involving just one sequence of measurements, and without any other chemical step to mask or eliminate interference effects. To conclude, ET system has shown an interesting potential as a fast screening tool for soil fertility or survey studies.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

This work was supported by Spanish Ministry of Science and Innovation through project MCINN CTQ2010-17099 and by program ICREA Academia.

The authors have declared no conflict of interest.

For Peer Review

5. References

- [1] S. Dragovic, A. Onjia, Classification of soil samples according to geographic origin using gamma-ray spectrometry and pattern recognition methods, *Applied Radiation and Isotopes* 2007, 65, 218-224.
- [2] B. Stenberg, R.A. Viscarra Rossel, A.M. Mouazen, J. Wetterlind, in: L. S. Donald (Ed.), *Advances in agronomy*, Academic Press 2010, pp. 163-215.
- [3] Y. Vlasov, A. Legin, A. Rudnitskaya, C. Di Natale, A. D'Amico, Nonspecific sensor arrays ("electronic tongue") for chemical analysis of liquids: (iupac technical report), *Pure and Applied Chemistry* 2005, 77, 1965-1983.
- [4] A. Riul Jr, C.A.R. Dantas, C.M. Miyazaki, O.N. Oliveira Jr, Recent advances in electronic tongues, *Analyst* 2010, 135, 2481-2495.
- [5] S.J. Birrell, J.W. Hummel, Real-time multi isfet/fia soil analysis system with automatic sample extraction, *Computers and Electronics in Agriculture* 2001, 32, 45-67.
- [6] T.H. Gieling, G. van Straten, H.J.J. Janssen, H. Wouters, Ise and chemfet sensors in greenhouse cultivation, *Sensors and Actuators B: Chemical* 2005, 105, 74-80.
- [7] C. Di Natale, F. Davide, J.A.J. Brunink, A. D'Amico, Y.G. Vlasov, A.V. Legin, A.M. Rudnitskaya, Multicomponent analysis of heavy metal cations and inorganic anions in liquids by a non-selective chalcogenide glass sensor array, *Sensors and Actuators B: Chemical* 1996, 34, 539-542.
- [8] B. Bhattacharya, D.P. Solomatine, Machine learning in soil classification, *Neural Networks* 2006, 19, 186-195.
- [9] Y. Cal, Soil classification by neural network, *Advances in Engineering Software* 1995, 22, 95-97.
- [10] P. Scull, J. Franklin, O.A. Chadwick, The application of classification tree analysis to soil type prediction in a desert landscape, *Ecological Modelling* 2005, 181, 1-15.
- [11] D. Güçlü, S. Dursun, Amelioration of carbon removal prediction for an activated sludge process using an artificial neural network (ann), *CLEAN – Soil, Air, Water* 2008, 36, 781-787.
- [12] P.H. Fidêncio, R.J. Poppi, J.C. de Andrade, Determination of organic matter in soils using radial basis function networks and near infrared spectroscopy, *Anal. Chim. Acta* 2002, 453, 125-134.
- [13] Z. Ramadan, X.-H. Song, P.K. Hopke, M.J. Johnson, K.M. Scow, Variable selection in classification of environmental soil samples for partial least square and neural network models, *Anal. Chim. Acta* 2001, 446, 231-242.
- [14] M. Kovacevic, B. Bajat, B. Gajic, Soil type classification and estimation of soil properties using support vector machines, *Geoderma* 2010, 154, 340-347.
- [15] M. del Valle, Electronic tongues employing electrochemical sensors, *Electroanalysis* 2010, 22, 1539-1555.
- [16] P. Ciosek, W. Wroblewski, Sensor arrays for liquid sensing - electronic tongue systems, *Analyst* 2007, 132, 963-978.
- [17] J. Gallardo, S. Alegret, M. del Valle, Application of a potentiometric electronic tongue as a classification tool in food analysis, *Talanta* 2005, 66, 1303-1309.
- [18] Á.A. Arrieta, M.L. Rodríguez-Méndez, J.A. de Saja, C.A. Blanco, D. Nimubona, Prediction of bitterness and alcoholic strength in beer using an electronic tongue, *Food Chemistry* 2010, 123, 642-646.

- [19] V. Parra, Á.A. Arrieta, J.A. Fernández-Escudero, M. Íñiguez, J.A.d. Saja, M.L. Rodríguez-Méndez, Monitoring of the ageing of red wines in oak barrels by means of an hybrid electronic tongue, *Anal. Chim. Acta* 2006, 563, 229-237.
- [20] M. Gutiérrez, S. Alegret, R. Cáceres, J. Casadesús, O. Marfà, M. del Valle, Application of a potentiometric electronic tongue to fertigation strategy in greenhouse cultivation, *Computers and Electronics in Agriculture* 2007, 57, 12-22.
- [21] M. Gutiérrez, S. Alegret, R. Cáceres, J. Casadesús, O. Marfà, M. del Valle, Nutrient solution monitoring in greenhouse cultivation employing a potentiometric electronic tongue, *Journal of Agricultural and Food Chemistry* 2008, 56, 1810-1817.
- [22] D. Said-Pullicino, L. Massaccesi, L. Dixon, R. Bol, G. Gigliotti, Organic matter dynamics in a compost-amended anthropogenic landfill capping-soil, *European Journal of Soil Science* 2010, 61, 35-47.
- [23] T.R. Sinclair, J. Amir, A model to assess nitrogen limitations on the growth and yield of spring wheat, *Field Crops Research* 1992, 30, 63-78.
- [24] S.L. Less, P.F. Leadlay, C.J. Dutton, J. Staunton, Biosynthesis of tetronasin: Part 7. Preparation of structural analogues of the tetraketide biosynthetic precursor to tetronasin, *Tetrahedron Letters* 1996, 37, 3519-3520.
- [25] D. Calvo, M. del Valle, Simultaneous titration of ternary alkaline-earth mixtures employing a potentiometric electronic tongue, *Microchemical Journal* 2007, 87, 27-34.
- [26] M. Cortina, A. Duran, S. Alegret, M. del Valle, A sequential injection electronic tongue employing the transient response from potentiometric sensors for anion multidetermination, *Anal. Bioanal. Chem.* 2006, 385, 1186-1194.
- [27] A. Mimendia, J.M. Gutiérrez, L. Leija, P.R. Hernández, L. Favari, R. Muñoz, M. del Valle, A review of the use of the potentiometric electronic tongue in the monitoring of environmental systems, *Environmental Modelling & Software* 2010, 25, 1023-1030.
- [28] J. Gallardo, S. Alegret, M.A. de Roman, R. Munoz, P.R. Hernandez, L. Leija, M. del Valle, Determination of ammonium ion employing an electronic tongue based on potentiometric sensors, *Analytical Letters* 2003, 36, 2893-2908.
- [29] P.R. Hesse, *A textbook of soil chemical analysis*, John Murray, London 1971.
- [30] J. Porta i Casanellas, M. López-Acevedo Reguerín, R. Rodríguez Ochoa, C. Col·legi d'Enginyers Agrònoms de, L. Escola Tècnica Superior d'Enginyers Agrònoms de, *Técnicas y experimentos en edafología*, Col·legi Oficial d'Enginyers Agrònoms de Catalunya, Barcelona 1986.
- [31] M. Pansu, J. Gautheyrou, *Handbook of soil analysis*, Springer, Heidelberg 2006, pp. 593-604.
- [32] M. Pansu, J. Gautheyrou, *Handbook of soil analysis*, Springer, Heidelberg 2006, pp. 15-63.
- [33] H.D. Chapman, P.F. Pratt, *Methods of analysis for soils, plants and waters*, University of California, Riverside CA 1961.
- [34] United States Department of Agriculture (Ed.), *Keys to soil taxonomy*, Natural Resources Conservation Service, Washington DC 2010.
- [35] Working Group IUSS, *World soil resources reports*, FAO, Rome 2006, pp. 1-10.
- [36] H. Demuth, M. Beale (Eds.), *Neural network toolbox user's guide*, Mathworks, Natick MA 1994.
- [37] G. Hanrahan, Computational neural networks driving complex analytical problem solving, *Anal. Chem.* 2010, 82, 4307-4313.

Table 1. Taxonomic groups and environmental factors of selected soils in the study.

Soil n°	I	II	III	IV	V	VI
Horizon	A1	A1	A1	A1	Ap	H3
Depth (cm)	2-20	3-12	3-25	2-15	0-20	40-80
Textural classe	Sand	Clay loam	Clay loam	Sandy clay loam	Clay loam	Silt loam
Classification [34]	Typic Xero-psamment	Mollic Hapludalf	Typic Eutrudept	Entic Haplustoll	Typic Calcixerept	Limnic Haplosaprist
Classification [35]	Haplic Arenosol Eutric	Haplic Luvisol Humic	Leptic Cambisol Calcaric	Haplic Kastanozem	Haplic Calcisol	Sapric Histosol Calcaric
Parent material	Granodiorite	Sandstone (Eocene)	Marl-limestone (Eocene)	Alluvial carbonated sediments (Pleistocene)	Carbonatic lutite (Miocene)	Delta carbonatic sediments (Holocene)
Moisture regime	xeric	udic	udic	ustic	xeric	aquic
Location	Vallgorguina	Serra Bellmunt	Montesquiu	Taradell	Bellaterra	Ebro Delta
Vegetation (dominant plant)	<i>Quercus ilex</i>	<i>Fagus sylvatica</i>	<i>Quercus pubescens</i>	<i>Buxus sempervirens</i>	<i>Pinus halepensis</i>	Wetland plants

Table 2. Formulation of the ion-selective membranes employed in the potentiometric sensor array.

Sensor	PVC(%)	Plasticizer (%)	Ionophore (%)	Code
NH ₄ ⁺	33	BPA (66)	Nonactin (1)	NH4+
Na ⁺	22	NPOE (70)	CMDMM (6)	Na+
K ⁺	30	DOS (66)	Valinomycin (3)	K+
Gen Cat I	29	DOS (67)	Dibenzo-18-crown-6 (4)	DIBENZO
Gen Cat II	27	DBS (70)	Lasalocide (3)	LAS
Gen An I	29	DBP(65)	Tetraoctylammonium bromide(4)	TOAB
Cl ⁻	28.3	o-NPOE (70)	TDMAC (0.3)	Cl
SO ₄ ²⁻	33	o-NPOE(66)	1,3[bis(3-phenylthioureidomethyl)]benzene(1)	SO4
Ba ²⁺	27	DBS (70)	Monensin (3.0)	MONE
NO ³⁻	30	DBP (67)	TOAN (3.0)	NO3
Mg ²⁺	32.7	o-NPOE(65.6)	ETH4030(1.0)	Mg2+
Ca ²⁺ 1	32.9	o-NPOE (66)	Tetronasin (1.0)	TETRO
Ca ²⁺ 2	33.3	o-NPOE (65.2)	ETH1001 (1.0)	ETH1001
Ca ²⁺ 3	30	DOPP (65)	BBTP (5.0)	BBTP
Cu ²⁺	57.2	NPOE (34.3)	o-Xylylenebis(N,N-diisobutyldithiocarbamate) (6.9)	Cu2+
Pb ²⁺ II	37.2	NPOE (49.6)	S,S'-Methylenebis(N,N-diisobutyldithiocarbamate) (11.2)	Pb II
Pb ²⁺ IV	33	NPOE(65.65)	tert-Butylcalix[4]arene-tetrakis(N,N-dimethylthioacetamide) (1)	Pb IV
Zn ²⁺ I	40.22	NPOE(53.62)	Tetrabutylthiuram disulfide (2.3)	Zn I
Zn ²⁺ II	55.25	DBBP (41.4)	3,7,12,17-Tetramethyl-8,13-divinyl-2,18-porphinedipropionic acid disodium salt (2.76)	Zn II
H ⁺	32.8	DOS (65.6)	tri-N-dodecylamine (1)	H+

Table 3. Particle size analysis, chemical composition and concentration of main ions present in the water extract from six selected soils.

<i>Soil n°</i>	I	II	III	IV	V	VI
Whole soil						
<i>pH (water)</i>	6.31	5.00	7.87	7.79	8.11	7.68
<i>CaCO₃ (%)</i>	0	0	29.1	16.9	20.0	23.1
<i>Organic C (%)</i>	0.73	1.81	1.75	1.03	1.94	20.26
<i>Total N (%)</i>	0.08	0.32	0.18	0.17	0.08	0.83
<i>CEC (cmol/kg)</i>	9.2	20.7	13.2	10.0	11.5	>25
<i>Salinity (d/m)*</i>	0.28	0.37	0.61	0.62	0.64	5.8
<i>Coarse sand (%)</i>	72.4	27.1	7.3	7.1	24.4	10.6
<i>Fine sand (%)</i>	12.1	26.7	47.3	58.7	21.6	10.5
<i>Silt (%)</i>	8.3	22.5	29.1	14.7	34.6	74.6
<i>Clay (%)</i>	7.2	23.6	16.2	19.6	19.4	4.3
Water extract						
<i>Cl⁻ (cmol/kg)</i>	5.25	7.39	6.15	5.50	3.98	75.60
<i>SO₄²⁻ (cmol/kg)</i>	0.61	1.42	0.98	1.68	0.74	24.58
<i>NO₃⁻ (cmol/kg)</i>	>0.14	>0.14	>0.14	>0.14	10.18	0.39
<i>Na⁺ (cmol/kg)</i>	7.83	11.31	9.57	10.00	7.39	82.64
<i>Ca²⁺ (cmol/kg)</i>	4.74	13.47	18.21	16.22	10.98	34.43
<i>Mg²⁺ (cmol/kg)</i>	3.58	3.91	1.56	2.88	1.23	11.11
<i>K⁺ (cmol/kg)</i>	0.92	1.51	1.02	1.66	0.72	10.74
<i>NH₄⁺ (cmol/kg)</i>	2.82	2.56	>0.01	5.84	>0.01	14.62

* Saturated paste method

Table 4. Confusion matrix obtained with the PCA-ANN classification for the three different extraction methods of soil samples. In every attempt, a reduced portion of soils were not classified to any of existing classes, due to unclear output. The sensitivity is related with the correct classification rate while the specificity considers the classes correctly rejected.

Acetic Acid										
Expected \ Classified	I	II	III	IV	V	VI	Unclassified	Sensitivity %	Specificity %	
	I	II	III	IV	V	VI	Unclassified	Sensitivity %	Specificity %	
I	90	0	0	0	0	0	10	90	100	
II	0	95	0	0	0	0	5	95	100	
III	0	0	90	0	0	0	10	90	100	
IV	0	0	0	95	0	0	5	95	100	
V	0	0	0	0	95	0	5	95	100	
VI	0	0	0	0	0	100	0	100	100	
Average								94	100	

BaCl ₂										
Expected \ Classified	I	II	III	IV	V	VI	Unclassified	Sensitivity %	Specificity %	
	I	II	III	IV	V	VI	Unclassified	Sensitivity %	Specificity %	
I	80	0	0	0	0	0	20	80	100	
II	0	100	0	0	0	0	0	100	100	
III	0	0	75	0	5	0	20	75	90	
IV	0	0	0	70	0	0	30	70	100	
V	0	0	0	0	100	0	0	100	90	
VI	0	15	0	0	0	75	10	75	95	
Average								83	97	

Distilled water										
Expected \ Classified	I	II	III	IV	V	VI	Unclassified	Sensitivity %	Specificity %	
	I	II	III	IV	V	VI	Unclassified	Sensitivity %	Specificity %	
I	95	0	0	0	0	5	0	95	100	
II	0	100	0	0	0	0	0	100	100	
III	0	0	90	0	10	0	0	90	90	
IV	0	0	10	70	0	0	20	70	100	
V	0	0	0	0	100	0	0	100	90	
VI	0	0	0	0	0	100	0	100	95	
Average								83	96	

Table 5. Results of the fitted comparison lines between obtained and reference values of the evaluated soil parameters. Comparison limited to estimations for the external test set, for the three extracting agents used (uncertainties calculated at the 95% confidence level).

Prop.	Extract	Acetic Acid		
	Slope	Intercept	R ²	RMSE
%C organic	0.991 ± 0.030	0.06 ± 0.56	0.998	0.72
Conductivity	0.988 ± 0.031	4 ± 18	0.998	20.7
Cl ⁻	0.991 ± 0.027	0.5 ± 3.3	0.998	4.42
NO ₃ ⁻	1.002 ± 0.051	0.16 ± 0.83	0.994	1.13
Mg ²⁺	0.985 ± 0.030	0.19 ± 0.40	0.997	0.47
K ⁺	0.989 ± 0.025	0.15 ± 0.48	0.998	0.66
Prop.	Extract	Barium Chloride		
	Slope	Intercept	R ²	RMSE
%C organic	0.968 ± 0.048	0.05 ± 0.77	0.996	1.2
Conductivity	0.934 ± 0.147	0.2 ± 73.4	0.963	94.9
Cl ⁻	0.968 ± 0.046	0.05 ± 5.10	0.995	7.6
NO ₃ ⁻	0.966 ± 0.018	0.02 ± 0.28	0.999	0.59
Mg ²⁺	0.954 ± 0.117	0.1 ± 1.2	0.925	2.5
K ⁺	0.959 ± 0.086	0.09 ± 1.41	0.980	2.3
Prop.	Extract	Water		
	Slope	Intercept	R ²	RMSE
%C organic	0.985 ± 0.051	0.051 ± 0.009	0.995	1.2
Conductivity	0.992 ± 0.052	0.05 ± -1.85	0.995	32.4
Cl ⁻	0.983 ± 0.047	0.05 ± 0.61	0.995	7.1
NO ₃ ⁻	0.965 ± 0.114	0.1 ± 1.3	0.967	2.6
Mg ²⁺	0.974 ± 0.048	0.05 ± 0.54	0.995	0.74
K ⁺	0.981 ± 0.040	0.04 ± 0.39	0.996	0.95

CAPTIONS FOR FIGURES

Figure 1. PC1 vs. PC2 loadings plot of the PCA performed to the measurements from the sensor array employed for the soil measurements.

Figure 2. 3-D scores plots including the three first principal components in function of the extraction method used: Acetic Acid (A), Barium Chloride (B), Doubly distilled Water (C). The grouping regions correspond to Vallgorguina (I), Bellmunt (II), Montesquiu (III), Taradell (IV), Bellaterra (V) and Delta de l'Ebre (VI) soils. In each case, 3-D axis orientation is optimized for best visual separation of clusters.



