Fax: +90 224 4428243
E-mail: aelmaci@uludag.edu.tr

FEB/ Vol 22/ No 4b/ 2013 – pages 1225 – 1229

AN APPLICATION OF VIS-NIR REFLECTANCE SPECTROSCOPY AND ARTIFICIAL NEURAL NETWORKS TO THE PREDICTION OF SOIL ORGANIC CARBON CONTENT IN SOUTHERN ITALY

Antonio P. Leone[1], Natalia Leone[2] and Salvatore Rampone[2,*]

[1]Italian National Research Council, Institute for Mediterranean Agriculture and Forest System, Ercolano, NA 80056, Italy [2]Dept. of Science for Biology, Geology and Environment, University of Sannio, Benevento, BN 82100, Italy

ABSTRACT

Understanding soil properties is an essential prerequisite for sustainable land management. Assessment of these properties has long been gained through conventional laboratory analysis, which is considered costly and time consuming. Therefore, there is a need to develop alternative cheaper and faster techniques for soil analysis. In recent years, special attention has been given to vis–NIR reflectance spectroscopy and chemometrics. In this study we evaluated the potential of vis–NIR spectroscopy and Back Propagation Neural Networks (BPNN) for prediction of organic carbon (OC) of soils representative of three Mediterranean agro-ecosystems from the Campania region, southern Italy. An Artificial Neural Network (ANN) model was developed based on Multi-Layer Perceptron (MLP) network and trained by a Back-Propagation algorithm on reflectance data. The training and validation phases, confirmed by a ten fold cross validation methodology, led to a very satisfactory calibration of the BPNN model.

KEYWORDS: Mediterranean pedo-environments; Southern Italy; Soil properties; Organic carbon; vis-NIR reflectance spectroscopy; Back Propagation Neural Networks.

1 INTRODUCTION

Soil is one of our most important natural resources, having properties derived from the combined effect of climate and biotic activities, as modified by topography, acting on a parent material over a period of time [1]. It varies greatly over both space and time. The chemical and physical properties of soils determine their potential and limitations for agricultural and non-agricultural uses. Therefore, our understanding of their properties, functions and of their variability in space and time is essential for sustainable land management [2, 3].

* Corresponding author

Historically our understanding of the soil system and assessment of their properties has been gained through conventional laboratory analysis [4]. The latter, although usefully and practically irreplaceable for detailed investigations are costly and time consuming, thus not very suitable when large numbers of soil samples need to be analysed, as for example in large soil surveys, or for high resolution soil mapping and precision agriculture. Therefore, there is a need to develop alternative techniques for soil analyses.

In recent years, vis-NIR reflectance spectroscopy has been shown to be a useful technique for the measurement of various soil properties. Compared to conventional analytical methods, vis–NIR spectroscopy is faster, cheaper, and non-destructive; it requires less sample preparation, with less or no chemical reagents, is highly adaptable to automated and *in situ* measurements, and has the potential to analyse various soil properties simultaneously [4-6].

Reflectance spectroscopy refers to the measure of spectral reflectance [7] i.e., the ratio of the electromagnetic radiation reflected by a soil surface to that which impinges on it [8]. Since the characteristics of the radiation reflected from a material are a function of the material's properties, observations of soil reflectance can provide information on the properties and state of the soil [9]. The reflectance spectra of soil in the vis–NIR are largely non-specific due to the overlapping absorption of soil constituents. This characteristic lack of specificity is compounded by scatter effects, caused by soil structure or specific constituents, such as quartz [10]. All of these factors result in complex absorption patterns that need to be mathematically extracted from the spectra and correlated with soil properties. Hence, the analysis of soil reflectance spectra requires the use of chemometrics and pedometrics [11].

The most common calibration methods for soil applications are based on linear regressions, namely stepwise multiple linear regression (SMLR), principal component regression (PCR), and partial least squares regression (PLSR) [12, 13]. PCR and PLSR are related techniques and in most situations their prediction errors are similar [10]. Actually there is a growing interest in processing vis-NIR spectral characteristics by using Artificial Neural Networks (ANNs) [14-16]. In particular, ANNs are powerful tools for dealing with function approximation and it has been shown that they are universal function approximators [17, 18]. They are particularly useful in situations where the physical processes governing input–output responses are not entirely understood [19, 20]. A recent paper [21] compared the performance of three calibration methods on soil samples collected from Belgium and Northern France, namely, PCR, PLSR and Back Propagation Neural Network (BPNN) [22-24] for the accuracy of measurement of selected soil properties. Both the leave-one-out cross validation and prediction for the three replicates showed that all BPNN models outperformed PCR and PLSR models.

A lot of research has successfully demonstrated the use of vis–NIR reflectance spectroscopy to characterise soils from different pedo-climatic environments of the world. However, few studies have been carried out with soils from the Mediterranean regions [25-27]. Many of these studies have looked at iron oxy-hydroxides, due to their strong absorptions in the vis–NIR [28-31]. In this paper, we aimed to evaluate the performance of vis–NIR reflectance spectroscopy and BPNN for predicting an important soil property, the Organic Carbon (OC) content from three agricultural areas of the Campania region in southern Italy.

2 MATERIALS AND METHODS

2.1 Study areas

Soil samples used in this paper come from the Tusciano-Asa-Picentini district (Tusciano), the province of Benevento (Fortore beneventano (Fortore)), and the coastal area of the low Volturno river basin (Volturno) (Fig. 1). The climate, in all the three areas, shows characteristics typical of the Mediterranean region, with the wettest period between late autumn (October–November) and early spring (March–April). Temperature and potential evapotranspiration are inversely related to rainfall, with the highest values during summer (June–August).

The Tusciano district is located south of the town of Salerno. Its landscape is varied and includes coastal systems dominated by depressed retrodunal areas, alluvial terraces near the Sele river, conglomeratic and clayey coastal hills, near Eboli and Cilento, and internal calcareous relieves with falls of volcanic ash [32]. The dominant soil types are Calcaric Gleysols, Calcari-Gleyic Cambisols, Gleyic Luvisols, Ferri-Profondic, Vertic and Calcic Luvisols, Calcaric and Calcari-Mollic Cambisols, Molli-Eutrisilic and Molli-Vitric Andosols [33]. The dominant land use is broad leaved-forest, fruit trees, olive groves, and vegetable crops [34].

The Fortore is primarily an agricultural land located in the northern-east part of Campania. The landscape is dominated by clayey hills [32], covered by soils with strong "vertic" properties, mainly Eutric, Pellic, and Calcic Vertisols and Calcari-Vertic Cambisols [33]. The dominant land use is crop, mainly wheat [34]. Telesina Valley is a typical vineyard and olive growing area.

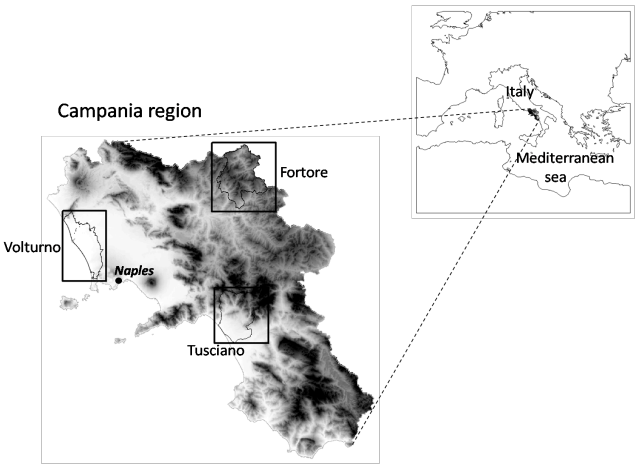

FIGURE 1 - Location of the study areas within the regional (Campania), national (Italy) and Mediterranean context.

The coastal belt of the low Volturno river basin is a rich irrigated agricultural land, mainly devoted to vegetal crops and fruit trees [34], located north and west of the city of Naples and the town of Caserta, respectively. The landscape is mainly characterised by areas with relatively low depressions from alluvial and land reclamation deposits, with intercalation of volcanic deposits, alluvial terraces composed by alluvial deposits mixed with volcanic tuff and ignimbrite and retrodunal lagoons composed by reclamation deposits, intercalated with peat layers [32]. The dominant soil types are Gleyic, Gleyi-Vertic, Calcari-Gleyic and Calcari-Fluvic Cambisols, and Calcaric Gleysols [33].

2.2 Soil sampling and laboratory analysis

Surface and sub-surface soil samples (0–30 cm, 30-60 cm) were collected from the three study areas [28]; we used 290 samples. Most of the sampling sites fall within agricultural fields, only few samples belong to forest areas. The samples were air-dried and ground to a size fraction of 2 mm. Each sample was divided into two sub-samples; one was used for conventional chemical and physical laboratory analysis, the other for the spectroscopic measurements.

The OC content was determined according to the Italian Official Methods for Soil Analysis [35]. Namely, OC is determined using Walkey-Black method.

2.3 Statistical analysis of soil properties

The measured soil property was statistically described in terms of minimum, maximum, mean and coefficient of variation (CV). According to Ameyan [36], a variable

shows small, moderate, or large variability when CV is below 20%, between 20 and 50%, or above 50%, respectively.

2.4 Spectroscopic measurements

The diffuse vis–NIR reflectance of soil samples was measured in the laboratory, under artificial light, using a FieldSpec Pro spectroradiometer [37]. This instrument combines three spectrometers to cover the portion of the spectrum between 350 and 2500 nm. The instrument has a spectral sampling distance of ≤ 1.5 nm for the 350–1000 nm region and 2 nm for the 1000–2500 nm region.

2.5 Spectroscopic data pre-processing

In order to enhance the predictive power of multivariate calibration models, spectroscopic data are pre-processed prior to data analysis. This is because variation in the **X** (predictor variables) data that is unrelated to **y** (response variable) may degrade the predictive ability of the models. The aim of pre-processing is to remove undesired variation in the data [38].

Common methods used to pre-process spectroscopic data include multiplicative signal correction (MSC) [39] and standard normal variance (SNV) correction [40, 41]. The MSC can be used to correct for light scattering variations while the SNV may be used to remove interferences due to light scattering and path length variations. Barnes *et al*. [40] also described the use of detrending using a quadratic polynomial together with the SNV transform to correct for any curvilinear trends and linear baseline shifts in the spectra. An alternative technique for the correction of these sources of error is based on the use of wavelet detrending and SNV with wavelet detrending [42].

A rapid and often utilised method for reducing scatter effects for continuous spectra consists of using derivatives [43]. The first derivative spectrum is the slope at each point of the original (or transformed) spectrum. It peaks where the spectrum has maximum slope and it crosses zero where spectrum has peaks. The second derivative is a measure of the curvature at each point in the original (or transformed) spectrum. The effect of the first derivative is usually to remove additive baseline ("offset"), whereas the effect of the second derivative also involves the removal of a linear baseline.

The problem with the use of derivatives is that they may reduce the signal and increase the noise of the spectra. To reduce the effects of random noise and improve signal-to-noise ratio, different filtering procedures can be used, prior to derivative transformation. The most commonly used filter in spectral reflectance analysis are median and Savitzky-Golay filters. The median filter can be implemented using:

$$x_i^{med} = Median(S_i) \qquad for\ i=0,1,\ldots,n-1$$

where x_i^{med} represents the filtered sequence, n the number of elements in the input sequence x, and S_i is a subset of x centered about its ith element. S_i may be described by:

$$S_i = \left(x_{i-rl}, x_{i-rl+1}, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{i+rr-1}, x_{i+rr}\right)$$

where rl is the filter left rank, and rr the filter right rank.

The Savitzky-Golay algorithm [44] uses a moving polynomial fit of any order and the size of the filter consists of (2n + 1) points, where n is the half-width of the smoothing window. The points between the 2n's are interpolated by the polynomial fit.

Before multivariate modeling, spectra can be also pretreated by mean centring and/or variance scaling the data.

For the purpose of the present paper, all the preprocessing procedures were assessed, either alone or in combination.

2.6 Principal Component Analysis

Principal component analysis [45] of the whole data set is needed in order to reduce the (highly redundant) dimension of the spectroscopic data. Principal Component Analysis (PCA) is a common data analysis technique used to reduce the dimensionality of large data sets displaying a high degree of covariance. The original data set undergoes a linear transformation whereby the axes of the new coordinate system are aligned such that the maximum variance is parallel to the first axis (principal component), the second highest variance lies along the second principal axis, etc. This is achieved by finding the principal eigenvalues of the covariance matrix, X, and then applying criteria to the eigenvalues to optimise the dimensionality of the data space.

2.7 Neural Network

An Artificial Neural Network, also simply called neural network, is a computational model which originally was intended for use in the simulation of the structure and/or functional aspects of biological neural networks [45]. It consists of an interconnected group of processing elements (neurons).

Neural networks are usually used to model complex relationships between inputs and outputs or to find patterns in data. Here we aim to express the relationship between the feature vectors of spectroscopic data and the corresponding OC values.

Three fundamental elements characterize any neural network:

– The network structure or topology, i.e. the way the neurons are linked;

– The activation function, which represents the answer of a neuron to the input stimuli;

– The learning algorithm.

The neural network structure we adopt is a feedforward one, the well known "Multilayer Perceptron" [23]. Here the neurons are organized into clusters or layers, called "input", "output" and "hidden" (i.e. those units

which are neither input nor output) layers. Each neuron of a given layer is connected to all the neurons of the next one, as depicted in Figure 2.

Our model is synchronous: at each time, each neuron receives as an input the weighted sum of the input values and/or of the other neuron outputs, as shown in the following equation:

$$O_k = f(\Sigma_n W_{kn} O_n) \quad (1)$$

where W_{kn} is the weight associated to the link from neuron n to neuron *k*, O_n is the output of neuron *n* or of the *n*-th input. The neuron output is a continuous and derivable function of its input. For our experiment we chose the sigmoid function,

$$f(x) = 1/(1 + e^{-x}) \quad (2)$$

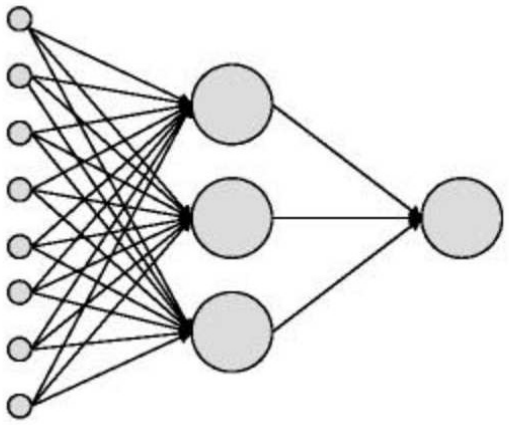

FIGURE 2 - Example of topological structure (architecture) of the Multilayer Perceptron used in our experiments.

The training procedure is the so called "back propagation". It makes use of a data set of feature vectors, each one labelled with the correct output, as examples of the correct input/output relationship. First a vector is presented to the input neurons and then the network gives its output. If it is not equal to the desired one, the difference (error) between these two values is computed and the weights W_{kn} are changed in order to minimize it. We repeat these operations for each input vector. Given the p-th vector in input, the error E_p is:

$$E = 1/2 \, \Sigma_j (t_j - O_j)^2 \quad (3)$$

where t_j is the *p*−th desired output value and O_j is the output of the corresponding neuron.

After an entire cycle of input vector presentations (*epoch*), we calculate the *Mean Square Error (MSE)* [45] for all the input patterns p. Then the epoch is repeated again until we reach an acceptable error. Then the network behaviour is validated by means of an independent set of data (validation set). In the end the Coefficient of Determination is measured [46].

The weight updating phase in the above algorithm, which is widely known in literature as the "generalized delta rule", aims at minimizing the error function E. This is done by a method known as Gradient Descent, where weight changes are made in the steepest downward direction of the error function [45].

The error back propagation algorithm has several parameters, and the most important ones are called, respectively, *learning rate* and *momentum term*. The first term is a measure of the influence degree, in the formula for updating weights, of the error term, whereas the latter one determines the influence of the past history of weight changes in the same formula. Moreover the weights W_{kn} are initially randomly chosen in a fixed range. The range choice can substantially improve or degrade the learning results. Performing optimal choices for all these parameters is a quite dark side of the procedure and must be achieved on the basis of personal experience and a trial and error approach.

2.8 Application environment

For our experiments we adopt an Excel-based system which simulates a neural network developed by Angshuman Saha and which is freely available all over the web[1]. By using this tool we can define feed-forward, back-propagation networks with 1 or 2 hidden layers, enter training data, set various learning parameters, start a learning phase and see the results (error rate, etc.) presented in various ways, including graphically. Categorical values are automatically converted into numerical ones.

As previously described, the data set is composed of labelled feature vectors called examples. Each example has the following structure

$$O_p, \; t_p = (x_1, x_2, \ldots, x_n), \; y$$

where the x_i's are the chosen features and y is the expected output.

The whole data set is partitioned into 10 non-overlapping subsets. From these 10 subsets, one can be chosen to evaluate the neural network performance (validation set), and the remaining 9 can be used to instruct it (training sets), i.e., setting the weights. This procedure is repeated 10 times, corresponding to all the possible choices of the validation set (ten-fold cross validation) [47], and the corresponding average performance was gauged in terms of an average error. It is worth underlining that in each trial the validation data is not used for training.

3 RESULTS AND DISCUSSION

3.1 Soils properties

Table 1 shows the descriptive statistics of the organic carbon content.

TABLE 1 - Descriptive statistics of the organic carbon content.

	Min	Max	Mean	CV (%)
OC (g kg^{-1})	0.4	215.6	17.9	114.9

[1] http://xoomer.virgilio.it/srampone/NNpred01.zip

Organic carbon content across the three study areas taken together is on average moderate (17.9 g kg^{-1}). However, it increases in the soils from Tusciano and Volturno (on average 19.5 and 20.0 g kg^{-1}, respectively) and decreases significantly in soils from Fortore (10.2 g. kg^{-1}).

3.2 Soil reflectance

Figure 3 shows the mean reflectance spectra and mean second-derivative of absorbance spectra of soils from the three study areas.

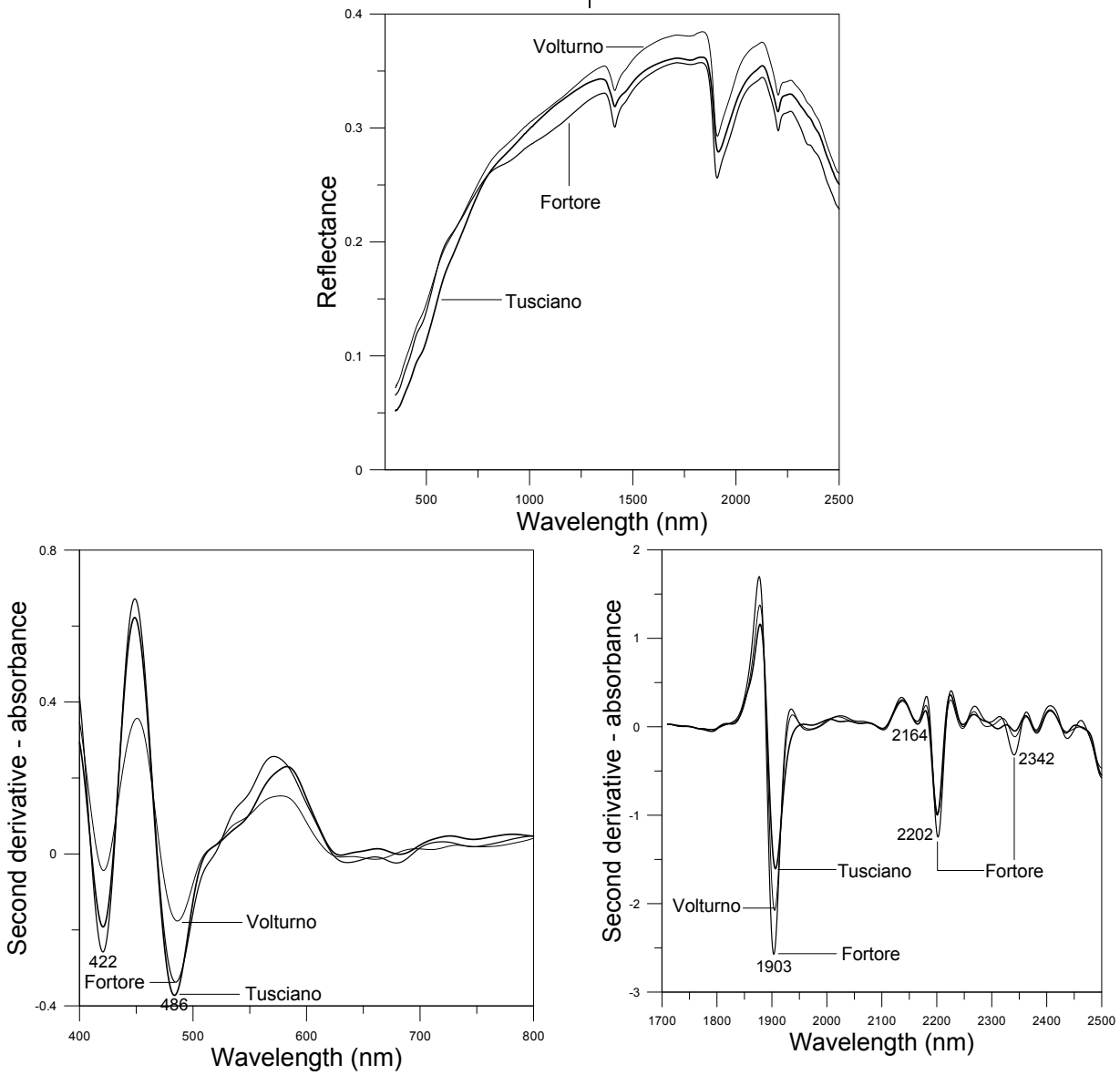

FIGURE 3 - Mean reflectance spectra and mean second-derivative of absorbance spectra of soils from the three study areas (Tusciano, Fortore and Volturno). Derivative spectra have been plotted in the 400 to 800 and 1700 to 2500 regions, where the main absorption bands occur, avoiding the region between 800 and 1700 nm.

The shape and overall reflectance of spectra are quite similar for all three areas. The only differences are the smaller reflectance of Tusciano soils in the visible, the larger reflectance of Volturno soils in the NIR and the lower reflectance of Fortore soils in this spectral range.

Derivative spectra of soils from all three areas show two evident adsorption bands in the visible range, near 422 and 486 nm. Both these bands, but particularly the one near 486 nm can be attributed to the presence of goethite iron-hydroxide [29]. The band at 486 nm is evidently deeper for soils from Tusciano and Fortore, than for soils from Volturno. Soil derivative spectra also show several important absorptions in the NIR, beyond 1700 nm. Specifically a strong water and hydroxyl band is present at 1903 nm. This band indicates vibrational stretching of H-O-H and OH^- ions in secondary clay minerals [48]. It is specifically diagnostic of the presence of montmorillonite [49]. The 1903 nm band is deeper in soils from Fortore and

Tusciano, than in those from Volturno, indicating a larger content of secondary clay minerals, particularly montmorillonite, in the soils from the first two areas, i.e. a more advanced stage of weathering in these areas. Several bands are present in the NIR beyond 2000 nm, due to the combination of vibrational processes in the silicate clay and carbonate minerals. A strong band apperas at 2202 nm, associated with a smaller band (doublet) at around 2164 nm, due to the combination of the OH-bond fundamental stretch with the Al-OH fundamental bending mode [50]. It must be observed that also montmorillonite has a band at 2200 nm, due to the OH-bond stretching. The 2200 nm band, resulting from the combined effect of kaolinite and montmorillonite (but also other clay minerals, such as vermiculite and illite), is deeper in the Fortore derivative spectrum, thus confirming the higher content of secondary clay minerals in soils from this area. Other "minor" bands associated to the presence of silicate clays [49, 51] are found beyond 2200 nm. A well-defined band apperas in all the three spectra at around 2342 nm, due to C-O stretching mode in $CaCO_3$ molecules. According to the results of chemical analysis (Table 1) this band is deeper in the average spectrum of Fortore.

Fig. 4 shows the values of correlation coefficients between OC contents and spectral reflectance.

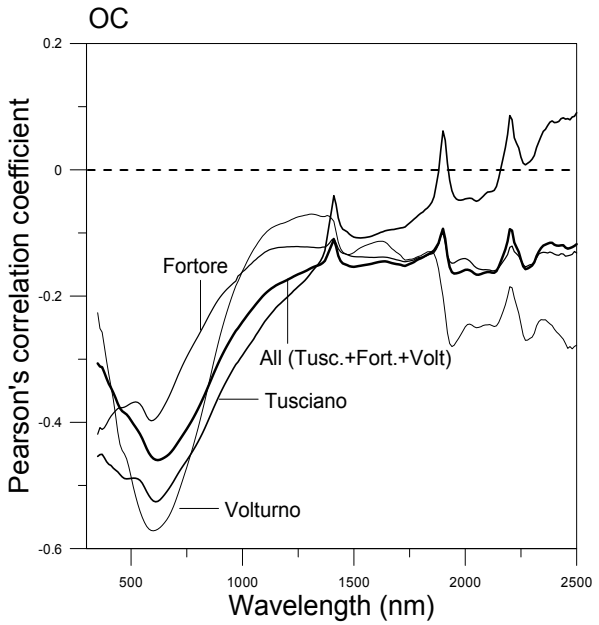

FIGURE 4 - Correlations between OC and spectral reflectance at each wavelength for soils from all the three study areas taken together and from the individual study areas.

Organic carbon is, in general, negatively correlated with reflectance throughout vis–NIR. However, the best correlations occur in the visible, with the largest one near 600 nm. The correlations are larger for soils from Volturno and Tusciano, which have larger OC contents (Table 1) compared to soils from Fortore.

3.3 PCA application

It was found that 30 principal components accounted for almost 100% of the variance for the whole data set. We accordingly encoded all data samples using 30-dimensional vectors of principal components.

3.4 Data Set

Finally we collected a data set of 290 examples, each one of 30 characteristics. The whole data set is partitioned into 10 non-overlapping subsets. From these 10 subsets, one can be chosen to evaluate the neural network performance (validation set), and the remaining 9 can be used to instruct it (training sets), i.e., setting the synaptic weights. This procedure is repeated 10 times, corresponding to all the possible choices of the validation set (Ten Fold Cross Validation) [45].

3.5 BPNN application

Several experiments were conducted over the data, varying the network structure and parameters. With reference to the network structure described in a previous Section, the number of hidden layers and the number of neurons in each hidden layer is arbitrary and should be neither too small nor too high. Too few neurons may deteriorate the network performance, whereas too many neurons may lead to over-fitting the network on the training data without actually improving its performance on the validation ones. Here, starting from a greater-than-needed BPNN, we apply a "pruning" technique to trim network size [52]. In this way the BPNN is defined as a single hidden layer of 13 neurons, while maintaining 30 input and one output (30-13-1 schema).

By a trial and error approach, the learning rate has been fixed to 0,8, and the momentum term is 0,27. The initial weights range in ±0,2. The best performance was achieved by iterating the learning for 500 cycles. (Table 2)

Results of BPNN applied to soil OC are provided in Table 3. These results were achieved by pre-processing the spectra prior to data analysis as follow: transformation of reflectance (R) to optical density (or absorbance, A = log 1/R), application of a median filter to absorbance, prior to computing their first derivative, and mean centring.

TABLE 3 - Results of BPNN applied to soil OC.

	Training Set		Validation Set	
Epoch	**MSE**	**%**	**MSE**	**%**
200	0,008	6,34%	0,006	4,83%

Table 4 reports the R^2 behaviour on the ten trials of the Ten Fold Cross Validation methodology we adopt. Overall, the selected architecture shows a R^2 value of 0,89 on the validation set.

TABLE 2 - Neural Network Architecture and parameters.

Neural Network Architecture and parameters			
Number of Inputs	30	Hidden Layer size	13
Number of Hidden Layers	1		
Learning parameter	0,8	Initial W_{kn} Range (0 +/- w):	0,2
Momentum	0,27		
Total #rows in data	290	No. of Training cycles	200

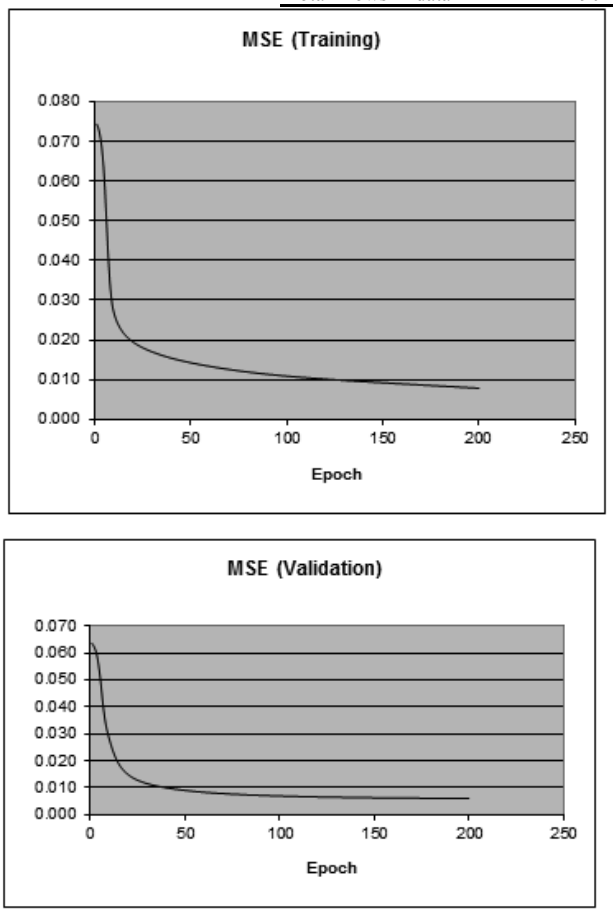

FIGURE 5 - Behaviour of training (7 a) and validation (7 b) MSE growing the number of epochs.

TABLE 4 - R^2 behaviour on the ten trials of the Ten Fold Cross Validation methodology we adopt.

Experiment	R^2
I	0.91
II	0.906
III	0.891
IV	0.908
V	0.866
VI	0.862
VII	0.866
VIII	0.904
IX	0.894
X	0.865

Figure 6 shows the relationships between observed and predicted soils data, using the whole data set.

The very good to excellent BPNN predictions of OC are, in general, in good agreement with the moderate to relatively large correlations of OC with visible reflectance.

BPNN appears to provide a method that can be used to accurately predict soil OC contents. It does so because the BPNN algorithm considers the multidimensionality of the spectra and takes into account the interrelations among the spectral reflectance values of individual properties across the vis-NIR spectral domain.

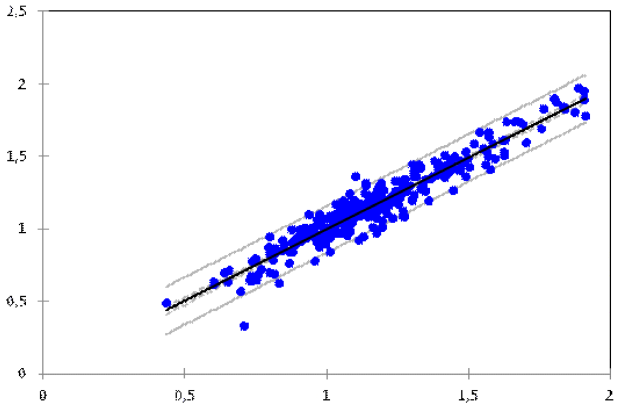

FIGURE 6 - Observed vs predicted soil OC, for soils from all the three investigated areas.

4 CONCLUSIONS

The present work demonstrated the potential of soil vis–NIR spectroscopy and BPNN for the analysis and prediction of soil OC contents in three different agricultural systems of Campania region, southern Italy. The investigated soil property correlates with spectral reflectance. Namely, organic carbon correlates negatively with reflectance, particularly in the visible range. Preliminary results, confirmed by a ten fold cross validation methodology, suggest that the approach is effective and robust.

The investigated soils from the Tusciano are characterised by moderate amounts of organic matter. Soils from Fortore have a small to moderate amounts of organic matter. Soils from Volturno have a large amount of organic matter[2]. By using vis-NIR spectroscopy and BPNN, Organic carbon is well predicted, using just few samples, in each of the three investigated areas. In this sense the performance of the prediction model should be then considered quite independent from the geographical origin of the investigated area (then from the soil units that characterise each of these areas).

More extensive work is obviously needed in order to validate/sharpen these results; however it is reasonable to assume that, through the use of a larger data set, the pre-

[2] Organic matter (= OC x 1.274).

diction model can be advantageously used in support to soil survey in other areas of the Campania region [53].

REFERENCES

[1] Brady, N. (1989) *The nature and properties of soils*. MacMillan Publishing Company: New York.

[2] McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S. and Shatar, T.M. (2000) An overview of pedometric techniques for use in soil survey. *Geoderma*, 97, 293–327.

[3] Florinsky, I.V., Eilers, R.G., Manning, G.R. and Fuller, L.G. (2002) Prediction of soil properties by digital terrain modeling. *Environ. Modell. Software*, 17, 295–311.

[4] Viscarra-Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. and Skjemstad, J.O. (2006) Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131, 59–75.

[5] McCarty, G.W., Reeves III, J.B., Reeves, V.B., Follett, R.F. and Kimble, J.M. (2002) Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Sci. Soc. Am. J*., 66, 640–646.

[6] Vasques, G.M., Grunwald, S. and Sickman, J.O. (2008) Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. *Geoderma*, 146, 14–25.

[7] Milton, E.J. (1987) Principles of field spectroscopy. *Int. J. Remote Sens*., 12, 1807-1827.

[8] Drury, S.A. (1993) *Image interpretation in geology*. Chapman & Hall: London.

[9] Irons, J.R., Weismiller, R.A. and Petersen, G.W. (1989) In: *Theory and Applications of optical remote sensing*; G. Asrar, Ed.; Wiley: New York, pp 66-106.

[10] Stenberg, B., Viscarra-Rossel, R.A., Mouazen, A.M. and Wetterlind, J. (2010) Visible and near infrared spectroscopy in soils science. *Advances in Agronomy*, 107, 163-215.

[11] Martens, H. and Naes, T. (1989) *Multivariate Calibration*. John Wiley & Sons: Chichester: UK.

[12] Ben-Dor, E. and Banin, A. (1995) Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Sci. Soc. Am. J.*, 59, 364–372.

[13] Dalal, R.C. and Henry, R.J. (1986) Simultaneous determination of moisture, organic carbon, and total nitrogen by infrared reflectance spectroscopy. *Soil Sci. Soc. Am. J*., 50, 120-123.

[14] Li, X., He, Y. and Fang, H. (2007) Non-destructive discrimination of Chinese bayberry varieties using Vis/NIR spectroscopy. *Journal of Food Engineering*, 81, 357–363.

[15] Li, X. and He, Y. (2008) Discriminating varieties of tea plant based on Vis/NIR spectral characteristics and using artificial neural networks. *Biosystems Engineering*, 99, 313–321.

[16] Liu, F. and He, Y. (2008) Classification of brands of instant noodles using Vis/NIR spectroscopy and chemometrics. *Food Research International*, 41, 562–567.

[17] Poggio, T. and Girosi, F. (1990) Networks for approximation and learning. *Proceedings of the IEEE*, 78, 1481-1495.

[18] Cybenko, G. (1989) Approximation by superposition of a sigmoidal function. *Mathematics of Control, Signals and Systems,* 2, 303-314.

[19] Freiwan, M. and Cigizoglu, H.K. (2005) Prediction of total monthly rainfall in Jordan using feed forward backpropagation method. *Fresen. Environ. Bull.*, 14, 142–151.

[20] Sengorur, B., Dogan, E., Koklu, R. and Samandar A. (2006) Dissolved oxygen estimation using artificial neural network for water quality control. *Fresen. Environ. Bull.*, 15, 1064–1067.

[21] Mouazen, A.M., Kuang, B., De Baerdemaeker, J. and Ramon, H. (2010) Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma*, 158, 23–31.

[22] Beale, R. and Jackson, T. (1990) *Neural Computing: an introduction*. Hadam Hilger, Bristol.

[23] Rumelhart, D.E. and McClelland, J.L. (1986) *Parallel Distributed Processing*. MIT Press, Cambridge.

[24] Garrett, J. (1994) Where and why artificial neural networks are applicable in civil engineering. *Journal of Computing in Civil Engineering*, 8, 129-130.

[25] Leone, A.P., Calabrò, G., Coppola, E., Maffei, C., Menenti, M., Tosca, M., Vella, M. and Buondonno, A. (2008) Prediction of soil properties with VIS-NIR-SWIR reflectance spectroscopy and artificial neural networks. A case study. *Advances in GeoEcology*, 39, 689-702.

[26] Calabrò, G., Leone, A.P. and Amenta, P. (2010) In: *Proceedings of the 11th Symposium on Statistical Methods for Food Industry*, Benevento, 23-26 February, 2010; Academy Scool: Afragola, Italy, pp 89-98.

[27] Leone, A.P., Viscarra-Rossel, R., Amenta, P. and Buondonno, A. (2012) Prediction of soil properties with PLSR and vis-NIR spectroscopy: Application to Mediterranean soils from southern Italy. *Current Analytical Chemistry*, 8, 283-299.

[28] Barrón, V. and Torrent, J. (1986) Use of the Kubelka-Munk theory to study the influence of iron oxides on soil colour. *J. Soil Sci*., 37, 499-510.

[29] Scheinost, A.C., Chavernas, A., Barrón, V. and Torrent, J. (1998) Use and limitations of second-derivative diffuse reflectance spectroscopy in the visible to near-infrared range to identify and quantify Fe oxide minerals in soils. *Clays and Clay Miner*., 5, 528 – 536.

[30] Leone, A.P. (2000) Bi-directional reflectance spectroscopy of Fe-oxides minerals in Mediterranean Terra Rossa soils: a methodological approach. *Agricoltura Mediterranea*, 130, 144-154.

[31] Sellitto, V.M., Fernandes, R.B.A., Barrón, V. and Colombo, C. (2009) Comparing two different spectroscopic techniques for the characterization of soil iron oxides: diffuse versus bi-directional reflectance. *Geoderma*, 149, 2-9.

[32] di Gennaro, A. (2002) *I sistemi di terra della Campania*. SELCA: Firenze.

[33] FAO. (1998) *World reference base for soil resources*. World Soil Resources, Reports n.84: Rome.

[34] Regione Campania, CUAS. http://sit.regione.campania.it/portal/portal/default/Home (accessed may 11, 2011).

[35] MIPAF, Ministero dele Politiche Agricole e Forestali (2000) *Metodi di Analisi Chimica del Suolo*. Franco Angeli: Milano.

[36] Ameyan, O. (1984) Surface soil variability of a map unit on Niger river alluvium. *Soil Sci. Soc. Am. J*., 50, 1289-1293.

[37] Analytical Spectral Devices (1995) *Technical Guide*, Bolder CO.

[38] Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C. and Wold, S. (2006) *Multi- and megavariate data analysis. Part I - Basic principles and applications*. 2th revised and enlarged edition, Umetrics AB Academy: Umeå, Sweden.

[39] Geladi, P., McDougall, D., Martens, H. (1985) Linearization and scatter correction for near infrared reflectance spectra of meat.

Appl. Spectrosc., 39, 491–500.

[40] Barnes, R.J., Dhanoa, M.S. and Lister, S.J. (1989) Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.*, 43, 772–777.

[41] Dhaona, M.S., Lister, S.J., Sanderson, R. and Barnes, R.J. (1994) The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformation of NIR spectra. *Journal of Near Infrared Spectroscopy*, 2, 43-47.

[42] Daubechies, I. (1992) *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics: Philadelphia, PA.

[43] Naes, T., Isaksson, T., Fearn, T. and Davies, T. (2002) *A User-Friendly Guide to Multivariate Calibration and Classification.* NIR Publications: Chichester, U.K., 93–104,

[44] Savitzky, A. and Golay, M.J.E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36, 1627–1639.

[45] Bishop, C. M. (1996) *Neural networks for pattern recognition.* Clarendon Press, Oxford.

[46] Nagelkerke, N.J.D. (1991) A Note on a General Definition of the Coefficient of Determination. *Biometrika*, 78, 691–692.

[47] Devijver, P. A. and Kittler, J. (1982) *Pattern Recognition: A Statistical Approach*. Prentice-Hall, London.

[48] Baumgardner, M.F., Silva, L.F., Biehl, L.L. and Stoner, R. (1985) Reflectance properties of soils. *Advances in Agronomy*, 38, 1-44.

[49] Hunt, G.R. and Salisbury, J.W. (1970) Visible and near-infrared spectra of minerals and rocks. I - Silicate minerals. *Modern Geology*, 1, 283-300.

[50] Hunt, G.R. and Ashley, R.P. (1979) Spectra of altered rocks in the visible and near infrared. *Econ. Geol.* , 74, 1613-1629.

[51] Clark, R.N., King, T.V.V., Klejwa, M. and Swayze, G.A. (1990) High spectral resolution reflectance spectroscopy of minerals. *J. Geophys. Res*., 95, 12653-12680.

[52] Suzuki, K., Horiba, I. and Sugie, N. (2001) A Simple Neural Network Pruning Algorithm with Application to Filter Synthesis. *Neural Processing Letters*, 13, 43-53.

[53] Rampone, S., and Valente A. (2012) Neural Network Aided Evaluation of Landslide Susceptibility in Southern Italy. *International Journal of Modern Physics C*, 23, 10-29.

Received: August 22, 2012
Revised: October 16, 2012
Accepted: November 13, 2012

CORRESPONDING AUTHOR

Salvatore Rampone
University of Sannio
Department of Science for Biology, Geology and Environment
Via dei Mulini 59/A - P.zzo Inarcassa
82100 Benevento
ITALY

Phone +39 0824 323663
Fax +39 02 700822249
E-mail: rampone@unisannio.it

FEB/ Vol 22/ No 4b/ 2013 – pages 1230 - 1238