Application of Fuzzy-Neural Network in Classification of Soils using Ground-penetrating Radar Imagery

Lameck O. Odhiambo  
*University of Nebraska - Lincoln*, lodhiambo2@unl.edu

Robert S. Freeland  
*University of Tennessee, Knoxville*

Ronald E. Yoder  
*University of Tennessee, Knoxville*

J. Wesley Hines  
*University of Tennessee, Knoxville*

Follow this and additional works at: [http://digitalcommons.unl.edu/animalscifacpub](http://digitalcommons.unl.edu/animalscifacpub)
Application of Fuzzy-Neural Network in Classification of Soils using Ground-penetrating Radar Imagery

Lameck O. Odhiambo, D. Tech. Sc.
Post-doctoral Research Associate¹

Robert S. Freeland, Ph.D., P.E.
Associate Professor¹

Ronald E. Yoder, Ph.D., P.E.
Professor and Head¹

J. Wesley Hines, Ph.D., MBA
Associate Professor²

¹Biosystems Engineering and Environmental Science
²Nuclear Engineering Department
The University of Tennessee, Knoxville, TN 37996-2210

Written for presentation at the
2002 ASAE Annual International Meeting/ CIGR XVth World Congress
Sponsored by ASAE
Hyatt Regency Chicago
Chicago, Illinois, USA
July 28-July 31, 2002

Abstract. Errors associated with visual inspection and interpretation of radargrams often inhibits the intensive surveying of widespread areas using ground-penetrating radar (GPR). To automate the interpretive process, this paper presents an application of a fuzzy-neural network (F-NN) classifier for unsupervised clustering and classification of soil profile using GPR imagery. The classifier clusters and classifies soil profiles strips along a traverse based on common pattern similarities that can relate to physical features of the soil (e.g., number of horizons; depth, texture and structure of the horizons; and relative arrangement of the horizons, etc). This paper illustrates this classification procedure by its application on GPR data, both simulated and actual real-world. Results show that the procedure is able to classify the profile into zones that corresponded with those obtained by visual inspection and interpretation of radargrams. Results indicate that an F-NN model can supply real-time soil profile clustering and classification during field surveys.

Keywords. Automation, Clustering, Soil mapping, Soil survey, Unsupervised classification, Fuzzy neural networks
Introduction

Ground-penetrating radar (GPR) probes soils nonintrusively, supplying high-resolution subsurface imagery of soil horizon profiles. The instrumentation transmits electromagnetic waves into the ground at user-selected frequencies. Soil boundaries having differing dielectric properties reflect radar waves back to a receiving antenna. The resulting GPR image (radargram) is a high-resolution profile image that highlights boundary interfaces of abrupt dielectric discontinuities, which often correlate to soil horizon boundaries and to regions of heterogeneity. This technology allows for continuous subsurface profiling; thereby yielding continual 3-D spatial data—more than can be practically obtained by borehole or auger sampling.

Soil scientists employ GPR to help assess soil properties that affect soil use, management, and classification (Doolittle and Amussen, 1992; Doolittle and Collins, 1995). A number of studies have demonstrated the ability of GPR to map important soil classification parameters. These studies have applied GPR technologies to map soil textural variations, organic matter content, thickness and depth of soil horizons and water tables, and soil compaction and plow pan development (Johnson et al., 1982; Doolittle, 1982; Doolittle, 1983; Olson and Doolittle, 1985; Collins and Doolittle, 1987; Schellentrager et al., 1988; Truman et al., 1988; Mokma et al., 1990; Raper et al., 1990; Doolittle and Amussen, 1992; Collins, 1992; Doolittle and Collins, 1995, Freeland et al., 1998). Traditional applications of GPR in soil mapping requires visual inspection and interpretation of the radargram regions into classes according to the perceived similarities in layers and other properties (Freeland, 1998; Adamson, 1999; Inman, 2000), with latter verification from ground-truthing. This method is somewhat similar to the traditional methods of soil judging, an evaluation procedure requiring considerable knowledge, skill, and experience, and often involving subjective judgment. In order to make full use of all the features collected in radar data and to reduce the requirement of subjective visual interpretation, we propose a quantitative procedure for systematic classification.

Neural networks and Fuzzy Classification Methods

Previous studies have traditionally used numerical and geostatistical methods to characterize soil variation (Cambell et al., 1970; Moore and Russel, 1972; Rea and Knight, 1998). However, neural networks (NN) and fuzzy system (FS) classifiers have recently emerged as promising alternatives to various conventional methods of pattern recognition and classification. For example, both methods have been used to classify radar images of groundcover taken from satellites.

Neural networks interpret a wide-variety of real-world problems including image analysis and classification. Zhang (2000) presented a comprehensive literature survey of important issues and recent NN developments for classification problems. Specifically, he examined the issues of posterior probability estimation, the link between neural and conventional classifiers, learning and generalization tradeoff in classification, the feature variable selection, as well as the effect of misclassification costs. Paola and Schowengerdt (1995) presented a detailed review and analysis of the use of NN-based classifiers for classification of remotely sensed multi-spectral imagery. Types of NN classifiers are supervised and unsupervised classifiers. In the supervised classifiers, the NN is trained to classify the data based on input-output examples presented to the network (i.e., the data are made up of distinct classes known a priori.) Typically, the supervised NN classifiers consist of multi-layer feed-forward networks that are trained using back-propagation algorithms. These algorithms employ recursive learning and gradient-decent search methods. Example applications include classification of images acquired by LANDSAT
Multispectral Scanner (Benediktsson et al., 1990), LANDSAT Thematic Mapper™ (Yoshida and Omatsu, 1994), SPOT HRV (Tzeng et al., 1994), and ASAS (Abuelgasim et al., 1996).

In contrast, unsupervised NN classifiers do not require input-output examples for training, but automatically classify the data based only on the information contained within the data. The network’s ability to cluster the input data into natural homogeneous sets determines classification, whereby the elements of each set are as similar as possible, and are as dissimilar as possible from those elements of the other sets. Users may select set numbers beforehand, or they may result from pre-imposed constraints. Most unsupervised NN-based classifiers fall in the categories of modified Adaptive Resonance Theory (ART) (Carpenter and Grossberg, 1988), the modified Learning Vector Quantization (Kohonen, 1989), and the Self-Organizing Maps (Kohonen, 1989). Applications that have utilized this type of classifiers for automatic classification of remotely sensed images include work by Hara et al. (1994). The performance of NN classifiers have been compared to conventional classifiers for a number of classification problems, and the results have shown that the accuracy of the NN approach is equivalent to, or slightly better than, conventional methods (Benediktsson et al., 1990; Hornik, 1991).

Real-world systems, such as soils and land cover, generally do not occur in discrete, internally uniform units with sharp boundaries; but they occur in continuous variations in the geographic space. This means that a soil or land cover can have partial membership in more than one class, and there exists the possibility of overlap in classes. Under such conditions, fuzzy systems (FS) provide a more natural setting for the formulation and approximate solution of classification than approaches based on crisp logic. Fuzzy methods have successfully been applied in the classification land cover using remotely sensed data (Zhang and Foody, 1998; Seong and Usery, 2001; and Tatem et al., 2001). Burrough et al. (1997) traces the development of conceptual paradigms in soil classification from the pre-1960’s model of crisp, non-overlapping classes to modern approaches using FS for handling continuous variation in both attributes and location. There are several examples of application of FS in the classification of soils using soil grid data obtained by soil sampling from various points and depths (Chang and Burrough, 1987; Burrough et al., 1992; Odeh et al., 1992; McBratney and De Gruijter, 1992; De Gruijter et al., 1997; Lark and Bolan, 1997; McBratney and Odeh, 1997, and Zhu, 1997).

Although individual applications of NN and FS have been successful in solving classification problems, integration of the two methods has received considerable attention. Integration appears to be more effective for managing the uncertainness in class boundaries of real world system. Neural network and fuzzy systems combine in a variety of ways including, fusion, serial combination, and/or parallel combination (Tsoukalas and Uhrig. 1997). Much research and development in fuzzy neural systems for classification has focused on fuzzy modification of the Adaptive Resonance Theory (ART) and self-organizing map networks. The fuzzy ART developed by Carpenter et al. (1991) generalizes unsupervised ART to learn patterns in analog and binary data. The fuzzy ARTMAP (Carpenter and Crossberg, 1994) uses two fuzzy ART modules coupled by a map field for supervised learning of the patterns. The resulting fuzzy ARTMAP has classified vegetation cover at species level from LANDSAT and terrain data (Carpenter et al., 1997). Gopal et al. (1999) used fuzzy ARTMAP to classify global land cover from remotely sensed imagery from AVHRR satellites. Others applications include Lin et al. (2000) who developed a cascaded neural-fuzzy network with feature mapping to cluster satellite images, In all these studies, the fuzzy neural networks performed better than the conventional and/or pure NN classifiers.

Objectives

Moving toward automation for classification of soil radargrams, we propose the application of a fuzzy-neural network (F-NN) based classifier that implements an unsupervised classification of
the soil using digitized ground-penetrating radar imagery. Our objectives are: 1) to adapt a F-NN classifier to the application of interpreting soil radargrams, illustrated both by actual and simulated GPR data, and 2) to compare these results to those obtained from human interpretation of radargrams into pattern classes.

**Materials and Methods**

In choosing the type of network for soil classification, we considered two limiting conditions: 1) the characteristics of radargram, and 2) the soil variability. The characteristics of the radargram depend upon a number of variable factors, such as frequency of the transmitted electromagnetic waves, type of antenna, antenna speed on the ground surface, climate, soil-water content, etc. This means that the characteristics of a radargram are not stable and are liable to have variations owing to different surveying conditions and equipment. As such, there are no fixed input-output examples for use in neural training. Furthermore, soil variation is more continuous than discrete; therefore, it calls for a continuous classification. With no fixed input-output examples and continuous soil variation, we require unsupervised classification and a system that is able to handle fuzzy boundary conditions. The network algorithm developed by Adeli and Hung (1995) provides such an approach in that it quantitatively assigns individual patterns to geographically and taxonomically continuous classes.

The Adeli-Hung algorithm was adapted for use in this study because of its suitability and simplicity. It consists of a NN and FS combined in series as shown in figure 1. The classification process proceeds in two stages. The first stage is an unsupervised NN clustering and classification process. In this stage, the NN classifies soil profile strips into a certain number of clusters determined dynamically. After all the profile strips have been classified, the values of the mean vector (prototype) for each cluster are stored in the weights associated with the connections between the input and output nodes. The second stage evaluates the fuzzy membership values for each profile strip in the set of classified clusters.

**Neural Network process**

The process of unsupervised clustering and classification uses a topology and weight change two-layer neural network (NN). The number of input nodes is equal to the number of patterns (M) in each soil profile strip. The number of output nodes is equal to the number of clusters. Since the number of clusters is determined through the classification process, the topology of the NN is changed and dynamically self-organized during the classification process. The profile strips are classified one by one. First, a NN with M input nodes and one output node, denoted as \( \Phi(M,1) \), is generated and the first profile strip inputted. At this point, the first profile strip belongs to the first cluster. Then the second profile strip inputs into the network. If the second strip classifies to the first cluster, the output node, representing the first cluster, becomes active. In this case, the topology of the neural network does not change, but the connection weights update using a recursive estimate algorithm. The topology of the neural network is still a \( \Phi(M,1) \) network. On the other hand, if the second strip classifies as a new cluster, an additional output node adds to the NN. The values of weights in the original NN do not alter. In this case, the topology of the NN modifies as an \( \Phi(M,2) \) network. The classification follows this procedure until all the profile strips are classified. This unsupervised classification process is shown schematically in figure 2.
The classification of a soil profile strip into an existing cluster, or a new cluster, uses the concept of maximum likelihood. We define a function $D(X, C)$, called the degree of difference, to represent the difference between a profile strip $X$ and a cluster $C$. This function maps two given vectors ($X$ and $C$) to a real number ($D$). The patterns of each cluster (means of the patterns of the strips in the cluster) are stored in the links (weights) of NN during the classification process. A threshold value $k$ is predefined as a crossover value.

The implementation scheme is as follows:

Equation 1. Calculate the degree of difference, $D(X, C)$, between the profile strip, $X$ and each cluster, $C_i$. The function $D(X, C)$ is defined as the Euclidean distance represented by:

$$D(X, C) = \left[ \sum_{j=1}^{M} (x_j - c_{ij})^2 \right]^{1/2}$$ .................................(1)

Equation 2. Find the smallest degree of difference, $D_{\text{min}}$, and assign the cluster with the smallest degree of difference as an active cluster.

$$C_{\text{active}} = \{C_i | \min \{D(X, C)\}, i = 1, 2, \ldots, P\}$$ ...........................................(2)

Equation 3. Compare the value of $D_{\text{min}}$ with a predefined crossover value $k$. If the value of $D_{\text{min}}$ is greater than the crossover value, the profile strip is classified as a new cluster.

$$C_{\text{new}} = X, \text{if } k < \min \{D(X, C), i = 1, 2, \ldots, P\}$$ .............................................(3)

The similarity function is the Euclidean distance between strip ($X$) and cluster ($C$). If the Euclidean distance for a given strip is less than the crossover value, the strip belongs to the cluster $C$ and the fuzzy membership value is between zero and one. Otherwise, the strip does not belong to the cluster and the fuzzy membership value is equal to zero.
Evaluating Fuzzy Membership Functions

After the process of neural classification is completed, classified clusters may be disjoint or partly overlapping. If the clusters are completely disjoint, each given strip in the soil profile belongs to only one of the classified clusters. However, if the classified clusters are partly overlapping, a given strip in the profile may belong to more than one cluster. In this case, the boundaries of the classified clusters are fuzzy rather than crisp. The algorithm for evaluating the membership functions assumes that there is a prototype for each cluster, defined as the mean of all profile strips in that cluster, and the degree of membership of each strip in the cluster is how similar this strip is to the prototype one. The similarity is a function of distance between the strip and the prototype of the cluster. Based on the triangular-shaped membership function, the fuzzy membership value of the X strip in the cluster C, \( \mu_c(X) \), is defined as:

\[
\mu_c(X) = f(D(X,C)) = \begin{cases} 
0 & \text{if } D(X,C) > k \\
1 - \frac{D(X,C)}{k} & \text{if } D(X,C) \leq k 
\end{cases}
\]

Program development

The program was implemented using MATLAB\textsuperscript{™} (2001), a software package for high-performance numerical computation and visualization. MATLAB offers programming features with hundreds of built-in functions that allow quick manipulation of data sets in a wide variety of ways. In addition, there are 'Toolboxes' for special application of neural networks and fuzzy systems. MATLAB offers graphical user interface (GUI) tools that allow the use of MATLAB as an application development tool.

Application and Results

The F-NN classifier accepts the GPR data in a two-dimensional format (lines x sample) that consists of continuous pixels. Each pixel value in the array represents the reflective intensity of multivariate soil properties in the actual soil profile. The digital intensity values range 0 - 65535, with zero representing no reflection and 65535 as maximum reflection. The image was decomposed into a reduced number of pixels by averaging the pixel values contained in 170 lines by 200 samples. The resultant line vectors each represent arbitrary soil horizons and the column vectors each represent a soil profile strip. The decomposed data was normalized to 0-100 scale to reduce the dimensionality of multivariate properties present. The classifier reads the pixel values of the profile strips through a classifier window passed along the traverse. The ability of the F-NN to classify soils on the basis of properties and sequence of similar horizons is illustrated by its application to classify an idealized hypothetical soil profile with distinct diagnostic horizons shown in figure 3.

The idealized profile consists of distinct diagnostic layers (black) located at two depths within a uniform soil profile formation (gray). Visual inspection and interpretation of the profile results in four soil classes (i.e. C1, C2, C3, and C4). Class C1 is uniform throughout its depth, class C2 has a diagnostic layer located in its lower half depth, class C3 has a diagnostic layer located in both halves of its depth, and class C4 has the diagnostic layer located in the upper half of its depth. Application of the F-NN to this data at a tolerance value of 1.0 resulted in 100% agreement with visual inspection and interpretation as illustrated by the step plot in figure 3. Each stair in the plot represents a class in the profile, and stairs at the same level infer similarity in soil properties. These results confirm the validity of the F-NN approach.
Next, the F-NN was applied to real world data that were obtained from GPR surveying of two sites located in the southeastern U.S. Site 1 is at the Ames Plantation located in the Coastal Plain Region of Tennessee. The soils are formed in variable depths (<1 to 2 m) of loess overlying alluvium deposits underlain by tertiary-aged sand deposits. Site 2 is located in the Cumberland Plateau near Crossville, TN. The soils are fine sandy loam and loam, both underlain by sandstone bedrock found in the upper one meter of the soil profile.

The radagrams and the resulting F-NN classifications at the two sites are given in figures 4 and 5 respectively. The horizontal scale on the radargram represents units of distance along the traverse, while the vertical scale represents soil depth. The radagrams consist of interfaces displayed in groups of multiple bands that represent different soil horizons, hard pans, strata, and/or water table that are present within the soil profile. The different tones represent other soil properties such as texture, color, structure, moisture content, and organic mater content. The horizontal changes in vertical properties of soils are inferred from changes in the intensity and width of the bands. The profile image at site 1 show gradual changes in the tone and width of the bands within line ranges 0-40, 80-110, and 340-370 as we move along the traverse. The profile image at the site 2 has a single distinct horizon (bedrock) cutting across a soil profile with rather uniform properties. The distinct horizon varies in depth, width, and consistency along the traverse. Sections of fractured bedrock are observed at different portions along the traverse (i.e. sample range 0 – 1100; 1600 –1800; and 2500 –2900). The locations of all the class division boundaries in both profiles are not obvious from visual inspection.

Application of the F-NN to the radagrams results in a number of clusters and classes determined by the crossover value. There are more clusters at lower crossover values than at higher values. For example, at site 1 (Fig. 4), crossover values 2.0, 3.0, and 3.7 results in 9, 4, and 2 classes respectively. At lower crossover values, the computed fuzzy membership values indicates no overlapping in clusters and at higher levels, there is some overlapping. The same trend of clustering and classification is observed with data at site 2 (Fig. 5). The number of clusters and classes at crossover values of 4.5, 5.5 and 6.5 where 7, 4, and 3 respectively. The choice of crossover value appears to depend on internal variability of the profile and the level of resolution required in the classification. Careful visual inspections of the radagrams indicate that F-NN used the depth, width, and dispersion of diagnostic features in the layers to cluster and classify the profile. Some obviously similar sections of profile clusters in the same class.
Conclusion

Soil classification has traditionally emphasized both the arrangement and properties of individual soil horizons. The F-NN classifier satisfied this tradition by making classification on the basis of properties and sequence of component horizons. The contribution of this method lies not only in the possibility of processing large amounts of data generated by GPR, but also provides a more rational classification of soils based on quantitative evaluation of soil properties rather than classifications based on descriptive properties. Possibilities exist supply for interfacing this method with other tools like GPS and GIS to provide real-time soil maps during field surveys.

Acknowledgements

Support for this project was provided by the United States Department of Agriculture-Water Resources Assessment Protection Program of the National Research Initiative, “Predicting Offsite Subsurface Migration of Agrochemicals—Noninvasive Surveying,” contract no. 99-35102-8273, and by the Trustees of the Hobart Ames Foundation, Ames Plantation, Grand Junction, Tennessee.

References


Figure – 5. The radargram at site 2, and the F-NN classes at crossover values of 4.5, 5.5, and 6.5


