Investigation Of A Fuzzy-Neural Network Application In Classification Of Soils Using Ground-Penetrating Radar Imagery

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

R. S. Freeland
University of Tennessee, Knoxville

R. E. Yoder
University of Tennessee, Knoxville

J. W. Hines
University of Tennessee, Knoxville

Follow this and additional works at: https://digitalcommons.unl.edu/biosysengfacpub

Part of the Bioresource and Agricultural Engineering Commons, Environmental Engineering Commons, and the Other Civil and Environmental Engineering Commons


This Article is brought to you for free and open access by the Biological Systems Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Biological Systems Engineering: Papers and Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
INVESTIGATION OF A FUZZY-NEURAL NETWORK APPLICATION IN CLASSIFICATION OF SOILS USING GROUND-PENETRATING RADAR IMAGERY

L. O. Odhiambo, R. S. Freeland, R. E. Yoder, J. W. Hines

ABSTRACT. Errors associated with visual inspection and interpretations of radargrams often inhibit the intensive surveying of widespread areas using ground-penetrating radar (GPR). To automate the interpretive process, this article presents an application of a fuzzy-neural network (F-NN) classifier for unsupervised clustering and classification of soil profiles using GPR imagery. The classifier clusters and classifies soil profile 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 article illustrates this classification procedure by its application on GPR data, both simulated and actual. Results show that the procedure is able to classify the profile into zones that corresponded with the classifications obtained by visual inspection and interpretation of radargrams. Application of F-NN to a study site in southwest Tennessee gave soil groupings that are in close correspondence with the groupings obtained in a previous study, which used the traditional methods of complete soil morphological, chemical, and physical characterization. At a crossover value of 3.0, the F-NN soil grouping boundary locations fall within a range of ±2.7 m from the soil groupings determined by the traditional methods. These results indicate that F-NN can supply accurate real-time soil profile clustering and classification during field surveys.

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

Gound-penetrating radar (GPR) probes soils non-intrusively, 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 that 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; Collins and Doolittle, 1987; Truman et al., 1988; Doolittle and Amussen, 1992; Collins, 1992; Doolittle and Collins, 1995, Freeland et al., 1998). Traditional applications of GPR in soil mapping require visual inspection and interpretation of the radargram regions into classes according to the perceived similarities in layers and other properties (Freeland et al., 1998; Adamson, 1999; Inman et al., 2001, 2002), with later verification by 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. Soil maps prepared from correctly interpreted GPR data can provide the basis for evaluating soil and site conditions for precision agriculture. For example, GPR generated soil maps have been used to plan the production of cranberries in Plymouth County (Turrene, 1997). 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 using neural networks for systematic classification.

NEURAL NETWORKS AND FUZZY CLASSIFICATION METHODS

Previous studies have traditionally used numerical and geostatistical methods to characterize soil variation (Campbell et al., 1970; Moore et al., 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 are used to 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 multispectral imagery. Two categories of NN classifiers are supervised and unsupervised classifiers. In 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 \emph{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-descent search methods. Example applications include classification of images acquired by LANDSAT Multispectral Scanner (Benediktsson et al., 1990), LANDSAT Thematic Mapper (Yoshida and Omatu, 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), modified Learning Vector Quantization (Kohonen, 1989), and 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 with 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 of land cover using remotely sensed data (Zhang and Foody, 1998; Seong and Usery, 2001). Burrough et al. (1997) traces the development of conceptual paradigms in soil classifica-
tion from a pre-1960s 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 (Odeh et al., 1992; McBratney and De Gruyтер, 1992; Lark and Bolan, 1997; McBratney and Odeh, 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 uncertainty in class boundaries of real-world systems. Neural network and fuzzy systems can be combined 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 Grossberg, 1994) uses two fuzzy ART modules coupled for supervised learning of the patterns. The resulting fuzzy ARTMAP has been used to classify vegetative 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. Other applications include the research work by 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**

To automate classification of soil radargrams, we proposed 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. The objectives of our study were to adapt a F-NN classifier to the application of interpreting soil radargrams, illustrated both by actual and simulated GPR data; and to verify the F-NN by comparing the results with the soil groupings obtained by methods of complete soil morphological, chemical and physical characterization.

**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 the frequency of the transmitted electromagnetic waves, the type of antenna, the antenna speed on the ground surface, the climate, the soil-water content, etc. Therefore, the characteristics of a radargram 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, unsupervised classification and a...
system that is able to handle fuzzy boundary conditions are required. 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. Other techniques such as the fuzzy c-means and adaptive fuzzy clustering algorithms could also be used. The Adeli-Hung algorithm 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 CLASSIFICATION PROCESS**

The process of unsupervised clustering and classification uses a two-layer neural network (NN). The number of input nodes is equal to the number of patterns (M) in each soil profile strip, and the number of output nodes is equal to the number of clusters. The number of clusters is determined during the classification process, and the NN topology is changed and dynamically self-organized in the 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 data is inputted. At this point, the first profile strip is assigned to the first cluster. Then the second profile strip data is inputted 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 estimation algorithm. The topology of the neural network is still an $\Phi(M,1)$ network. On the other hand, if the second strip classifies as a new cluster, an additional output node is added 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.

![Figure 1. Adeli-Hung algorithm consisting of a NN and FS combined in series.](image1)

![Figure 2. Unsupervised neural networks classification process.](image2)

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

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

where, $x_j$ and $c_j$ are elements in the column vectors representing patterns for $X$ and $C$, and $j$ is the row number, and $M$ is the total number of rows.

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

$$C_{active} = \{ C | \min D(X,C), i \in P \}$$  \hspace{1cm} (2)

where $P$ is the number of classified clusters

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

$$C_{new} = X \text{ if } k < \min D(X,C), i \in P$$  \hspace{1cm} (3)

If the degree of difference between a given profile strip, $X$, and a cluster, $C$, 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 neural classification process 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 by the mean of all profile strips in that cluster, and the degree of membership of each strip to the cluster is a measure of how similar the strip is to the prototype. 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:

$$
\mu_C(X) = \begin{cases} 
0, & \text{if } D(X,C) > k \\
1 - \frac{D(X,C)}{k}, & \text{if } D(X,C) \leq k 
\end{cases}
$$  

Program Development

The program was implemented using MATLAB® (The Mathworks, Inc., 2000), 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 radargram in a digitized format that consists of continuous array of pixels. The horizontal scale on the radargram represents units of distance along the traverse, while the vertical scale represents soil depth. The spacing of pixels in both the horizontal and vertical scales defines the spatial resolutions of the radargram. Horizontal resolution is determined by the speed at which the GPR antennas are moved on the ground surface. Slower antenna speeds result in greater degree of resolution with more reflection traces (scans) recorded per meter of ground covered. The vertical resolutions depend on the depth of penetration of the transmitted electromagnetic waves and are determined by the antenna frequency and the electrical conductivity of the soil being imaged. Horizontal resolutions in the range of 85 to 100 scans/m of ground cover and vertical resolutions in the range of 102 to 128 samples/m depth was observed in the data used in this study.

Each pixel value in the array of GPR data represents the reflective intensity of multivariate soil properties in the actual soil profile. The digital intensity values range 0 to 65535, with zero representing no reflection and 65535 as maximum reflection. The image was decomposed to reduce the dimensionality by averaging the pixel values contained in 170 scans by 200 samples. The resultant horizontal vectors each represent arbitrary soil horizons (approximately 1.33 to 1.67 m horizon depth) and the column vectors each represent a soil profile strip (approximately 2.0 to 2.35 m wide). The decomposed data were normalized to a 0 to 100 scale. 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 (dark) 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, 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.

![Figure 3. Idealized simulated soil profile with distinct diagnostic horizons and F-NN classification at crossover value = 1.0.](image-url)
The F-NN was applied to real world data that were obtained from GPR surveying of two sites located in the southeastern United States. Site 1 is at the Ames Plantation located in the Major Land Resource Area (MLRA) 134. The soils are formed in variable depths (<1 to 3 m) of loess overlying alluvium deposits underlain by tertiary-aged sand deposits. The soils at this site have been thoroughly investigated and grouped based on complete soil morphological, chemical and physical characterization (Inman et al., 2002). Site 2 is located in the Cumberland Plateau in MLRA 125. The soils are fine sandy loam and loam, both underlain by sandstone bedrock found in the upper one meter of the soil profile.

The radargrams and the resulting F-NN classifications with various crossover values for site 1 are given in figure 4, and for site 2 in figure 5. The radargrams consist of interfaces

Figure 4. The radargram for the study site at Ames Plantation (Site 1), and F-NN classification at crossover values 2.0, 3.0, and 3.7.
displayed in groups of multiple bands that represent different soil horizons, hard pans, strata, and/or water tables that are present within the soil profile. The different tones represent other soil properties such as texture, color, structure, moisture content, and organic matter 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 (fig. 4) shows gradual changes in the tone and width of the bands along the traverse. The profile image at site 2 (fig. 5) 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., pixel column ranges 0 to 1100; 1600 to 1800; and 2500 to 2900). The locations of all the class division boundaries in both profiles are not obvious from visual inspection.

Figure 5. The radargram for the study site in Cumberland Plateau (Site 2), and F-NN classification at crossover values 4.5, 5.5, and 6.5.
Application of the F-NN to the radargrams results in a number of clusters and classes determined by the crossover value. The crossover value indicates the range of influence of a cluster when the data space is considered as a unit hypercube. The crossover value is adjusted up or down to obtain the required fineness in classification results. Specifying a small crossover value will usually yield many small clusters in the data and specifying large crossover value will usually yield a few large clusters. 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 results in no overlapping clusters, and at higher levels, there are some overlapping clusters. 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 were 7, 4, and 3, respectively. The choice of crossover value appears to depend on the internal variability of the profile and the level of resolution required in the classification. Careful visual inspections of the radargrams indicate that the F-NN used the depth, width, and dispersion of diagnostic features in the layers to cluster and classify the profile. The classification results show that some obviously similar sections of the profile clusters into the same class.

The results of F-NN classification of soils at site 1 were next compared with soil groupings determined by methods of complete soil morphological, chemical and physical characterization (Inman et al., 2001, 2002). In their study of soils at site 1, Inman et al. (2001) used a hydraulic excavator to trench a 36-m long transect to a depth of approximately 3 m. They chose a total of six pedons spaced 6 m apart along the northeast face of the trench for an in-depth soil morphological investigation and described each pedon according to Soil Survey Staff (1993). Based on these analyses, they were able to divide the soil profile along transect into three groups (Inman et al., 2001, 2002). Comparison of the results of F-NN classification of soils at site 1 with the grouping determined by Inman et al. (2001, 2002) shows a close correspondence (fig. 6). At crossover value = 2.0, the F-NN classification reveals internal group variations not indicated.

![Figure 6](image-url)

*Figure 6. A comparison of the soil grouping determined by the traditional methods and visual interpretation (groups 1, 2, and 3 in radargram), and F-NN classification using crossover values = 2.0 and 3.0.*
by the traditional methods. For example, it shows that group 2 is a transition between group 1 and group 3. At a crossover value of 3.0, the smaller group variations are averaged and the F-NN boundary locations fall within range of ±2.7 m from those determined by traditional methods (fig. 6). Since soil variation is more continuous than discrete, variation in the exact location of group boundaries in the range of only ±2.7 m appears to confirm the reliability and accuracy of the F-NN method.

**CONCLUSION**

Soil classification has traditionally emphasized both the arrangement and properties of individual soil horizons. The F-NN classifier satisfied this tradition by performing classification on the basis of both the properties and the sequence of component horizons. Close correspondence was observed between the soil groupings obtained by F-NN and those obtained by methods of complete soil morphological, chemical and physical analyses. The potential contribution of this method lies in its ability to fast process large amounts of GPR data, and to provide a classification based on numerical evaluation of the radargrams rather than classifications based on subjective description by visual inspection of radargrams. Also capabilities exist to interface this method with other tools 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**


