2004

Soil Characterization Using Textural Features Extracted from GPR Data

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

Robert S. Wright
University of Tennessee, Knoxville

Ronald E. Yoder
University of Tennessee, Knoxville, ryoder2@unl.edu

The Society for engineering in agricultural, food, and biological systems

The Canadian Society for Engineering in Agricultural, Food, and Biological Systems

An ASAE/CSAE Meeting Presentation *Paper Number: 042108*

Soil Characterization Using Textural Features Extracted from GPR Data

Lameck O. Odhiambo Post-Doctoral Res. Associate

Robert S. Freeland Professor

Ronald E. Yoder Professor and Head

Biosystems Engineering and Environmental Science Department
The University of Tennessee, Knoxville, TN 37996

Written for presentation at the
2004 ASAE/CSAE Annual International Meeting
Sponsored by ASAE
Fairmont Chateau Laurier, The Westin, Government Centre
Ottawa, Ontario, Canada
1 - 4 August 2004

Abstract. *Soils can be non-intrusively mapped by observing similar patterns within ground-penetrating radar (GPR) profiles. We observed that the intricate and often indiscernible textural variability found within a complex GPR image possesses important parameters that help delineate regions of similar soil characteristics. Therefore, in this study, we examined the feasibility of using textural features extracted from GPR data to automate soil characterizations. The textural features were matched to a "fingerprint" database of previous soil classifications of GPR textural features and the corresponding ground truths of soil conditions. Four textural features (energy, contrast, entropy, and homogeneity) were selected for inputs into a neural-network classifier. This classifier was tested and verified using GPR data obtained from two distinctly different field sites. The first data set contained features that indicate the presence or lack of sandstone bedrock in the upper 2 m of a shallow soil profile of fine sandy loan and loam. The second data set contained columnar patterns that correspond to the presence or the lack of vertical preferential-flow paths within a deep loess soil. The classifier automatically grouped each of these data sets into one of the two categories. Comparing the results of classification using extracted textural features to the results obtained by visual interpretation found 93.6% of the sections that lack sandstone bedrock correctly classified in the first set of data, and 90% of the sections that contain pronounced columnar patterns correctly classified in the second set of data. The classified profile sections were mapped using integrated GPR and GPS data to show surface boundaries of different soil categories. These results indicate that extracted textural features can be utilized for automatic characterization of soils using GPR data*.

Keywords: *Soils, Characterization, Textural features, Ground-penetrating radar, Neural network*

Introduction

Subjective interpretation of ground-penetration radar (GPR) patterns, followed by ground-truth corroboration, is a common method by which one can noninvasively delineate and identify subsurface features. Examples are: (1) Identifying preferential subsurface flow pathways through which pollutant loaded water may flow (Freeland *et al.*, 2002 a; Gish *et al.*, 2002); (2) Detecting water table depths, variations of soil water content, and wetting front (Freeland *et al.*, 1998; Huisman *et al.*, 2002; Schmaltz *et al.*, 2002; Smith *et al.*, 1992); (3) Estimating the thickness and volume of organic materials in soils (Doolittle *et al.*, 1990); (4) Characterizing landfill sites (Doolittle *et al.*, 1997; Orlando and Marchesi 2001; Porsani *et al.*, 2004); and (5) Mapping tree root systems (Butmor *et al.*, 2003; Hruska *et al.*, 1999; Stokes *et al.*, 2002).

A few studies report on employing automated methods, rather than subjective visual interpretation, for the rapid characterization of GPR data. Al-Nuaimy *et al.* (2000) developed a system of automated targeting of buried utilities and solid objects within GPR patterns. The system consisted of a neural-network classifier, a pattern-recognition stage, and pre-processing, feature extraction, and image processing stages. They tested the system on GPR patterns containing pipes, cables, and anti-personnel landmines. Their results indicated that effective automated mapping is possible for such structures. Scott *et al.* (2000) also proposed a procedure that uses image processing and pattern recognition methods to automate characterization of GPR data to detect distress on bridge decks, with preliminary testing providing good results. Shihab *et al.* (2002) developed a neural network target identifier based on statistical features extracted from GPR patterns. The neural network discriminated between signals and other spurious sources of reflections such as clutter. They applied this classifier to a variety of GPR data sets gathered from a number of sites and the results showed that the classifier was capable of outlining regions of extended targets such as disturbed soil or storage tanks, and was able to pinpoint the location of localized targets such as landmines and pipes. In a previous study, the authors (Odhiambo *et al.*, 2004) investigated an application of a fuzzy-neural network (F-NN) classifier for unsupervised clustering and classification of soil profiles using GPR imagery, and found that F-NN can supply accurate soil clustering and classification based on both the arrangement and properties of individual soil horizons.

The need for an automated classification system becomes apparent whenever one attempts visual interpretation, as GPR data sets collected during a routine field-scale survey are massive. The difficulties associated with visual interpretation often limit the use of GPR as a practical, widespread tool for soil investigations. A technique that provides automatic characterization of

vast quantities of GPR data to classify soils into categories associated with known environmental conditions would greatly enhance the usefulness of GPR for environmental management, not only by saving time, but also by reducing the probability of misclassification.

Objectives

In this study, four textural features (energy, contrast, entropy, and homogeneity) based on a co-occurrence matrix, were extracted and used as inputs to a neural network classifier. The classifier was used to partition soil profile regions into categories and the results matched to a database of previous soil classification that relates textural parameters to known soil characteristics. We examined the applicability of such features to automate characterization and mapping of soil sections into categories associated with known environmental conditions. The method was tested and verified using GPR data sets from two sites.

Methods and Materials

Data Collection

The data used in this study were collected at two sites using a GSSI Subsurface Interface Radar (SIR) System 10-A and 200-MHz antenna (Model 3105) (Geophysical Survey Systems, Inc., New Salem, NH). This system measures the time that it takes electromagnetic energy to travel from the antenna to an interface and back. The control settings used on the SIR 10-A unit were as shown in table 1. The first site is located at the University of Tennessee Agricultural Experiment Station (Plateau Experiment Station), near Crossville, TN. The soils at this site are

Table 1. Control settings used on SIR 10-A unit

Parameter	Site 1	Site 2
Antenna Model 1	3105	3105
Range	60 ns	75 ns
Samples/Scan	512	512
Bits/Sample	16	16
Scans/Second	50	50
# Gain points	5	5
Horizontal IIR Running Average	5	5
Vertical IIR High Pass	#poles=2, Freq=65	#poles=2, Freq=130
Vertical IIR Low Pass	#poles=2, Freq=600	#poles=2, Freq=1065

fine sandy loam and loam, and are underlain by sandstone bedrock in the upper 2 m of the soil profile. The second site is located at the Ames Plantation near Grand Junction, TN. The soils at this site consist of loess overlying alluvium deposits underlain by tertiary-aged sand deposits. This site was specially prepared for a study of the preferential flow paths by applying water into a large ring infiltrometer constructed at the center of the site, and taking GPR surveys in a spiral path around the infiltrometer at intervals after water application.

Feature Extraction

Ground penetration radar data sets are typically very large and contain a lot of information that is redundant and superfluous for soil characterization. The purpose of feature extraction is to reduce the dimensionality of the data and convert it to variables that are more suitable for discrimination between soil categories. The GPR data are displayed as a two-dimensional array of numbers, where each value in the array represents the reflective intensity of multivariate soil properties in the soil profile. The vertical direction of such a display is time, which can be converted to depth once the signal velocities are known, and the horizontal direction is linear distance on the ground surface. The reflective intensities are represented in the data by values that range from 0 to 65535, where 0 and 65535 represents the maximum limits of reflection, and the value 32768 represents no reflection as shown in figure 1. For computational efficiency, the data were normalized to the range 1 to 256.

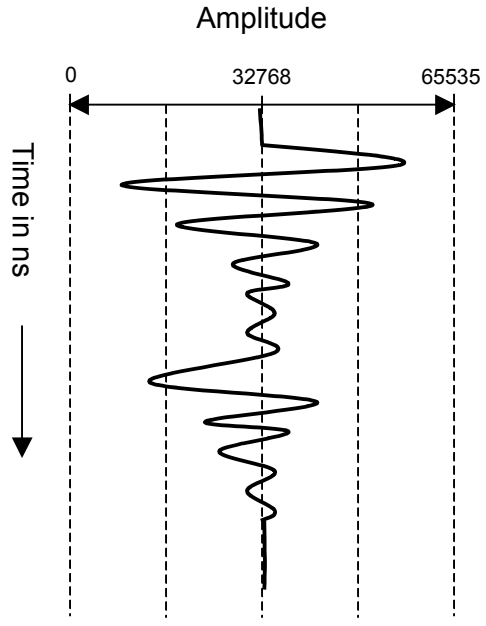

Figure 1. A typical single waveform showing the maximum limits of reflection to the left and right side.

We have observed that the intricate and often indiscernible textural variability found within a complex GPR image possesses important parameters that help delineate regions of similar soil characteristics. Several methods have been used to extract textural features from digital images for use in image classification. Haralick *et al.* (1973) developed a conceptual framework of measures from which textural features are derived. The framework is based on the co-occurrence matrices, which define the spatial relationship of pairs of values of pixels in a digital image. The co-occurrence matrix of a GPR data set, P(i,j,d,θ), is the frequency of occurrence in the data set of pairs of reflective intensity levels (i and j), that are separated by a certain distance (d) and lie along a certain direction (angle θ). When the GPR data set is read through a classifier window passed along the linear distance of the GPR display, the frequencies for angles quantized to 45° intervals for each window are expressed as follows:

$$P(i,j,d,\theta=0^o)=\#\{((k,l),(m,n))\in(M\times N)\times(M\times N)\,|\,k-m=0,|l-n|=d,I(k,l)=i,I(m,n)=j\} \quad (1)$$

$$P(i,j,d,\theta=45^o)=\#\{((k,l),(m,n))\in(M\times N)\times(M\times N)\,|\,(k-m=d,l-n=d),I(k,l)=i,I(m,n)=j\} \quad (2)$$

$$P(i,j,d,\theta=90^o)=\#\{((k,l),(m,n))\in(M\times N)\times(M\times N)\,|\,|k-m|=d,l-n=0,I(k,l)=i,I(m,n)=j\} \quad (3)$$

$$P(i,j,d,\theta=135^o)=\#\{((k,l),(m,n))\in(M\times N)\times(M\times N)\,|\,(k-m=d,l-n=d),I(k,l)=i,I(m,n)=j\} \quad (4)$$

where # denotes the number of elements in the set, M×N is the size of the classifier window, i,j = 1…256 (number of possible of reflective intensity levels), k,m = 1…M (image width), and l,n =1…N (image height). The frequencies of occurrence are inherently not invariant under rotations. To alleviate these directional biases, the frequencies were summed as follows:

$$P_{ij}=P(i,j,d,0^o)+P(i,j,d,45^o)+P(i,j,d,90^o)+P(i,j,d,135^o) \quad (5)$$

Haralick *et al.* (1973) proposed 14 measures of textural features, which are derived from the co-occurrence matrices, and each represents certain image properties such as coarseness, contrast, homogeneity, and texture complexity. For this study four commonly used textural features (equations 7 to 8), were extracted and used as inputs to the neural network classifier.

1. Energy: $$f_1=\sum_{ij}P_{ij}^2 \quad (7)$$

2. Entropy: $$f_2=\sum_{ij}P_{ij}\log P_{ij} \quad (8)$$

3. Homogeneity: $$f_3 = \sum_{ij} \frac{P_{ij}}{|i-j|} \quad (9)$$

4. Contrast: $$f_4 = \sum_{ij} |i-j| P_{ij}^2 \quad (10)$$

where P_{ij} is the sum of frequency of occurrence in the data set of pairs of reflective intensity levels (i and j) calculated in equation (5).

Relational Database

A relational database of previous soil classifications of GPR textural patterns, soil conditions, and corresponding ground-truths was constructed. The textural features were extracted from representative sections of GPR data sets that contain patterns associated with known environmental conditions. At Plateau Experiment Station, textural features were extracted from representative sections of the GPR data that are associated with the absence of sandstone bedrock (see figure 2). At Ames Plantation, textural features were extracted from representative sections of the GPR data that show pronounced columnar patterns occurring in and around the alluvium/Tertiary sand interface (see figure 3). Freeland *et al.*, (2002 a) found these columnar patterns to be associated with vertical preferential flow paths.

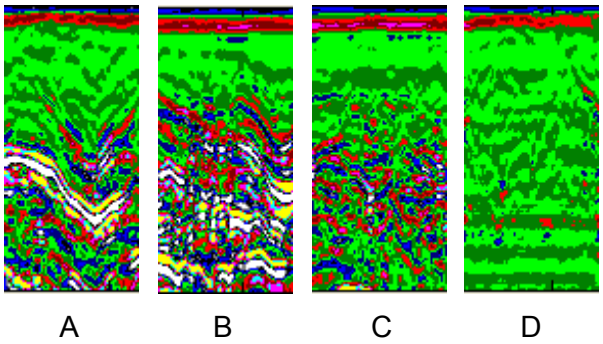

Figure 2. Examples of GPR profile sections at site 1. **A** and **B** represent conditions with solid bedrock, **C** represents conditions with fractured bedrock, and **D** represents conditions with no bedrock.

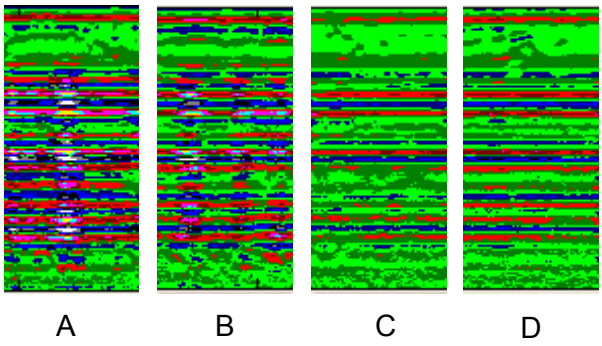

Figure 3. Examples of GPR profile sections at site 2. **A** and **B** represent conditions with pronounced columnar patterns associated with vertical preferential flow paths. **C** and **D** represent conditions with few or no columnar patterns.

The database items were organized into two tables from which data can be accessed and reassembled to determine soil categories. Table 2 contains the extracted textural parameters and the assigned class. Each row contains a unique instance of data for the categories defined by parameters in the columns. The textural parameters in table 2 are defined in equations (7) to

(10). Table 3 contains ground-truth information on soil condition at the sites and the assigned category. The two tables relate through the class fields in table 2 and the category field in table 3. The soil condition is determined from GPR data by using the relationship between tables 2 and 3. The relational database has the important advantage of being easy to extend when new data on soil categories becomes available.

Table 2. Extracted textural parameters and the assigned classes

Class	Textural Parameters			
	f_1	f_2	f_3	f_4
1	0.8675	0.0065	0.0879	0.0065
2	*	*	*	*
3	0.0239	0.0039	1.5812	0.0470
4	0.0949	0.0444	2.2415	0.0397

* Class 2 is not a unique category. It represents a conglomeration of different bedrock depths, thickness, and state, i.e., solid and/or fractured.

Table 3. Ground-truth information on soil conditions and the assigned categories

Soil Conditions	Category
Site # 1	
Sandstone-bedrock absent	1
Sandstone-bedrock present	2
Site #2	
Preferential flow paths	3
No preferential flow path	4

Neural Network Classifier

Neural networks have become popular in classifying complex data sets because of their adaptive, accurate, and rapid processing properties. Several types of neural network classifiers have been used in characterization and classification of digital data. These include the multi-layer perceptron (MLP), the learning vector quantization (LVQ), the self-organizing feature maps, and the radial basis function classifiers (Looney,

1997). In this study we used a two–layer perceptron (figure 4) that performs supervised

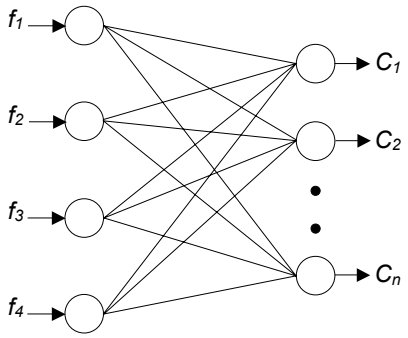

Figure 4. Neural network classifier

classification of soil profile strips by comparing each strip's textural features to samples in the database (C_1, C_2, …, C_n) that represent known soil conditions. The four textural features (f_1, f_2, f_3, and f_4) extracted from each strip are used as inputs to the network, and the number of output nodes is equal to the number of pre-determined soil categories (n). The classification of a soil profile strip into the categories existing in the database 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 category C. This function maps two given vectors (X and C) to a real number (D). The patterns of each soil category are stored in the links (weights) of the neural network during the classification process. A threshold value φ 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 category, C, in the database. The function D(X,C) is defined as the Euclidean distance represented by:

$$D(X,C) = \left[\sum_{j=1}^{M} \left(x_j - c_j \right)^2 \right]^{1/2} \tag{11}$$

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. Next the smallest degree of difference, D_{min}, is found and compared with a predefined crossover value φ. If the degree of difference between a given profile strip, X, and a category, C, is less than the crossover value, the strip belongs to

the category C. Otherwise, the strip does not belong to the category and is rejected. The procedure is repeated for each of the unique categories in the database.

Implementation and Results

The method was implemented using a MATLAB program developed for extracting textural features from GPR data, characterization of soil profile strips using the neural network classifier and a relational database, and mapping using an integrated GPR and GPS data sets to show surface boundaries of different soil categories. The method was tested and verified using GPR data collected from the study sites at Plateau Experiment Station and Ames Plantation. The crossover parameter (φ) was optimized for each site based on texture type. The database of the two study sites consisted of GPR images having different types of texture, and therefore, a common φ value was rather difficult to find. At the Plateau Experiment Station, the data show underlying bedrock, which contains features that are associated with three known environmental conditions: solid bedrock, fractured bedrock, and no bedrock. The data for this site was divided into 604 profile strips (510-pixels depth by 100-pixel width), which were classified to identify areas with no bedrock in the 2 m depth from the rest of the area using a φ value of 1.2. The results are shown in figure 5 (a-c). Figure 5(a) shows how the program separated the soil profile into two categories (areas without bedrock and areas with bedrock). Out of the 604 profile strips, 47 were classified as having no bedrock in the 2 m depth. Careful visual interpretation of the data found 44 profile strips indicating the absence of bedrock in the 2 m depth. Comparing the results of classification using extracted textural features to visual interpretation found 93.6% of the profile strips lacking sandstone bedrock correctly classified and 6.4% misclassified. Figure 5(b) shows how markers on the GPR data were assigned to the different categories and surface plotted. Figure 5(c) is a surface map showing areas without and with sandstone bedrock in the 2 m depth. The locations of the areas identified as having no bedrock matched closely to the locations identified as having the bedrock at a depth greater than 1.5 m by Freeland *et al.* (2002 b) through a study, which used physical probing to determine the depth to bedrock. The areas with depth to bedrock greater than 1.5 m included areas without bedrock in the 2 m depth.

At the Ames Plantation site, sections of the GPR data shows pronounced columnar patterns occurring in and around the alluvium/Tertiary sand interface. These columnar patterns have been associated with vertical preferential flow paths (Freeland *et al.*, 2002 a). The data for this site was divided into 305 soil profile strips (510-pixel depth by 100-pixel width), which were classified to identify areas with pronounced columnar patterns from the rest of the area using a φ value of 0.5. The results are shown in figure 6(a-c). Figure 6(a) shows how the program

separated the soil profile into two categories (areas that exhibited pronounced columnar patterns and areas that exhibited few or no columnar patterns). Out of the 305 profile strips, 126 were classified as exhibiting pronounced columnar patterns, while careful visual interpretation identified 140 as exhibiting pronounced columnar patterns. Comparing the results of classification using extracted textural features to visual interpretation found 90.0% of the profile strips having pronounced columnar patterns correctly classified and 10.0% misclassified. Figure 6(b) shows how markers on the GPR data were assigned to the different categories and surface plotted, and figures 6(c) is a surface map showing areas with pronounced columnar patterns and hence high soil water content, and areas with few or no columnar patterns and hence low soil water content.

Conclusions

The results of this study indicate that textural features extracted from GPR data can be used to automate soil characterization. The method was demonstrated using fairly simple GPR data obtained from two distinctly different field sites. The soil conditions were determined by matching the extracted textural features to "fingerprints" in a relational database of previous soil classifications of GPR textural features and the corresponding ground truths of soil conditions. Only four textural parameters were used in this study, but the effects of additional textural parameters on the accuracy of prediction could be investigated in future studies. The relational database can be expanded when new data on soil categories become available. A neural network classifier was used to assign data to the known soil categories. The φ values are optimized based on the type of texture, and hence no single φ value is applicable to all types of GPR data. The results of soil characterization using extracted textural features was found to be in close agreement with results obtained by careful visual interpretation of the data (93.6% correct classified for site 1 and 90% correct classified for site 2.) The classified soil profile sections were mapped using an integrated GPR and GIS data to show surface boundaries of different soil categories.

Acknowledgments

We thank the Tennessee Agricultural Experimental Station for financial and technical support. We gratefully acknowledge support from the Trustees of the Hobart Ames Foundation, Ames Plantation, Grand Junction, Tennessee.

References

Al-Nuaimy, W., Y. Huang, M. Nakhkash, M. T. C. Fang, V. T. Nguyen, and A. Eriksen. 2000. Automatic detection of buried utilities and solid objects with GPR using neural networks and pattern recognition. J. Appl. Geophy. 43(2-4):157-165.

Butnor, J. R., J. A. Doolittle, K. H. Johnsen, L. Samuelson, T. Stokes, and L. Kress. 2003. Utility of ground penetrating radar as a root biomass survey tool in forest systems. Soil Sci. Am. J. 67:1607-1615.

Doolittle, J. L., P. Fletcher and J. Turenne. 1990. Estimating the thickness and volume of organic materials in cranberry bogs. Soil Surv. Horiz. 31(3): 73-78.

Doolittle, J., L. Hernandez, and J. Galbraith. 1997. Using ground-penetrating radar to characterize a landfill site. Soil Surv. Horiz. 38(2): 60-67.

Freeland, R. S., J. C. Reagan, R. T. Burns, and J. T. Ammons. 1998. Sensing perched water using ground-penetrating radar – a critical methodology examination. Appl. Eng. Agric. 14(6): 675-681.

Freeland, R. S., D. J. Inman, R. E. Yoder, and J. T. Ammons. 2002 a. Detecting vertical anomalies within loessial soils using ground penetration radar. Appl. Eng. Agric. 18(2):263-264.

Freeland, R. S., R. E. Yoder, J. T. Ammons, and L. L. Leonard. 2002 b. Integration of real-time global positioning with ground-penetrating radar surveys. Appl. Eng. Agric. 18(5):647-650.

Gish, T. J., W. P. Dulaney, K. J. S. Kung, C. S. T. Daughty, J. A. Doolittle, and P. T. Miller. 2002. Evaluating use of ground-penetrating radar for identifying subsurface flow paths. Soil Sci. Soc. Am. J. 66(5):1620-1629.

Haralick, R. M., K. Shanmugan, and I. Dinstein. 1973. Textural features for image classification. IEEE Trans. Systems, Man and Cybernetics, SMC-3(6): 610-621.

Hruska, J., J. Cermak, and S. Sustek. 1999. Mapping tree root systems with ground-penetrating radar. Tree Physiol. Victoria, Canada: Heron Pub. 19(2): 125-130.

Huisman, J. A., J. J. J. C. Snepvangers, W. Bouten, and G. B. M. Heuvelink. 2002. Mapping spatial variation in surface soil water content: comparison of ground-penetrating radar and time domain reflectrometry. J. Hydrol. Amsterdam: Elsevier Science B.V. 269(3/4): 194-207.

Looney, C. G. 1997. Pattern recognition using neural networks. Theory and algorithms for engineers and scientists. Oxford University Press, New York.

Odhiambo, L. O., R. S. Freeland, R. E. Yoder, and J. W. Hines. 2004. Investigation of a fuzzy-neural network application in classification of soils using ground penetrating radar imagery. Appl. Eng. in Agric. 20(1): 109-117.

Orlando, L. and E. Marchesi. 2001. Georadar as a tool to identify and characterize solid waste dump deposits. J. Applied Geophys. 48(3): 163-174.

Porsani, J. L., W. M. Filho, V. R. Elis, F. Shimeles, J. C. Dourado, and H. P. Moura. 2004. The use of GPR and VES in delineating a contamination plume in a landfill site: a case study in SE Brazil. J. Appl. Geophys. 55(3/4): 199-209.

Scott, M., J. C. Duke, N. Davidson, G. Washer, and R. Weyers. 2000. Automated characterization of bridge deck distress using pattern recognition analysis of ground penetrating radar data. Materials Evaluation, 58(11): 1305-1309.

Schmaltz, B., B. Lennartz, and D. Wachsmuth. 2002. Analyses of soil water content variations and GPR attribute distributions. J. Hydrol. Amsterdam: Elsevier Science B.V. 267(3/4): 217-226.

Shihab, S., W. Al-Nuaimy and Y. Huang. 2002. Neural network target identifier based on statistical features of GPR signals. 9th Intern. Conf. on Ground Penetrating Radar, Proceedings of SPIE Vol. 4758. p 135-138.

Smith, M. C., G. Vellidis, D. L. Thomas, and M. A. Breve. 1992. Measurement of water table fluctuations in a sandy soil using ground-penetrating radar. Trans ASAE. 35(4): 1161-1166.

Stokes, A., T. Fourcaud, J. Hruska, J. Cermak, N. Nadyezdhina, V. Nadyezhdin, and L. Praus. 2002. An evaluation of different methods to investigate root system architecture of urban trees in situ. I. Ground-penetrating radar. J. Arboric. Champaign, IL. International Society of Arboriculture. 28(1): 2-10.

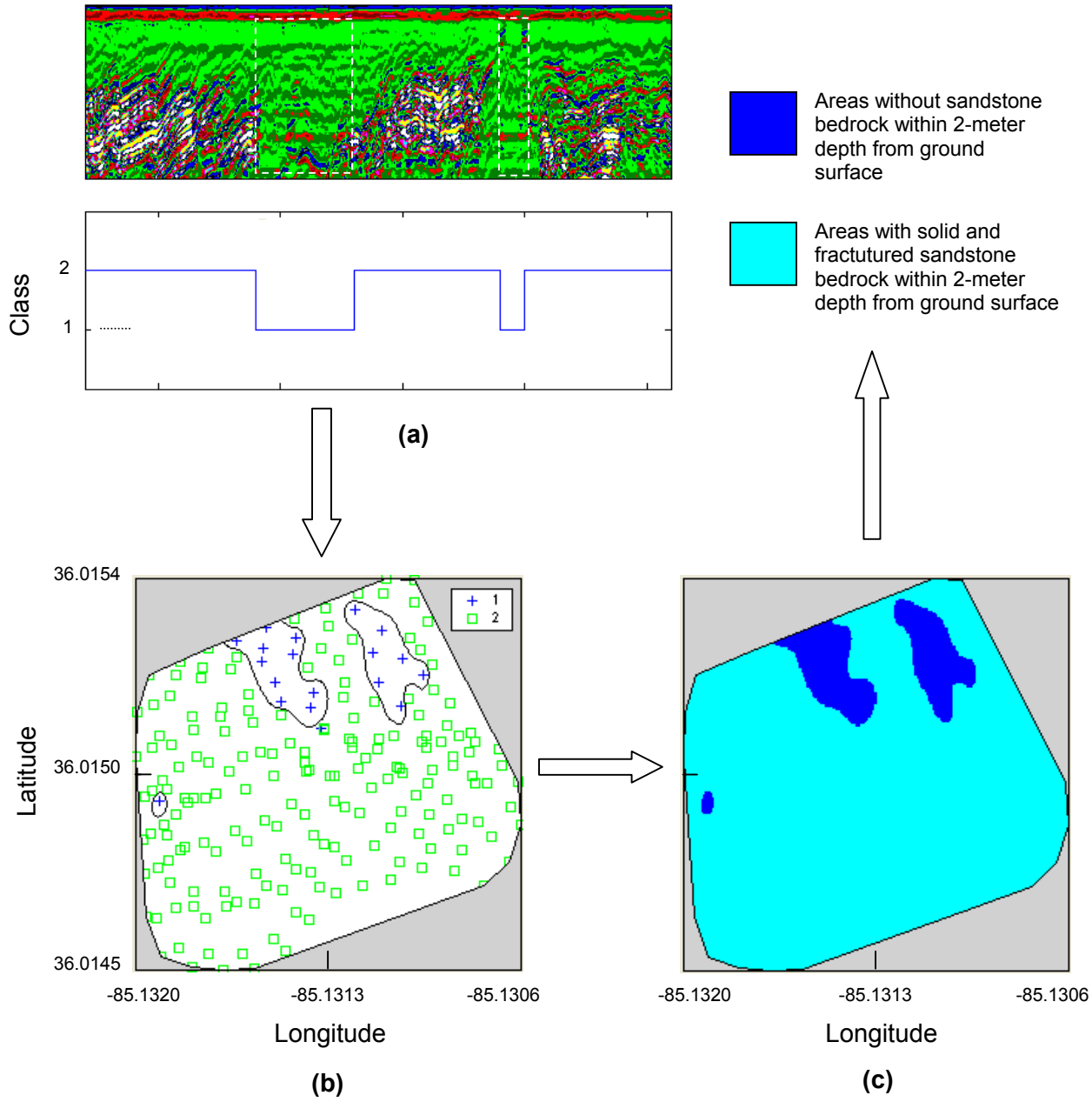

Figure 5. (a) Shows how the program separated the soil profile into two categories (areas without bedrock and areas with bedrock), (b) shows how markers on the GPR data were assigned to the different categories and surface plotted, and (c) is a surface map showing areas without and with sandstone bedrock.

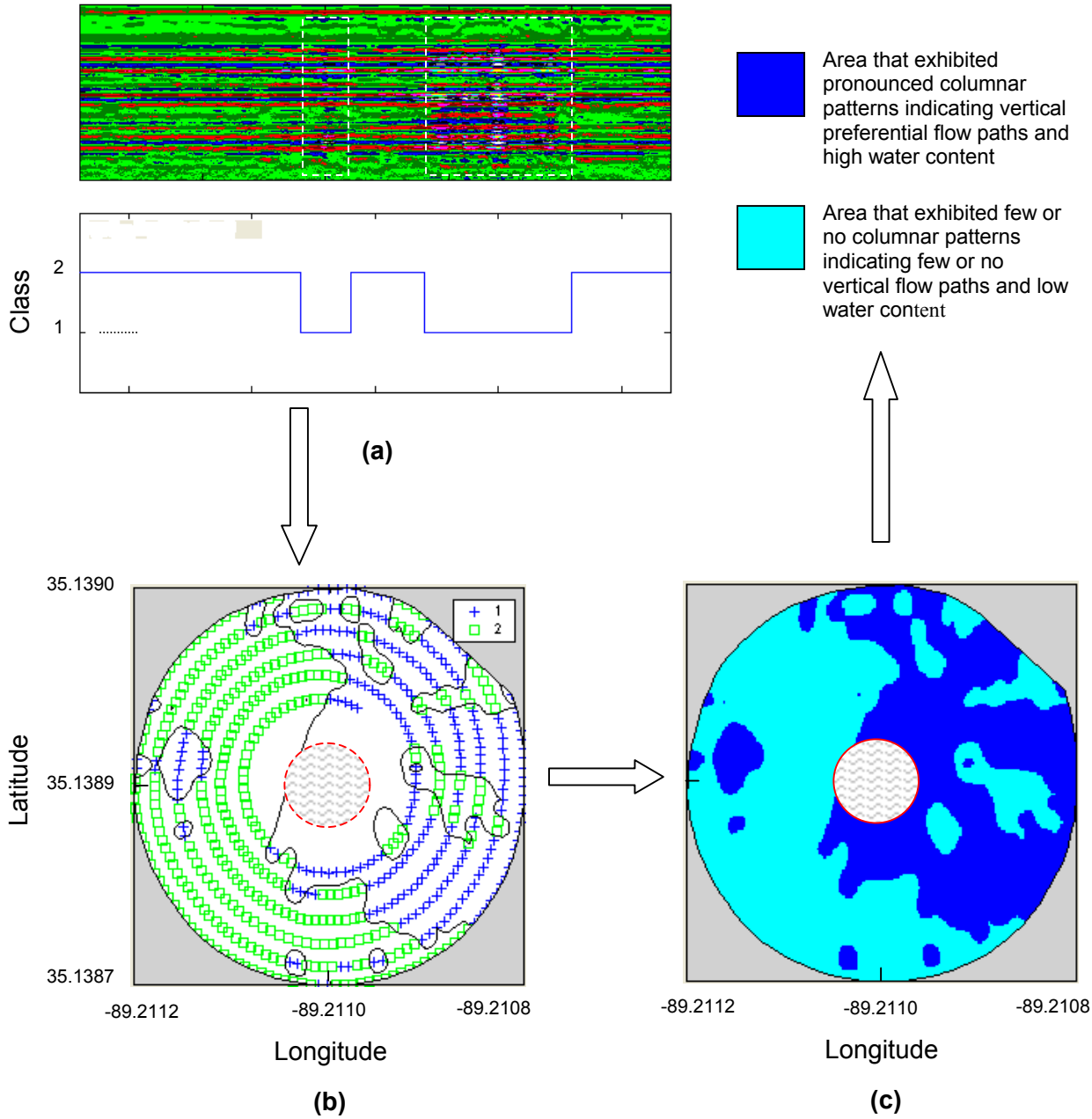

Figure 6. (a) Shows how the program separated the soil profile into two categories: areas which exhibited pronounced columnar patterns as class 1, and areas which exhibited few or no columnar patterns as class 2, (b) shows how markers on the GPR data were assigned to the different categories and surface plotted, and (c) is a surface map showing areas with pronounced columnar patterns and hence high soil water content, and areas with few or no columnar patterns and hence low soil water content.