Levee Soil Moisture Assessment Based On a Backpropagation Neural Network Using Synthetic Aperture Radar Data

Majid Mahrooghy[1,2], James Aanstoos[2], Khaled Hasan[2], and Nicolas H. Younan[1,2]

[1]Department of Electrical and Computer Engineering, Mississippi State University, Mississippi State, MS 39762, USA-
(younan@ece.msstate.edu)

[2]Geosystems Research Institute, Mississippi State University, Mississippi, Mississippi State, MS 39762, USA-
(majid, aanstoos, khaled@gri.msstate.edu)

Abstract- a methodology of soil moisture assessment around a levee based on a backpropagation neural network and using Synthetic Aperture Radar (SAR) data is developed. Soil moisture changes along a levee over a period of time can help to monitor and find potential failure indicators such as slides and sand boils around the levee. Several analytical and empirical models have shown relationships between SAR backscatter and soil moisture. The algorithm includes three steps (1) segmentation of levee area into a 300 meter buffer zone from levee centerline, and removal of the trees using a threshold; (2) extracting the backscatter and texture features from each pixel within the buffer; (3) soil conductivity calculation using a backpropagation neural network. We use SAR data from UAVSAR and conductivity measurements as ground-truth data. With 10% of the data set allocated to testing samples, the result shows an RMSE around 8 mS/m and the correlation 0.64.

Keywords: Synthetic Aperture Radar (SAR), feature extraction, neural networks,

1. INTRODUCTION

Soil moisture plays a significant role in many applications such as agriculture, atmospheric science, and hydrology. The estimation of soil moisture using synthetic aperture radar has been investigated in recent decades by many researchers. The radar backscattering coefficients are affected by soil moisture, surface roughness, and incidence angles. In addition, the backscatter data also are sensitive to low vegetation and trees. Different approaches have been developed to estimate soil moisture using SAR data. Most empirical models developed are for soils without significant vegetative cover. The empirical models of Oh et al. (1992) and Dubois et al (1995) are popular models for bare surfaces in which they use inversion techniques to retrieve the surface roughness and relative dielectric constant from bare soil. Theoretical scattering models show that the backscattering coefficients are more affected by surface roughness than soil moisture (Oh, 2004); therefore assessing soil moisture is a difficult in areas of varying roughness. Also the presence of vegetation increases the complexity of the backscattering. Thus, as our study area is covered by thick grasses and occasional shrubs, new methods of soil moisture estimation are being investigated in place of the existing empirical models (Oh, Dubois, etc) built on bare soil conditions (Hajnsek, 2009).

There are other approaches to estimating soil moisture and soil prosperities using matching learning and training samples (Liou, 2001; Chang 2000). Since the relationship between the SAR backscattering coefficients and the soil moisture is complex and non-linear, a neural network can be employed and trained to obtain the soil moisture.

The study area encompasses portions of levees built along the Mississippi River. Parts of this levee system are occasionally weakened by combinations of extreme meteorological and hydrologic events. These can result in slope failure in the form of slough slides on the riverward side and as sand boils behind the landward slope of the levee. Since increase in soil moisture usually accompanies these failures, monitoring the moisture content can provide an indication of vulnerability and impending failure of a levee segment. Changes in spatial and temporal patterns of soil moisture can reveal signs of instability and help identify zones of weakness. Such information can be used effectively by levee managers to stabilize the susceptible segments before failure occurs.

2. DATA

The broader study region covers an area of the United States extending bounded by 32°N to 34.5°N and 90.5°W to 91.5°W. SAR data was acquired over this area on 25 January 2010 by the UAVSAR platform flown by NASA. This data is quad-polarized L-band (23.79 cm) measurements collected over a 22 km wide swath at a 25-60° look angle range. The ground cell size of the multi-look, orthorectified radar data was approximately 5 by 7 meters. As a proxy for soil moisture, apparent electrical conductivity measurements were collected on the same day in a small subset of this area. These measurements were collected using an electromagnetic soil conductivity meter, model EM38-MK2 made by Geonics Corporation. The sensor in this device reads the strength of the electromagnetic fields which is proportional to the soil's electrical conductivity over a given depth range (Morris, 2009), which in turn is proportional to the moisture content (Grisso, 2009). The output is then presented in units of mS/m (millisiemens per meter). The range of values measured in our study area was 30 to 85 mS/m. The instrument was towed back and forth across a levee section of interest, creating an irregular collection of spatial samples. Samples that fell within the ground cells of pixels in the radar imagery were assigned to those coordinates.

Figure 1 shows a UAVSAR polarimetric color composite image of the study area. The inset shows the detailed area of interest within which a small subset of the levee area was sampled for soil conductivity. Figure 2 is a high-resolution aerial photograph of the portion of the detail area which includes an unrepaired slough slide, overlaid with the soil conductivity measurements that were taken. Unrepaired slide areas generally have different roughness compared to the non-slide healthy area of the levee. Because such slide areas are a major focus of the broader investigation of levee condition assessment, it was chosen as the area of interest for this soil moisture estimation study. Successful estimation of spatial patterns of soil moisture variability in such areas are expected to contribute to improvements in the levee vulnerability classification application.
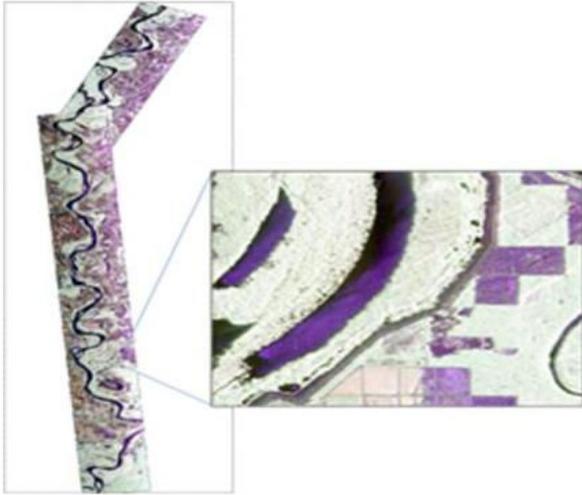
Figure 1. Radar image of the study area along the Mississippi river; Inset: close-up of area used in this investigation
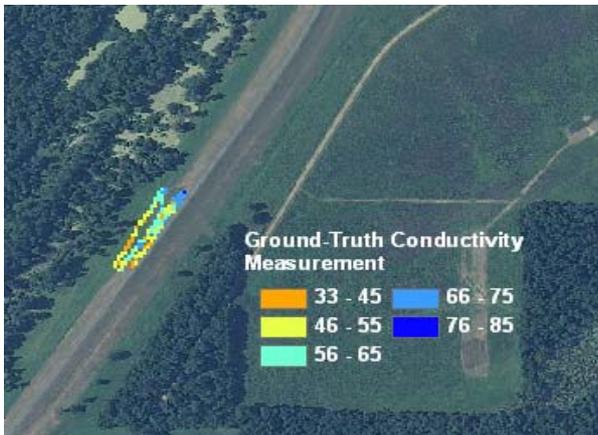

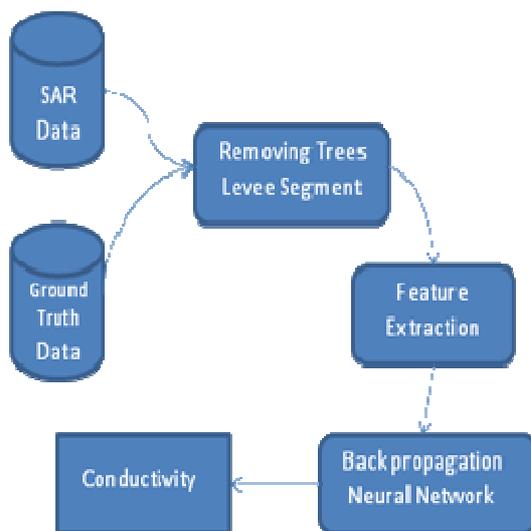Figure 2. Location of soil conductivity measurements

Figure 3. Flow chart of the methodology

3. METHODOLOGY

A flow chart of the soil moisture assessment using a back propagation neural network is depicted in Figure 3. First the areas dominated by trees along and beyond the edges of the levee are removed from the area of interest so as not to distort the statistics of the sample. Training sets are then selected based on field information from slide and non-slide areas. Features are extracted and put into a backpropagation neural network (BNN). The BNN weights and other parameters are trained based on the ground-truth data. In testing mode, the trees and the area out of the levee buffer are also removed. Then, the features at each pixel in the study area are computed and are input into the BNN using the weights derived in training mode. The output then gives the estimated soil conductivity at each pixel location.

3.1 Segment the Levee Area

Due to the difficulty of computing accurate soil moisture over tree covered areas, these areas are removed using the HV backscatter values. Figure 4 shows a histogram of two classes identified as levee area (blue) and trees (red). As can be seen, the HV value can separate the two classes effectively using a threshold of HV = -20 (db). The result of tree removal in the levee buffer segment is depicted in Figure 5 (blue mask).
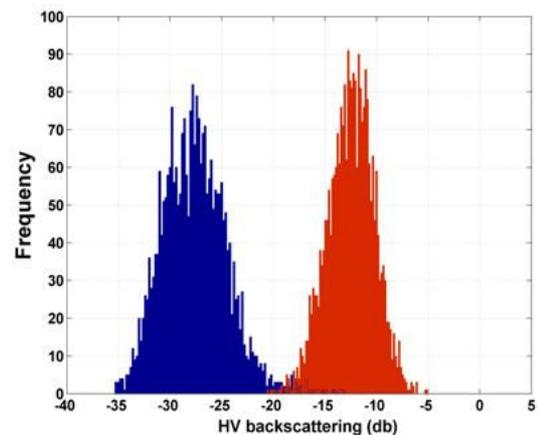

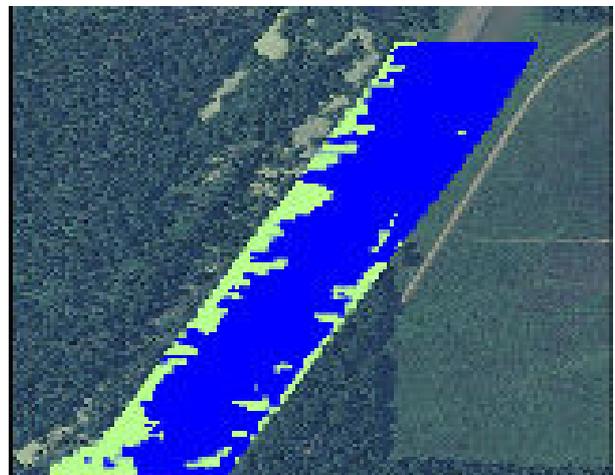Figure 4. HV backscattering histogram of levee area and trees

Figure 5. Segmented area (blue) after tree removal

3.2 Feature Extraction

In the next step, feature data are extracted from the segmented area. The features used include the magnitudes of the polarimetric backscattering coefficients HH, VV, and HV, as well as their ratios HH/VV and HV/VV, along with texture features such as window statistics (mean and variance), and wavelet features. Wavelet features used are the mean and standard deviation of the energy of approximation and vertical, horizontal, and diagonal detail coefficients of a two-level decomposition of each pixel and its neighbors (sliding window size 7). A total of 51 features were thus extracted at each pixel location.

3.3 Backpropagation Neural Network

A backpropagation neural network is a multilayer, feed-forward network trained by the backpropagation method (Fausett, 1993). It is defined by the input, output, and hidden layers; the weight parameters; and the specified transfer function. The first layer has weights (W) which are applied to the input feature values. The number of hidden layers is dependent on the application. The weights of each hidden layer are applied to the outputs of the previous layer. All layers have biases, and the last layer is the network output. The BNN used in this work includes one hidden layer, and one output neuron. Figure 6 shows the scheme of the BNN used in this work.
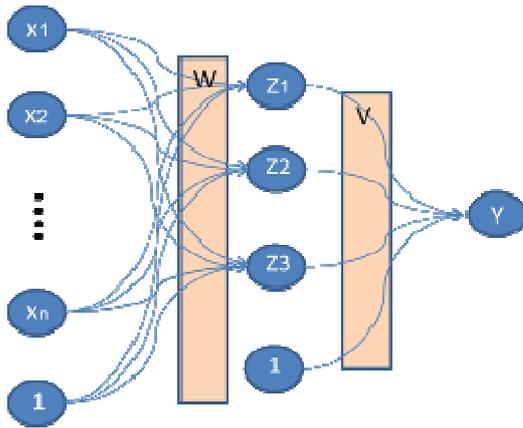

Figure 6. Backpropagation neural network with one hidden layer

The features x_i are placed into the input layer. The weights in the hidden layer (W) and output layer (V) are initialized and trained using a portion of the ground-truth measurements allocated to training. During the training phase, the errors are backpropagated from the output layer to hidden layer using the delta rule such that the total squared error of the output becomes minimum.

4. RESULTS

The BNN was trained using ground-measured conductivity data and the corresponding features derived from the SAR data.
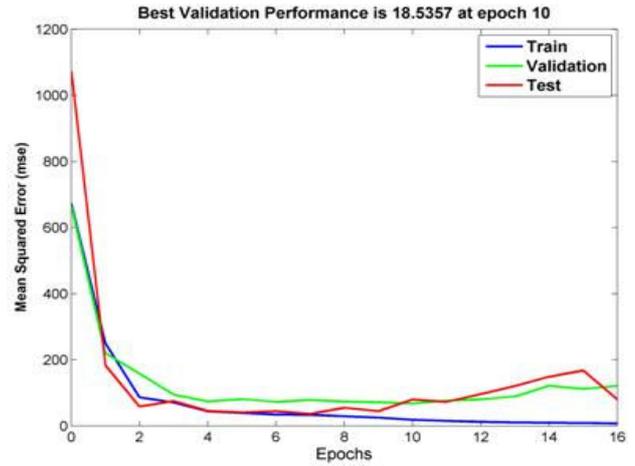

Figure 7. Training , validation, and testing the data set.

Figure 7 shows how the MSE (mean squared error) of the BNN output converges. Of the available ground truth data, 70% were used for training the BNN, 20% were used for validation, and the remaining 10% for testing. In this case, the total number of ground-truth samples is 125 pixels. Therefore the number of training, validation, and testing pixels are 87, 25, and 12, respectively. In the figure the blue, green, and red curves show the MSE of the training, validation, and test data, respectively, against the numbers of training epochs. The minimum error occurs at epoch 10 with MSE = 18 (RMSE = 4.3). The BNN is stopped at this epoch to prevent over fitting as evidenced by the validation MSE increasing beyond that point.
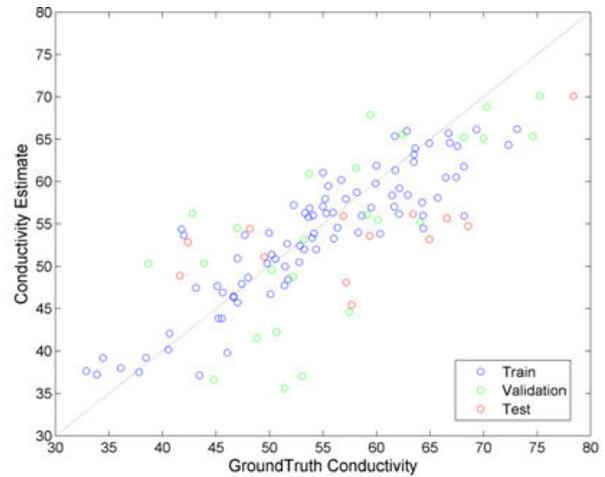

Figure 8. Scatter plot of conductivity estimates for training, validation, and testing data against the corresponding ground truth data

The scatter plot of the training (blue), validation (green) and test data (red) against the corresponding estimated data is shown in the Figure 8. It can be seen that the data cluster fairly well around the zero line (output = target), with somewhat higher deviations for the validation and test data than the training data. The statistics from these data are summarized in Table 1.

Table 1. Statistical results for each data set

Data/ Validation	Train	Validation	Test
RMSE	4.3	8.2	8.9
Correlation coefficient	0.89	0.68	0.64

Table 1 shows the RMSE and correlation coefficients for the three data sets. The RMSE for the training, validation and test data are 4.3, 8.2 and 8.9 mS/m, respectively.   The corresponding correlation coefficients are 0.89, 0.68, and 0.64.

The significant increase in error level for the validation and testing data sets is probably indicative of an insufficient number of training samples, particularly in light of the relatively high number of features used.

Figure 9 shows the spatial distribution of the conductivity estimates for the segmented area (the blue mask from Figure 5). This figure shows that the river side (left side) of the levee has higher conductivity than the land side (right side). It can be seen that the slide area has the highest conductivity (dark blue). This is consistent with physical aspects of this environment.
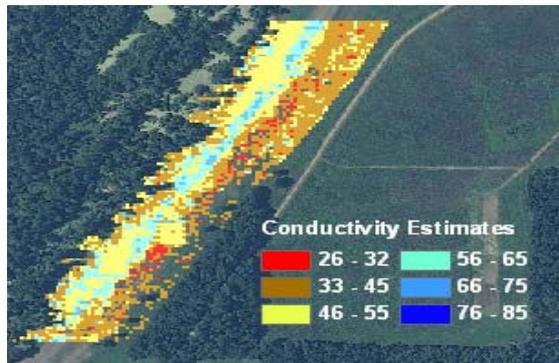

Figure 9. The conductivity estimate for the segmented area

5. CONCLUSION

In this work, an algorithm based on a backpropagation neural network using a variety of radiometric and textural features derived from polarimetric radar backscatter was developed to estimate the conductivity, and thus relative soil moisture, of an area around an earthen levee system. Ground truth data for training and validation was obtained using a Geonics EM38-MK2 measuring apparent soil conductivity. Results for the limited study area included an RMSE around 8 mS/m with a correlation coefficient of 0.64. The conductivity estimates for the area of interest including a slough slide area are physically consistent for soil moisture in such an environment, and may prove useful in the application of airborne SAR to the problem of levee condition assessment.

Many more samples of ground-truth measurements are needed, and taken over a wider area, before broader conclusions can be drawn. This study was limited by the number of samples that were acquired during an airborne radar data collection flight. Future work will address these limitations. Another ongoing aspect of the study is the evaluation of the use of satellite-based SAR versus the airborne UAVSAR used in this investigation. The currently available spaceborne SAR platforms use shorter-wavelength instruments, which do not penetrate soil as much as the L-band UAVSAR, thus adding and additional complications to their use for soil moisture estimation.

REFERENCES

D. Chang, S. Islam, "Estimation of Soil Physical Properties Using Remote Sensing and Artificial Neural Network," *Remote Sensing of Environment*, vol.74, Issue 3, pp.534-544, December 2000.

P. C. Dubois, J. van Zyl, and E. T. Engman, "Measuring soil moisture with imaging radar", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, pp.916 -926, 1995

L.V. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*, Prentice Hall, 1993.

Geonics Limited, http://www.fugroinstruments.com/pdf/Fugro_EM38-Mk2.pdf

R. Grisso, M. Alley, W.G. Wysor, D. Holshouser, and W. Thomason, "Precision Farming Tools: Soil Electrical conductivity," Virginia Tech university, Virginia Cooperative Extension , May 1, 2009. http://pubs.ext.vt.edu/442/442-508/442-508.html

I. Hajnsek, T. Jagdhuber, H. Schon, K.P. Papathanassiou, "Potential of Estimating Soil Moisture under Vegetation Cover by Means of PolSAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol.47, no.2, pp.442-454, Feb. 2009.

Y. Liou, S. Liu, W. Wang, "Retrieving soil moisture from simulated brightness temperatures by a neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol.39, no.8, pp.1662-1672, Aug 2001.

E. R. Morris, "Height-above-ground effects on penetration depth and response of electromagnetic induction soil conductivity meters," *Computer. Electronic. Agric*, vol. 68, 2, pp. 150-156, 2009.

Y. Oh, "Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces," *IEEE Transactions on Geoscience and Remote Sensing*, vol.42, no.3, pp. 596- 601, March 2004.

Y. Oh, K. Sarabandi, and F. T. Ulaby, "An empirical model and an inversion technique for radar scattering from bare soil surfaces", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp.370 -382 1992.