Characterizing Lithofacies from Geophysical Data Using the Bayesian Model coupled with a Fuzzy Neural Network

Jinsong Chen and Yoram Rubin

Department of Civil and Environmental Engineering
University of California at Berkeley

Abstract: *A Bayesian model coupled with a fuzzy neural network (BFNN) is developed to alleviate the difficulty of using geophysical data in lithology estimation when cross correlation between lithology and geophysical attributes is nonlinear. The prior estimate is inferred from borehole lithology measurements using indicator kriging based on spatial correlation, and the posterior estimate is obtained from updating of the prior using the geophysical data. The novelty of the study lies in the use of a fuzzy neural network for the inference of the likelihood function. This allows incorporating spatial correlation as well as a nonlinear cross correlation into lithology estimation. The effectiveness of the BFNN is demonstrated using synthetic data generated from measurements at the Lawrence Livermore National Laboratory (LLNL) site.*

1. Introduction

Heterogeneity of lithology has an important effect on determination of hydrogeological parameters. Since traditional methods for characterizing lithofacies rely heavily on expensive and invasive lithology core measurements, many efforts have been made to incorporate geophysical data into lithology estimation. The crucial part of the incorporation is to connect geophysical data to lithology through a possibly complex cross correlation [*Copty and Rubin*, 1995].

Several models have been used to estimate lithology from lithology measurements and geophysical data, such as indicator kriging, indicator cokriging, and neural networks or fuzzy neural networks. Indicator kriging uses only borehole lithology measurements but completely ignores geophysical information. Neural networks or fuzzy neural networks, however, use only geophysical data but ignore borehole lithology measurements. Indicator cokriging does use both borehole lithology measurements and geophysical data, but it is limited when cross correlation between lithology and geophysical attributes is highly nonlinear. This study develops an innovative model to incorporate geophysical data into lithology estimation using spatial correlation of lithology as well as a nonlinear cross correlation between lithology and geophysical attributes.

2. Bayesian Model

The developed model combines geophysical data with borehole lithology measurements to estimate lithology using a Bayesian framework. Let $Z(x)$ be a categorical random variable at location x defined on $K = \{1, 2, \Lambda, d\}$, where d is the total number of lithofacies. Let $z(x_i)$ be a lithology measurement at location $x_i, i \in \{1, 2, \Lambda, n\}$ and $g_1(x)$ and $g_2(x)$ be geophysical data at location x. As the Markov assumption is valid [*Almeida and Journel*, 1994], the Bayesian model is given by

$$f_{post}(Z(x) = k) = CL(Z(x) = k \mid g_1(x), g_2(x)) f_{prior}(Z(x) = k),$$

where C is a normalizing constant and $L(\cdot)$ is a likelihood function. $f_{prior}(Z(x) = k)$ is the prior probability estimated from lithology measurements using indicator kriging, and $f_{post}(Z(x) = k)$ is the posterior probability obtained from updating of the prior using collocated geophysical data through the likelihood function.

The key to using this model is to infer the likelihood function from a nonlinear cross correlation using a fuzzy neural network. The structure of the fuzzy neural network is similar to the one given by *Takagi and Sugeno* [1985], which consists of several inference rules. The input of the system is geophysical data, and the output is the log likelihood with the normalizing constant. We apply all fuzzy rules to a given input, and the final result is a combination of the outputs from each rule. Training the fuzzy neural network requires estimating the number of rules and parameters associated with those rules from a training data set. In this study, we use the fuzzy cluster analysis method to identify the number of rules and the Levenberg-Marguardt method to identify the parameters.

3. Case Study

This case study demonstrates the effectiveness of the BFNN in lithology estimation using synthetic data generated from measurements at the LLNL site by comparing the BFNN with indicator kriging, indicator cokriging and the fuzzy neural network (FNN) that does not use lithology measurements. We will generate three two-dimensional random fields: a lithology field with sand and silt from borehole lithology measurements using the sequential indicator simulation (SIS) [*Deutsch and Journel*, 1998], a gamma-ray shaliness field conditioned to the previously generated collocated lithology and borehole gamma-ray shaliness using the sequential Gaussian simulation (SGS) [*Deutsch and Journel*, 1998], and a resistivity field conditioned to the collocated lithology and gamma-ray shaliness using the parameters given by *Ezzedine et al.* [1999].

The generated lithology and geophysical data will be used to evaluate the performance of each model. We first randomly select eight columns from each generated random field to mimic boreholes in a real situation and then use data at those boreholes to train each model. The trained models are used to estimate lithology at any testing location, and the total numbers of misclassifications are counted according to the minimum distances of testing locations from the boreholes for each model.

Figure 1 shows cross correlation between gamma-ray shaliness and electrical resistivity according to the data at the boreholes, and it is nonlinear and non-unique. Figure 2 compares performances of indicator kriging, indicator cokriging, FNN and the BFNN in terms of percentages of misclassifications. It is evident that spatial correlation is important when a testing location is in the close vicinity of the boreholes and that cross correlation is important when a testing location is in the region far away from the boreholes. Otherwise, both spatial correlation and cross correlation are important for lithology estimation.
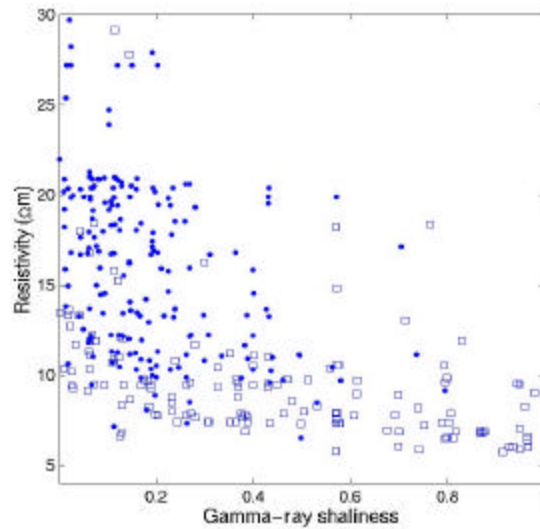
Figure 1. Cross-plot based on data at the boreholes (dots—sand and squares—silt)
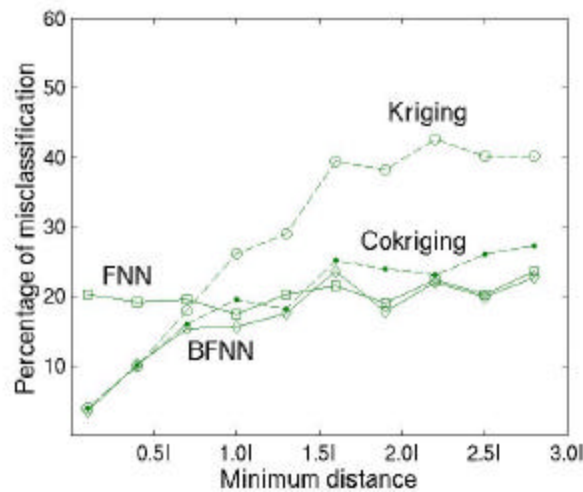

Figure 2. Model comparison in terms of misclassification, where I=10m is the integral length of sand at the LLNL site. The horizontal coordinate is the minimum distance of a testing location from the boreholes.

To compare the BFNN with indicator cokriging further and explore nonlinearity effects of cross correlation between lithology and geophysical attributes, we generate three two-dimensional lithology fields with two, three and four lithofacies, respectively. Following a similar procedure as before, we generate two-dimensional geophysical data for each lithology field and select several columns as boreholes. After training each model using the data at the boreholes, we can estimate lithology at any given location and compare the estimated results with the true values to evaluate model performances. The nonlinearity of cross correlation between lithology and geophysical attributes generally increases with the increase of the number of lithofacies. Testing results show that when there are two lithofacies, the BFNN and indicator cokriging have similar performances in

lithology estimation. However, the BFNN has a much better performance than indicator cokriging as the number of lithofacies or nonlinearity of cross-correlation increases.

4. Discussion

The BFNN is the most effective model for incorporating geophysical data into lithology estimation among indicator kriging, indicator cokriging, FNN and the BFNN. The BFNN has a similar performance as kriging when an estimating location is close to boreholes and a similar performance as FNN when an estimating location is far away from boreholes. The BFNN is particularly useful compared to indicator cokriging when cross correlation between lithology and geophysical attributes is highly nonlinear.

Although the BFNN is oriented toward the LLNL project where we have two different geophysical attributes that have been shown most informative to lithology estimation, it can be directly used for the cases where there are more than two types of geophysical data, such as in *Doveton* [1986]. The reason is that the fuzzy neural network can be used to extract complex patterns inherent in multidimensional data, which are very difficult to be extracted using other methods.

The limitation of the fuzzy neural network lies in the assumption that each variable is approximately parallel or perpendicular to axes, which is valid for many applications. In other cases, however, we need to either rotate coordinates using the principal component analysis or develop a more general neural network to estimate likelihood functions.

References:

Almeida, Alberto S. and Andre G. Journel, Joint simulation of multiple variables with a Markov-type coregionalization model. *Mathematical Geology*, 26(5), 1994.

Copty, Nadim and Yoram Rubin, A stochastic approach to the characterization of lithofacies from surface seismic and well data. *Water Resources Research*, 31(7), 1995.

Deutsch, Clayton and Andre Journel, *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 1998.

Doveton, John H., *Log analysis of subsurface geology concepts and computer methods*. John Wiley & Sons, New York, 1986.

Ezzedine, Souheil, Yoram Rubin and Jinsong Chen, Hydrological-geophysical Bayesian method for subsurface site characterization: theory and application to the LLNL superfund site. *Water Resources Research*, 35(9), 1999.

Takagi, Tomohiro and Michio Sugeno, Fuzzy identification of systems and its applications to modeling and control. *IEEE Transaction on Systems, Man, and Cybernetics*, SMC-15(1), 1985.