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 lithofacies estimation when cross correlation between lithofacies and geophysical attributes is nonlinear. The prior estimate is inferred from borehole lithofacies 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 likelihood function. This allows incorporating spatial correlation as well as a nonlinear cross correlation into lithofacies estimation. The effectiveness of the BFNN is demonstrated using synthetic data generated from measurements at the Lawrence Livemore National Laboratory (LLNL) site.*

1. Introduction

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

Several models have been used to estimate lithofacies from lithofacies measurements and geophysical data, such as indicator kriging, indicator cokriging, and neural networks or fuzzy neural networks. Indicator kriging uses only borehole lithofacies measurements but ignores geophysical information, and neural networks or fuzzy neural networks [*Rogers et al.,* 1992], however, use only geophysical data but ignore borehole lithofacies measurements. Indicator cokriging does combine borehole lithofacies measurements and geophysical data, but it is limited when cross correlation between lithofacies and geophysical attributes is highly nonlinear. This study develops an innovative model to incorporate geophysical data into lithofacies estimation using spatial correlation of lithofacies as well as a nonlinear cross correlation between lithofacies and geophysical attributes.

2. Bayesian Model

The developed model integrates geophysical data into lithofacies estimation using a Bayesian framework. Let $Z(x)$ be a categorical random variable at location x defined on $K = \{1,2,\cdots,d\}$, where d is the total number of lithofacies. Let $z(x_i)$ be a lithofacies measurement at location $x_i, i \in \{1,2,\cdots,n\}$, and $g_1(x)$ and $g_2(x)$ be the geophysical data at location x. Consider the Markov assumption [*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 lithofacies 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 given data set. Consider mathematical tractability, we often assume that the geophysical attributes at each location have a multivariate normal distribution. This assumption, however, is not valid in some situations, especially when the number of attributes is more than two and nonlinearity of cross correlation is high. In this study, we use a fuzzy neural network to infer the likelihood function directly from the given data set without making any assumption about the form of the likelihood function.

Figure 1 shows the structure of the fuzzy neural network, which is similar to the one given by *Takagi and Sugeno* [1985]. This system consists of several inference rules and has been shown very efficient in fitting nonlinear functions from low quality data. The input of the system is the 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. The training of the fuzzy neural network includes structure identification and parameter identification. The structure identification essentially is to determine the number of inference rules using the fuzzy cluster analysis, which is a method to find structure inherent in the given data set. The parameter identification is to estimate all the parameters associated with those rules using a hybrid algorithm combining the least squares estimation and the Levenberg-Marquart method.

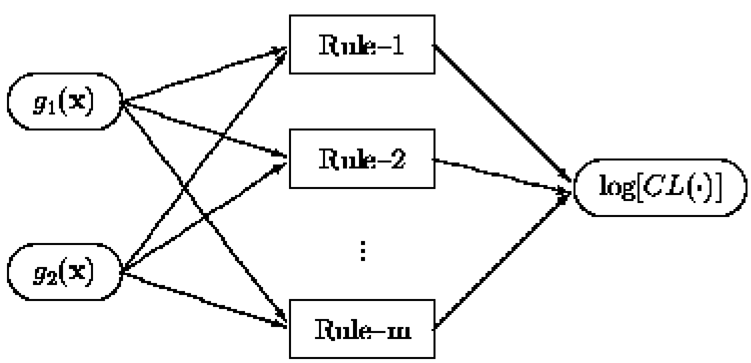

Figure 1. Structure of the fuzzy neural network

3. Case Study

This case study demonstrates the effectiveness of the BFNN in lithofacies 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 without using lithofacies measurements (FNN). We generate a two-dimensional lithofacies field with sand and silt from borehole lithofacies measurements using the sequential indicator simulation (SIS), a two-dimensional gamma-ray shaliness field conditioned to the previously generated collocated lithofacies and borehole gamma-ray shaliness using the sequential Gaussian simulation (SGS) [*Deutsch and Journel*, 1998], and a two-dimensional resistivity field conditioned to the collocated lithofacies and gamma-ray shaliness using the parameters given by *Ezzedine et al.* [1999].

The generated lithofacies and geophysical data will be used to evaluate performances of each model. We first randomly select eight columns from each generated random field to mimic boreholes in a real situation and then use the data at those boreholes to train each model. The trained models are used to estimate lithofacies at any testing location, and the total numbers of misclassification 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 misclassification. It is evident that spatial correlation is important when a testing location is in the short 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 lithofacies estimation.

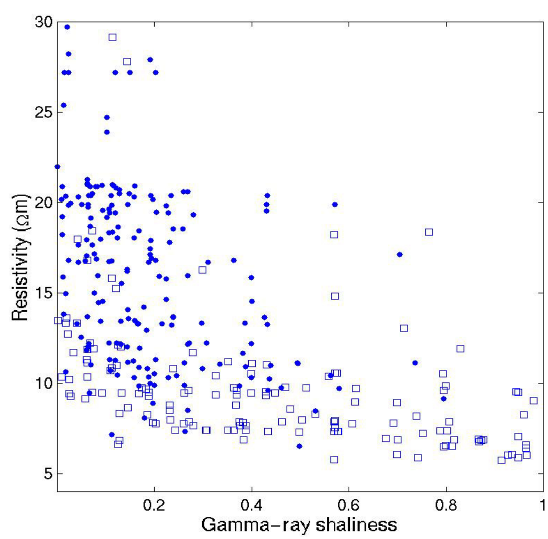

Figure 2. Cross-plot based on data at the boreholes (dots—sand and squares—silt)

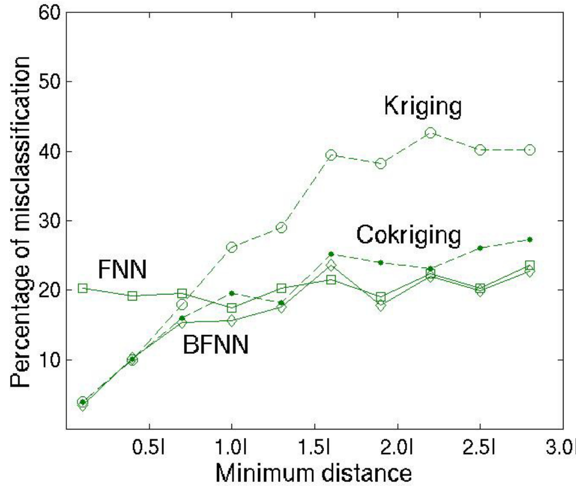

Figure 3. Model comparison in terms of misclassification, where I=10m is the integral length of sand at the LLNL site. The horizontal coordinate is the minimum distance of a testing location from the boreholes.

To compare the BFNN with indicator cokriging further and explore nonlinearity effects of cross correlation between lithofacies and geophysical data, we generate three two-dimensional lithofacies fields with two, three and four lithofacies, respectively. Following a similar procedure as before, we generate two-dimensional geophysical data for each lithofacies field and select several columns as boreholes. After training each model using the data at the boreholes, we can estimate lithofacies at any given location and compare the estimated results with the true values to evaluate model performances. The nonlinearity of cross correlation between lithofacies 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 lithofacies estimation. However, the difference between BFNN and indicator cokriging becomes more obvious as the number of lithofacies or nonlinearity of cross correlation increases.

4. Discussion

The BFNN is the most effective model for integrating geophysical data into lithofacies estimation compared to indicator kriging, indicator cokriging and FNN in terms of the previous case study. This is because the method takes advantages of the Bayesian model for combining different types of information and the fuzzy neural network for fitting a nonlinear cross correlation. 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 when cross correlation between lithofacies and geophysical attributes is highly nonlinear compared to indicator cokriging.

Although the BFNN is oriented toward the LLNL project where we have two different geophysical attributes that have been shown most informative to lithofacies estimation, it can be directly used to the cases where there are more than two geophysical data, such as in *Doveton* [1986]. The reason is that the fuzzy neural network can be used to extract complex patterns inherent in multi-dimensional data, which otherwise are very difficult 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, we need either rotate coordinates using the principal component analysis or develop a more general neural network to estimate likelihood functions.

Acknowledgement - This study was support by NSF Grant EAR-9628306.

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.

Rogers, Samuel J., J. H. Fang, C. L. Karr, and D. A. Stanley, Determination of lithology from well logs using a neural network. *The American Association of Petroleum Bulletin*, 76(5), 1992.

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.