NEURAL NETWORKS AS AN ALTERNATIVE TO STATISTICAL MODELING IN THE SEMIVARIOGRAM ANALYSIS PORTION OF CO–KRIGING PROCEDURES

Elaine M. Giroux
Master of Science Candidate
Geomatics Program
Department of Forest Sciences
Colorado State University

Denis J. Dean
Associate Professor
Geomatics Program
Department of Forest Sciences
Colorado State University

ABSTRACT

The purpose of this study is to create and evaluate a hybrid interpolation procedure based on an extension of co-kriging. This hybrid will replace the statistical model used in the semivariogram analysis portion of traditional co-kriging with an artificial neural network, but retain all other portions of the co-kriging approach. This hybrid should combine the optimal estimation qualities of co-kriging with the superior pattern recognition capabilities of neural network analysis.

Co-kriging is regarded as the optimal linear estimator for interpolating spatial point data when covariates are available. However, co-kriging is dependent upon regression analysis to quantify relationships between covariates, semivariance and distances between sample points. This regression analysis typically involves very noisy data and consequently has relatively poor predictive power.

Previous studies have shown that when compared to traditional regression techniques, artificial neural networks (ANNs) have superior capabilities in recognizing patterns and discovering relationships in noisy data sets. ANNs provide an obvious and appealing alternative to the regression techniques used in co-kriging.

This study will develop interpolation models based both on traditional co-kriging procedures and upon a hybrid co-kriging/ANN approach for a standard climatic data set. The absolute and relative accuracies of these interpolation procedures will be compared, and the strengths and weaknesses of the two approaches will be identified.

INTRODUCTION

Various mathematical techniques such as ordinary kriging, co-kriging, inverse distance weighting, minimum curvature, and so on are used for spatial interpolation and extrapolation of many types of sample data. Some of these procedures rely upon spatial covariates; e.g., variables whose values are spatially correlated with the variable being interpolated. Co-kriging appears to be one of the most popular of these covariate-dependent interpolation techniques.

Co-kriging

Co-kriging is an advanced geostatistical procedure largely based on semivariograms and cross-variograms. In practice, a semivariogram simply enumerates the relationship between the degree of similarity between two measurements of some variable *X* (as measured by the squared difference between the two measurements; this value is called the semivariance) and the distance between the two points at which *X* was measured (this distance is termed the lag) (Figure 1).

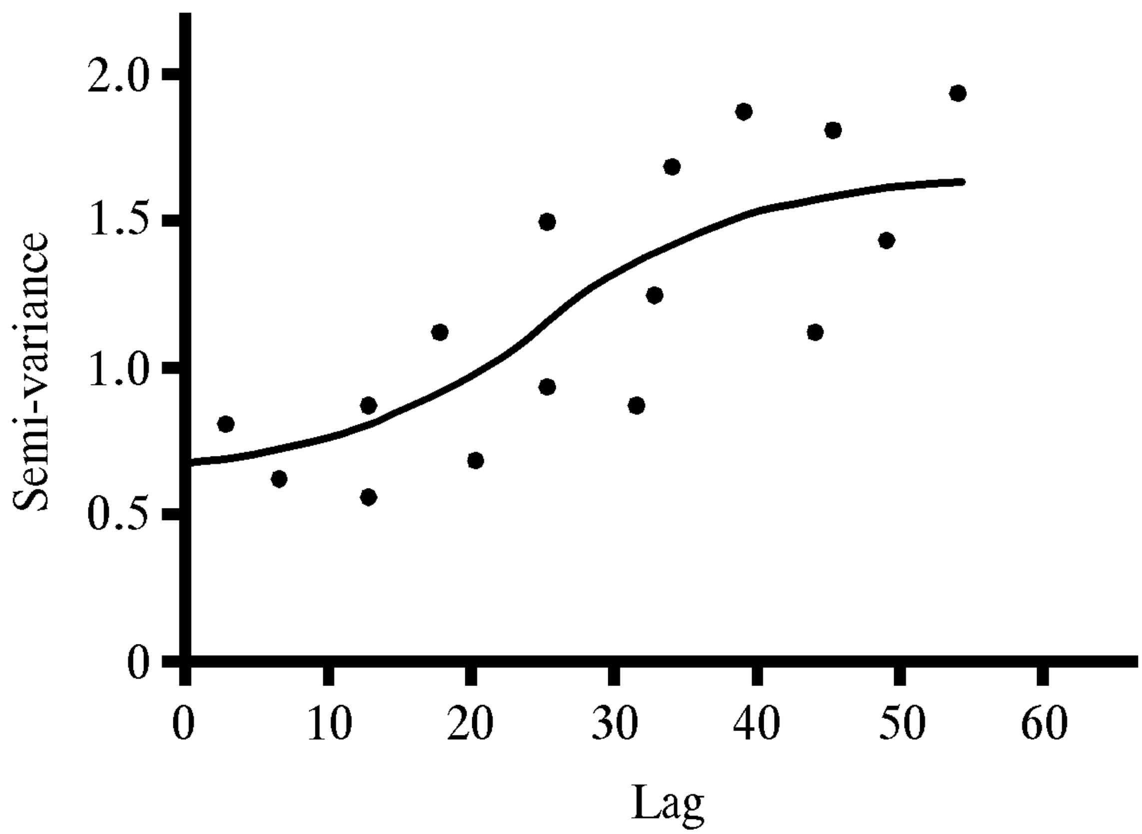

Figure 1. Example semivariogram

A cross-variogram extends this approach into the multivariate arena by measuring the relationship between the covariance between two variables (as measured by the squared difference between measurements of variable 1 multiplied by the squared difference between measurements of variable 2) and the distance between the points at which these variables were measured (Figure 2) (Cressie, 1991; Kaluzny et al., 1998).

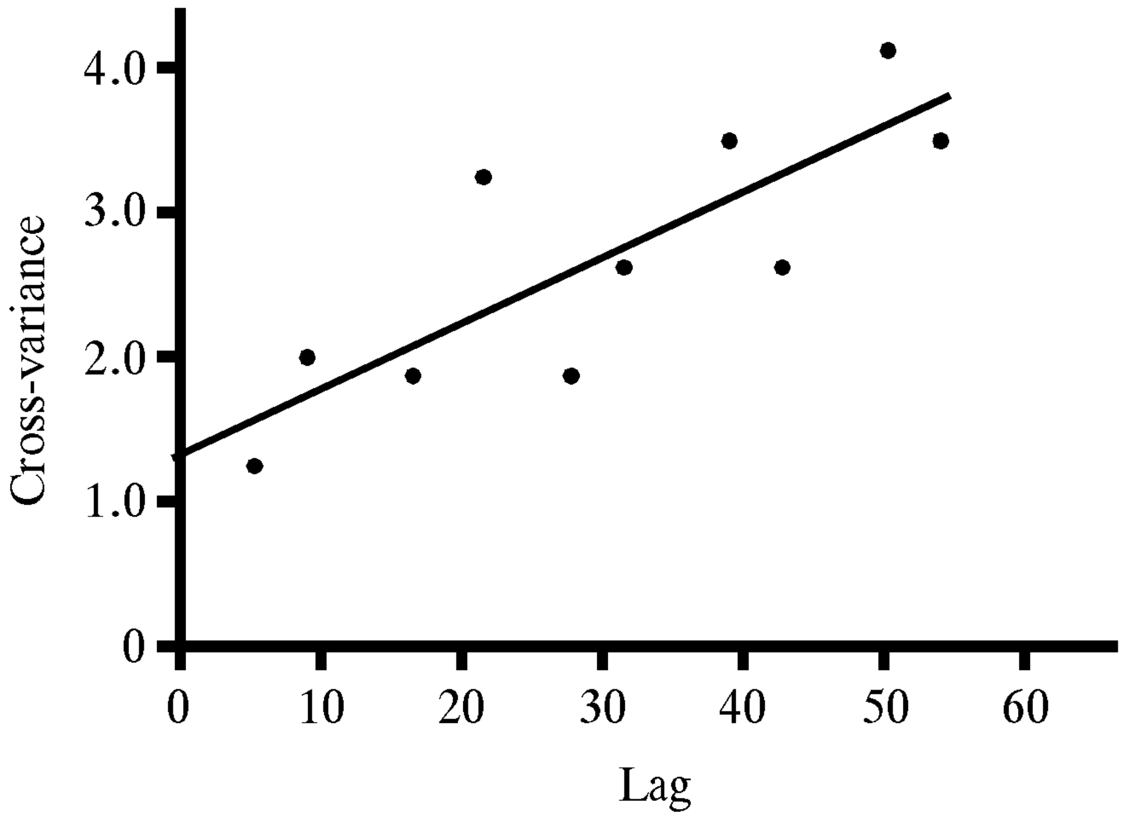

Figure 2. Example cross-variogram

For both the semivariogram and the cross-variogram, the relationship between semivariance/covariance and lag is statistically quantified by fitting some form of intrinsically linear equation (e.g., spherical, exponential, linear, gaussian, etc) to the observed sample data via some form of regression analysis.

In a typical co-kriging analysis, a variable being interpolated (call this variable *Z1*) is measured at a series of sample locations. Values of a second covariate variable (call this variable *Z2*) are also measured at most or all of these sample locations, as well as at a variety of other locations. The *Z1* semivariogram, *Z2* semivariogram, and *Z1* ? *Z2* cross-variogram are then computed using regression analyses. Once these regression analyses are complete, interpolated values can be computed based on the distance from the interpolated point to nearby sample locations for *Z1*, the variogram for *Z1*, the distance to nearby sample locations for *Z2*, the variogram for *Z2*, and the cross-variogram for *Z1* ? *Z2*. Estimated values computed in this way can be more accurate than estimates computed without the use of covariate *Z2*, especially if *Z1* and *Z2* are significantly correlated.

In practice, the regression analyses used to derive the semivariograms and cross-variograms used in co-kriging are based on data that does not conform to all the assumptions inherent in regression analysis. The data points used in these analyses are derived from pairs of sample points. For example, two data points where both *Z1* and *Z2* were measured can be used to obtain a single data point for the *Z1* semivariogram regression analysis (the dependent variable for the regression data point is the squared difference between the two measurements of *Z1* and the independent variable is the distance separating the sample points), a data point for the *Z2* semivariogram regression analysis (the dependent variable being the squared difference between the *Z2* measurements and the independent variable once again being distance between the sample points), and a data point from the *Z1* ? *Z2* cross-variogram analysis (with the dependent variable being the squared difference between the *Z1* measurements times the squared difference between the *Z2* measurements and the independent variable again being the distance between the sample points). Unfortunately, experience has shown that while there typically is a relationship between the independent and dependent variables in these regression analyses, the relationship is often hidden amongst a great deal of extraneous data variability (noise). Furthermore, the data typically exhibits heteroskedasticity (unequal variances over the range of the data). It is also true that since the data points used in these regression analyses are based on pairs of sample points, and since a single sample point may be part of multiple pairs of points, it is not correct to assume that the regression data points are truly independent of one another. As a result of all these problems, it is doubtful that the standard regression analyses used in the co-kriging process reliably find the best relationship between dependent and independent variables.

An Alternative to Regression Analysis

Artificial neural networks (ANNs) provide an alternative to regression analysis for quantifying the relationship between sets of independent and dependent variables. ANNs have the advantage of making no assumptions regarding data variance or distribution, sample independence, or mathematical form of the relationships between variables. In comparative studies, they have been found to produce at least as good, and frequently better, predictive models than standard statistical techniques (Blackard and Dean, 1999).

An ANN consists of a number of simple, interconnected processing units called nodes (Gurney, 1997). The nodes are arranged in layers (Figure 3). The input layer consists of one node for each independent variable and the output layer contains one node for each dependent variable. Between the input and output layers are one of more hidden layers containing as many nodes as the analyst feels is required to fully capture the complexity of the relationship between the independent and dependent variables.

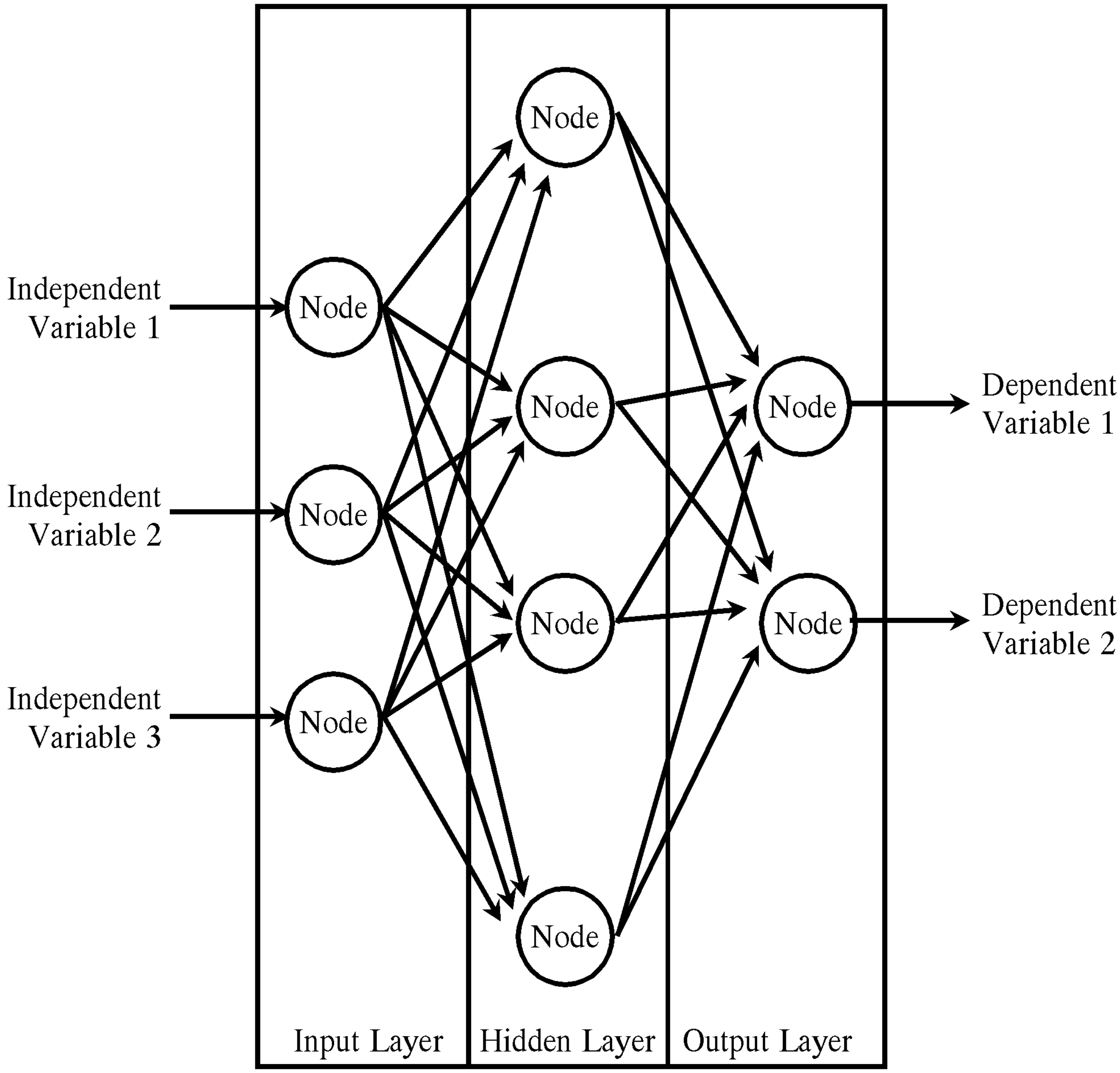

Figure 3. Schematic diagram of a simple neural network

Each node in a neural network receives input signals from either the independent input variables (in the case of the input layer) or from the nodes to which its is connected (for all nodes in layers other than the input layer). Internally, each node contains a simple activation function that determines if the sum of the node's inputs are sufficient for it to send output signals to nodes in the subsequent layers, and if so, how strong a signal should be sent. In the case of the output layer, a node's output signal is used as the value of the dependent variable the node represents (Figure 4).

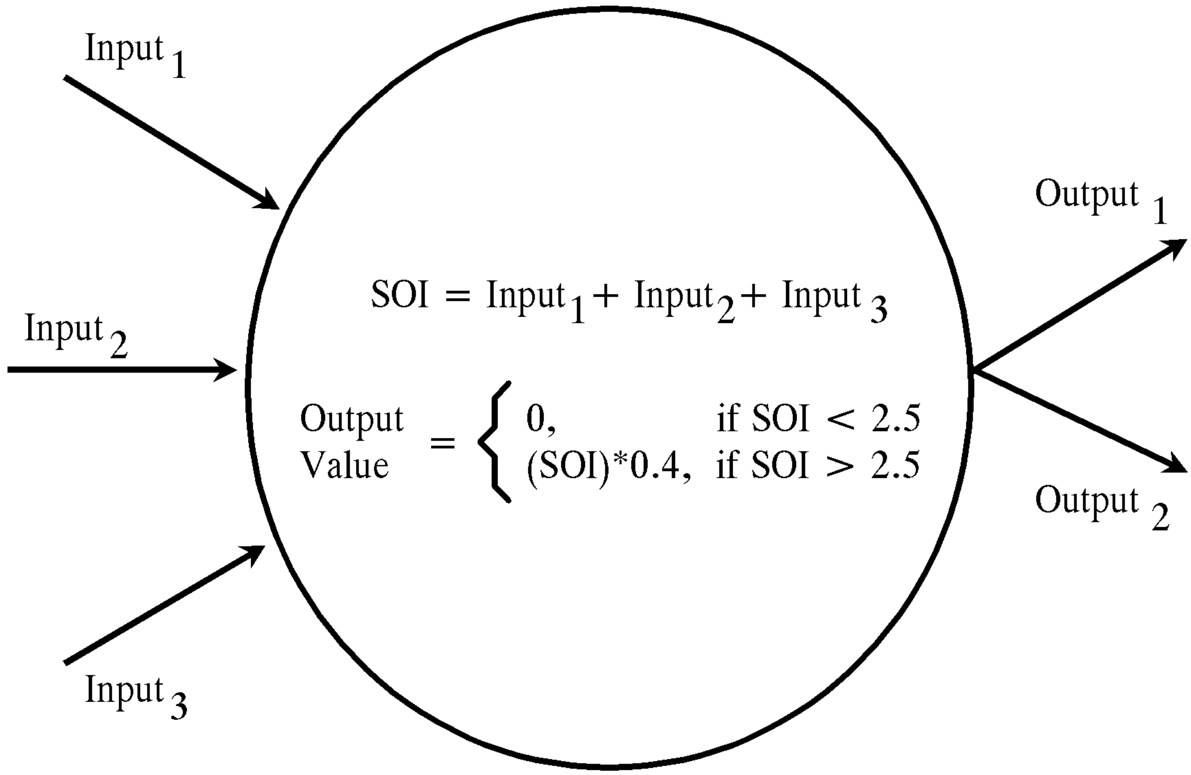

Figure 4. A hypothetical activation function within a single node of an ANN.

Training an ANN involves quantifying each node's activation function. This is usually accomplished by first specifying *a priori* what mathematical form each activation function should take. The ANN is then presented with a training data set containing multiple matched sets of independent and dependent variables. A training process then tries to parameterize each node's activation function so that some measure of error (involving a comparison of known dependent variable values to predicted values) is minimized. There are numerous training algorithms available, but by far the most widely used is the backpropagation algorithm. Backpropagation is described in detail in many existing texts, including Gurney (1997), Bishop (1995) and Ripley (1996).

ANNs make no assumptions about the mathematical form of the relationship between independent and dependent variables, the statistical distributions of these variables, the variance properties of the data, or the independence of the observations in the training data set. Given these characteristics, it is plausible to assume that ANNs should do at least as good a job as regression techniques in quantifying the relationships between semivariance/cross-variance and lag distance. Consequently, this study proposes to develop and test a hybrid co-kriging/ANN interpolation system that replaces co-kriging's regression-based approach to quantifying semivariance/cross-variance – lag distance relationships, but otherwise retains all of the properties of traditional co-kriging. It is hoped that this hybrid technique will produce more accurate interpolations than the traditional co-kriging approach.

PREVIOUS WORK

While no previous studies have proposed or examined the hybrid co-kriging/ANN approach envisioned in this study, a number of previous authors have compared co-kriging and other ANN-based interpolation techniques. Most of these comparisons have involved traditional co-kriging and an alternative interpolation system that was entirely dependent on an ANN. Typically, these ANN-based interpolators use as inputs (1) the distances between a point at which a value is to be interpolated and nearby sample points, and (2) the values of either the variable being interpolated and/or a covariate variable measured at those sample points. As outputs, these ANNs directly predict the value variable of the variable being interpolated at the point of interest.

Matsoukas et al. (1999) found that with the increased availability of rainfall measurements from multiple sensors having different characteristics, issues of sensor fusion and intercomparison became critical. Cokriging was used as one method of choice because of its best linear unbiased estimator (BLUE) properties. However, the authors found that the results of a co-kriging analysis on their data produced obviously unrealistic results. They attributed this to poor estimation of semivariograms and cross-variograms. As an alternative, they used an ANN-based approach and found that the ANN provided much more realistic results.

Mukhopadhyay (1999) compared co-kriging and ANNs for estimating aquifer transmissivity. Mukhopadhyay found that ANNs gave a better estimation of transmissivity at well locations, provided training data covering the whole spectrum of variation within the study area was available.

In a study done by Yoshioka et al. (1996), the porosity-thickness of oil reservoirs was estimated using three different approaches. Simple kriging, ANNs, and a combined ANN/co-kriging approach were evaluated. The combined co-kriging/ANN approach presented by Yoshioka et al. was sequential in nature and differed in quite a few ways from the approach proposed in this study. Yoshioka et al. found that the combined co-kriging/ANN approach produced inconsistent results, and that neural networks alone did a relatively poor job predicting porosity values. However, they also found that the simple kriging analysis produced results that were almost as inaccurate as those produced by the ANN, so the appropriateness of their data for any form of spatial interpolation has to be questioned.

Wang et al. (1999) conducted a somewhat similar study. The authors combined neural networks and kriging to model reservoir properties. Their integrated technique first used neural networks to estimate porosity, then optimized the network performance by analyzing the variograms of the residuals at conditioning points. Gaussian simulation was performed and the resulting maps were combined with porosity trends obtained from the ANNs. The authors found the results to be realistic and true to the known properties of the oilfield.

Wang et. al. (1999) performed another analysis where an ANN was used to produce a direction basis function that was then used to create anisotropic semi-variograms that were used in a kriging analysis. This produced results that were very similar to those obtained in the previous portion of their study. They found this technique to be fast and straightforward and eliminated any need for cross-correlation modeling.

PROPOSED EXPERIMENT

Using a yet to be determined model development data set, we propose to construct both standard co-kriging and hybrid co-kriging/ANN interpolation models. These models will be applied to a validation data set and the absolute and relative accuracies of these interpolation procedures will be compared.

Once selected, the data set will record values and identify locations where at least two covariates were measured. This data set will be divided into training and validation subsets. Traditional co-kriging and hybrid co-kriging/ANN interpolation models will then be constructed using the training data subset. The accuracy of both of these interpolation models will be determined by comparing their interpolated values to known values recorded in the validation subset.

Previous studies have shown that artificial neural networks have at least equal, and frequently superior, capabilities than standard regression techniques in recognizing patterns and discovering relationships in noisy data sets. Therefore, it is logical to hypothesize that a hybrid co-kriging/neural network approach should produce more accurate interpolations than traditional co-kriging.

REFERENCES

Bailey, T. C., Gatrell, A. C. 1995. Interactive spatial data analysis. Longman Scientific & Technical, England.

Bishop, C. M. 1995. Neural Networks for Pattern Recognition. Clarendon Publishing, Oxford, England. 482 pages.

Blackard, Jock A. and Denis J Dean. 1999. *Comparative Accuracies of Artificial Neural Networks and Discriminant Analysis in Predicting Forest Cover Types from Cartographic Variables*. **Computers and Electronics in Agriculture**. 24(1999):131 151.

Cressie, N. 1991. Statistics for spatial data. John Wiley and Sons, New York, New York. 900 pages.

Environmental Systems Research Institute, 1998. ArcInfo Version 7.2.1 on-line help documentation. Kriging (TIN). Environmental Systems Research Institute, Redlands, CA.

Gamma Design Software, 2000. http://www.geostatistics.com., http://www.GammaDesign.com. Accessed: July 25, 2000.

Gurney, K. 1997. An introduction to neural networks. UCL Press Limited, London.

Kaluzny, S.P., Vega, S.C., Cardoso, T.P. and Shelly, A.A. 1998. S+ spatial stats: user's manual for windows? and UNIX? . Springer-Verlag, New York, New York. 327 pages.

Matsoukas, C., Islam, S., and Kothari, R. 1999. Fusion of radar and rain gage measurements for an accurate estimation of rainfall. Journal of Geophysical Research 104(D24), 31, 437-31, 450.

Mukhopadhyay, A. 1999. Spatial estimation of transmissivity using artificial neural network. Ground Water 37(3), 458-464.

Ripley, B.D. 1996. Pattern Recognition and Neural Networks. Cambridge University Press, New York, New York. 403 pages.

Trochim, W. 1999. Research methods knowledge base, 2nd ed. http://trochim.human.cornell.edu. Accessed: August 8, 2000.

Wang, L., Wong, P., Kanevski, M., and Gedeon, T.D. 1999. Combining neural networks with kriging for stochastic reservoir modeling. In Situ 23(2), 151-169.

Wang, L., Wong, P., and Shibli, S.A.R. 1999. Modeling porosity distribution in the A'nan oilfield: Use of geological quantification, neural networks, and geostatistics. SPE Reservoir Eval & Eng 2(6), 527-532.

Yoshioka, K., Shimada, N., and Ishii, Y. 1996. Application of neural networks and cokriging for predicting reservoir porosity-thickness. GeoArabia 1(3).