Predicting the Form of a Surface from Random and Nonexperimental Samples Using Neural Networks

James E. Mitchell, Hao Sun
Geohydrology Section
Kansas Geological Survey
and Swapan Chakrabarti
Department of Electrical and Computer Engineering
University of Kansas
Lawrence, Kansas 66047
913-864-3965
JEMKGS@UKVAX.BITNET
JEMKGS@KUHUB.CC.UKANS.EDU
FAX: 913-864-5317

Neural Networks and Applications

Kansas Geological Survey Open-file Report No. 92-5

Predicting the Form of a Surface from Random and Nonexperimental Samples Using Neural Networks

James E. Mitchell*, Hao Sun*
and Swapan Chakrabarti**

*Geohydrology Section
Kansas Geological Survey
**Department of Electrical
and Computer Engineering
The University of Kansas
Lawrence, KS 66047

ABSTRACT

The performance of two types of neural networks are compared with a conventional approach for contouring three-dimensional data. The problem is of particular importance in geohydrology because of the nonexperimental nature of the data collection process. Geoscience applications typically rely on data collected by different investigators. Often the data are sampled using different techniques. Most conventional algorithms for contouring surfaces perform best when the data are regularly sampled from a symmetric grid.

Using data from a Digital Elevation Model (DEM), the predictive performance (based on mean absolute error) of four approaches are compared. The four algorithms include 1. a conventional, linear projection (CLP) method, 2. a first-order multi-layer linear neural network (LNN), 3. a quadratic nonlinear neural network with no hidden layer (QNLNN-0), and 4. a quadratic nonlinear neural network with one hidden layer (QNLNN-1).

All of these networks demonstrate the capability of using sample points to approximate a natural surface with high-order (nonlinear) information. These surfaces are difficult to describe in the form of polynomials. In the LNN, a surface is simulated by using planes in three dimensions, combined as a function of weight vectors. In a QNLNN, quadratic surfaces replace the planes and appear to provide more accurate surface predictions. Results indicate that at small and intermediate sample sizes the neural networks outperform the conventional approach. With large sample sizes, there appears to be no significant difference among the CLP, QNLNN-0, and QNLNN-1, but the LNN experiences a much higher mean absolute error than the other three algorithms. In addition, the QNLNNs experience a consistently lower loss than all the others across all sample sizes.

1. Introduction

Predicting the form of a three-dimensional surface from nonexperimental data is a problem of long-standing interest to hydrologists and other geoscientists [Da86]. Examples of this include the prediction of land surface elevation, groundwater elevation, geologic formation boundaries, and calculation of saturated aquifer thickness. The problem arises from the investigator's inability to control the sampling of data. As a result, irregularly, randomly spaced and typically "clumped" samples provide the only source of data.

Most contouring algorithms rely on fitting some type of linear or nonlinear function to a regular grid of data. If this grid is not available, an additional step is required to transform irregular (in 2-dimensional space) samples to a grid. Typically, this results in some loss of accuracy compared to regular grids [Da86]. In addition, these approaches depend on some parametric function with parameters that must be fitted, usually using a least-squares approach. These functions are not based on any geophysical relationships. Therefore the process amounts to an exercise in surface-fitting.

Neural networks have demonstrated an impressive ability to predict or interpolate. Most previous work focuses on function approximation and time-series prediction where data are defined as two-dimensional and not spatially correlated. These works have shown that a multi-layer LNN provides reliable predictions of nonlinear function mappings [Jo90]. So-called higher order neural networks (HONN) have been intensively studied in recent years in the area of invariant pattern recognition and have been shown to be more effective than LNNs for nonlinear discrimination [Sp90]. The neural network approach also has the advantage of being non-parametric and makes no assumptions about the form of the input data. Most conventional statistical methods assume that the response surface is a specific functional form [Mu90].

Here, two kinds of feed-forward network are employed to interpolate a nonlinear three-dimensional surface using irregular and random control points. One network is designed to have two hidden layers and first-order neurons. After training, the weight vectors in this network form a combination of planes in three dimensions that simulates different surfaces based on input samples. The second type of network has quadratic neurons in the input layer forming quadratic surfaces. Theoretically, the latter approach is closer to a natural surface because it contains higher-order information. Two forms of this network were tested and compared, and the results support this assumption. All the networks are based on a back-propagation (BP) learning algorithm.

This experiment represents the initial step in exploring neural network applications to surface contouring. The results demonstrate some advantages that these methods may have over conventional approaches.

2. Methodologies and Experimental Design

The data for this investigation were taken from USGS Digital Elevation Models (DEM) for the Kansas City area Wolcott quadrangle (1:24,000 scale) in northeastern Kansas. The data represent a grid of points from the contours on this map. From a 512 × 512 grid, subsamples of 15 × 15 were randomly selected for training sets. Each training set was randomly sampled for a variable number of points to begin the training of the LNN. Thirty training sets were selected. The DEM data were also used as control points to compare the results from different algorithms.

The LNN is a two-dimensional two-hidden-layer feed-forward structure based on a BP learning algorithm (see Fig. 1). The size of input, hidden, and output layers are 15 × 15, 7 × 7, 5 × 5, and 15 × 15 respectively. Input and output layers represent the geographic locations, and the value at each point is the elevation of the surface. This network takes the first-order information as input, and the resulting output is the surface. This global surface is divided into several regional subplanes, and each of these planes is described by a local weight vector. The prototypic sample points are partitioned into groups based on their characteristics and similarity; then each group forms a weight vector. The weight space is therefore divided into grid samples.

When the input neurons are replaced by quadratic models, the network (see Fig. 2) is called a quadratic nonlinear network (QNLNN). The output of the network has the form:

$$z = w_0 + w_1x + w_2y + w_3x^2 + w_4xy + w_5y^2 \quad (1)$$

where x and y represent the geographic location and z is the elevation. Similarly, Fig. 3 shows a QNLNN with one hidden layer (QNLNN-1).

These networks approximate the surface by using several local quadratic surfaces that formed by combining weight vectors and provide more accurate results than using planes. There is no training process in QNLNN; every sample set is fed into the network to generate a surface after the network converges. Because the structure is simple, this process is fast.

The networks were compared to a conventional linear projection algorithm CLP. This is a two-part procedure in which a weighted average of slopes projected from nearest neighbor data points around each grid node are used to estimate the value at the node. The nearest n neighboring observations around a grid node are found and each is weighted inversely to its distance from the point. A trend surface is then fitted to these weighted observations [Sa88].

3. Results

Figure 4 compares the four candidate approaches for sample sizes of 10, 20, 38, 42, 52, 61, 69, 100, and 169. Prediction errors were calculated from the same random 15 × 15 sample for each algorithm. Each was sampled repeatedly, increasing the sample size, as described earlier.

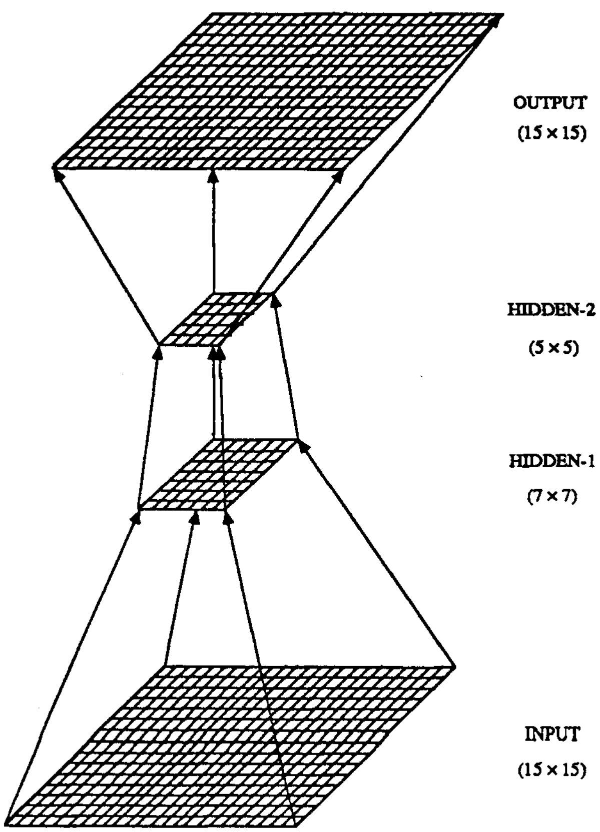

Figure 1. Two-dimensional structure for multilayer feed-forward neural network. Each layer is fully connected to the upper layer.

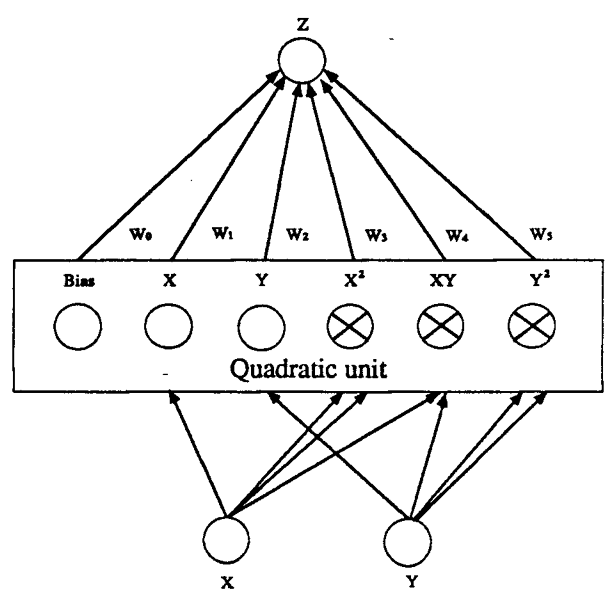

Figure 2. Organization for quadratic nonlinear neural network with no hidden layer. The quadratic unit converts input X and Y into a second order polynomial form.

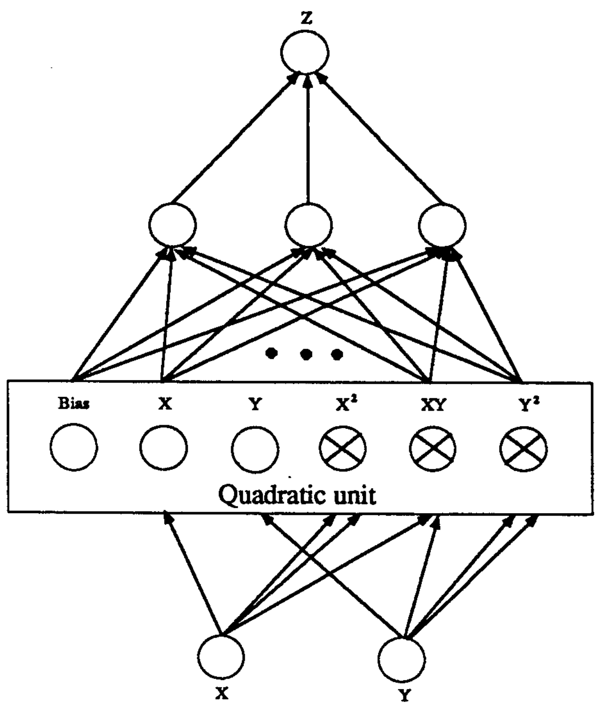

Figure 3. Organization for quadratic nonlinear neural network with one hidden layer which consists of three neurons.

The performance of the neural networks suggests that small sample sizes fail to provide sufficient information to the LNN and CLP. The QNLNN approaches have similar mean absolute error and are consistent across the various sample sizes for predictive purposes. In additional, large sample sizes apparently over-train the LNN, resulting in erroneous predictions.

4. Conclusions

The training sets for this network contain different random sample sizes and 15 × 15 grids. The ability of this network to generalize and the size of the training sets needed to attain an adequate level of performance are not known in general The results suggest that too many training sets or sample points may overfit the surface. This appears to inhibit the LNN network's ability to generalize.

In this research, higher-order neural networks have been shown to be more consistent and well suited for nonlinear prediction. Networks of an order higher than quadratic (such as cubic networks) may approximate and generalize better, because natural surfaces appear to be much more complicated than low order polynomial surfaces. This has yet to be tested.

The global surfaces generated by these networks are considered an average global estimate of local expected features. Some local detailed information is certainly missed or ignored if there is not enough information available in an area. One way to solve this problem may be to break the whole region into subareas based on some surface characteristics. This idea is still under development.

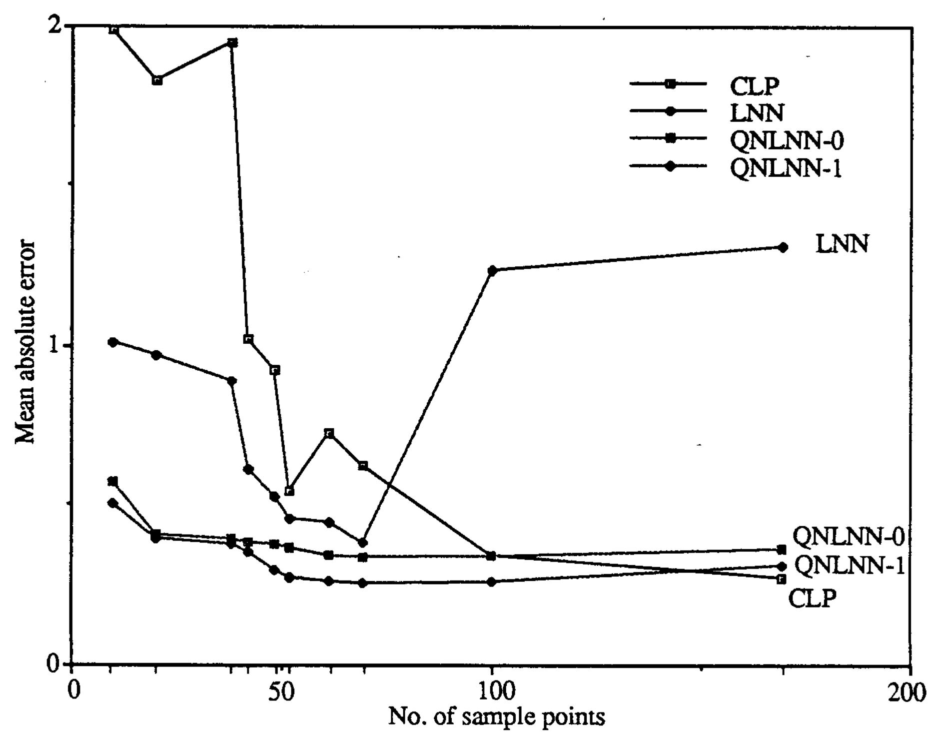

Figure 4. The comparison of mean absolute error for sample size of 10, 20, 38, 42, 52, 61, 69, 100 and 169.

REFERENCES

[Mu90] B. Muller and J. Reinhardt, Neural Network: An Introduction. Springer-Verlag, Berlin, Heidelberg, 1990.

[He91] John Hertz, Anders Krogh, and Richard G. Palmer, Introduction to the Theory of Neural Computation. Addison-Wesley, Redwood City, California, 1991.

[Da86] John C. Davis, Statistics and Data Analysis in Geology. John Wiley & Sons, 1986

[Sa88] Robert J. Sampson, Surface III. Interactive Concepts Incorporated, Lawrence, Kansas, 1988.

[Vo90] Dennis J. Volper and Steven E. Hampson, "Quadratic Function Nodes: Use, Structure and Training," *Neural Networks*, Vol. 3, pp. 93-107, 1990.

[Ha90] Steven E. Hampson and Dennis J. Volper, "Representing and Learning Boolean Functions of Multivalued Features," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 20, No. 1, pp. 67-79, 1990.

[Sp90] Lilly Spirkovska and Max B. Reid, " Connectivity Strategies for Higher-order Neural Networks Applied to Pattern Recognition," *Proceedings of IEEE International Joint Conference on Neural Networks*, Vol. 1, pp. I-21 - I-26, 1990.

[Jo90] R. D. Jones and Y. C. Lee, "Function Approximation and Time Series Prediction with Neural Networks," *Proceedings of IEEE International Joint Conference on Neural Networks*, Vol. 1 pp. 649-959, 1990.