IN SITU, 23(2), 151-169 (1999)

COMBINING NEURAL NETWORKS WITH KRIGING FOR STOCHASTIC RESERVOIR MODELING

L. Wang and P.M. Wong
School of Petroleum Engineering
University of New South Wales
Sydney 2052, Australia

M. Kanevski
Institute of Nuclear Safety
B.Tulskaya 52, Moscow 113191, Russia

T.D. Gedeon
School of Computer Science and Engineering
University of New South Wales
Sydney 2052, Australia

ABSTRACT

Stochastic reservoir modeling is being increasingly used for modeling reservoir heterogeneity. This paper describes a methodology to model the distribution of reservoir properties using well data and soft geological knowledge in the form of sedimentary and diagenetic patterns. The technique, developed based on a combined use of radial basis function (RBF) neural networks and geostatistical kriging, is demonstrated with an application to interpolating porosity in the A'nan Oilfield, located onshore north China. The integrated technique first uses neural networks to estimate the porosity trends from high-dimensional geological patterns. Optimization of the network performance is done by variogram analysis of the residuals at the conditioning points. Gaussian simulation of the residuals is then performed, and the resulting residual maps are combined with the porosity trends obtained from neural networks. From the case study, the results are realistic and honor the geological rules of the oilfield. The technique is

fast and straightforward, and provides a computational framework for conditional simulation.

INTRODUCTION

Spatial descriptions of reservoir properties such as porosity and permeability are key components for performance evaluation and field development planning. When wells are sparse and limited, the statistics of the well data become unrepresentative, and this poses a great challenge to reservoir model building. The recent trend in reservoir characterization is to incorporate soft information such as extensive seismic attributes for improved descriptions of spatial continuity of extremes. However, many natural phenomena are far too complex, and information such as two-point statistics extracted from conventional reservoir data (well and seismic) may not be realistic enough to model such phenomena.

Object-based simulation models[1] such as the Markov random field technique present a promising concept to incorporate stochastic geological objects into reservoir models. The approach, however, requires many model parameters, which are difficult to interpret or infer, and can only handle relatively simple objects such as fluvial channels. This explains why many reservoir scientists are still reluctant to investigate stochastic algorithms for reservoir modeling. It is because they are able to produce realistic results by using conceptual geological models[2], such as hand drawings by expert geologists.

Very often, hand drawings from expert geologists contain valuable information about the spatial continuity of reservoir properties. The drawings or "templates" incorporate important soft geological knowledge (rules) and present a comprehensive summary of complex structural, sedimentary and diagenetic patterns which are beyond the present tools of two-point statistics (histograms and covariances). Wang *et al.*[3], Wong *et al.*[4] and Caers and Journel[5] show that the neural network approach is a promising tool to handle such information for stochastic reservoir modeling, and it may be a viable alternative to object-based algorithms.

The objective of this paper is to introduce an integrated use of neural networks, particularly radial basis function neural networks (RBFNN), and kriging[6] to simulate porosity from well data and geological templates. The next section will first present a review of RBFNNs, followed by a detailed description of model setup and parameter specification. The technique will be demonstrated via a case study in the A'nan Oilfield, located onshore north China.

RADIAL BASIS FUNCTION NEURAL NETWORKS

A radial basis function (RBF)[7] is a symmetrical transfer function, such as the Gaussian function. In spatial interpolation, it transforms the Euclidean distance between two multidimensional vectors into a function value. Like most neural network methods, radial basis function neural networks (RBFNNs) also attempt to mimic simple biological learning processes. The biological representation is not apparent, but dendritic trees in the human brain are known to simulate a Gaussian-type response when neurons in the retina of the eye produce center-weighted responses to small receptive fields. A RBFNN can learn from a given set of input-output patterns and is a universal function approximator[7]. Its application to reservoir characterization includes log evaluation, such as permeability prediction from multiple well logs[8], and reservoir mapping, such as porosity interpolation[4,5] and reservoir top delineation[9,10]. The technique is robust and can be applied to model non-linear and non-stationary events in a multivariate environment.

The Estimator

A typical RBFNN contains three layers of processing elements or neurons: input, hidden and output layers. Each neuron is connected to every neuron in the preceding layer by a simple weighted link. The number of input neurons and output neurons depend on the application domain. If we use X-Y coordinates to infer porosity, the network will be represented by two input neurons (X,Y) and one output neuron (porosity). Fig. 1 shows the schematic diagram of a RBFNN.

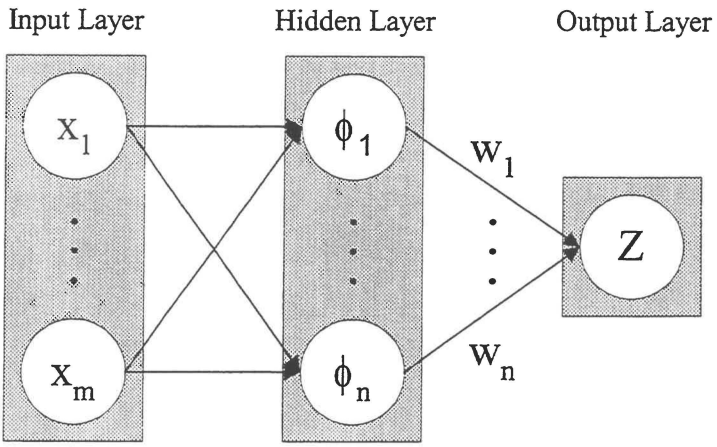


FIG.1. A schematic diagram of a RBNN.

For the sake of mathematical simplicity, we will limit our discussion to only single output networks.

Each hidden neuron represents a RBF center which is parameterized by a position or reference vector $\mathbf{c}$ located in the m-dimensional covariate input space $\mathbf{x}$. The estimator $z^*(\mathbf{x})$ is a weighted sum of the basis function values from the hidden layer:

$$z^*(\mathbf{x}) = \sum_{j=1}^{n} w_j \phi_j\left(\left\|\mathbf{x} - \mathbf{c}_j\right\|\right) \tag{1}$$

where $\{w_j\}_1^n$ are the weights, n is the number of RBF centers, $\|.\|$ denotes the Euclidean norm, and $\phi(.)$ represents a RBF.

The most popular RBF is the Gaussian function $\phi(x) = \exp(-x^2/2\sigma)$, where σ is a constant controlling the radius of influence of the basis function. A smaller σ would give more weight to the samples that are closer to the prediction location. In contrast, a larger σ would distribute weightings more evenly across

all the samples. Other examples of RBFs are multiquadratic and inverse-multiquadratic[7]. For demonstration purpose, we will use the Gaussian RBF in our study.

From Equation (1), it is obvious that RBFNNs are posed as a general linear least-squares problem: the regression of a target variable $z(\mathbf{x}^0)$ on an input set of covariates $\mathbf{x}^0$ given the training data pairings $\{(\mathbf{x}_1^0, z_1), (\mathbf{x}_2^0, z_2), \ldots, (\mathbf{x}_N^0, z_N)\}$. The unknowns are the weights $\{w_j\}_1^n$ and can be obtained by setting

$$\frac{\partial\left(\sum_{k=1}^{N}\left(z_k(\mathbf{x}^0) - z_k^*(\mathbf{x}^0)\right)^2\right)}{\partial w_j} = 0 \qquad (2)$$

Radial Basis Centers

The most critical decision for setting up a RBFNN is the number of basis centers in the hidden layer. If all the training inputs $\mathbf{x}^0$ are employed as the basis centers $\mathbf{c}$ ($n = N$), the weights can be obtained analytically by solving a system of linear equations and the estimator becomes an exact interpolator[4], or $z(\mathbf{x}^0) = z^*(\mathbf{x}^0)$. The results would be deterministic. If multiple realizations are desirable, we may need to obtain some measure of prediction confidence such as estimation variance.

If, however, only a portion of the training inputs are retained as the basis centers ($n < N$), numerical methods such as gradient descent are required to iteratively minimize the errors (residuals) between the model outputs and target values as shown in Equation (2). This process is commonly known as "learning" in neural network jargon. In fact, the number of centers controls the degree of freedom in the estimator, and hence the degree of determinism in the output. The residuals may be sufficiently small, but zero residuals are practically impossible. Hence, this estimator is not an exact interpolator. Because of this inexactitude property, this creates an opportunity to perform stochastic simulation on the residuals (see later sections). This paper emphasizes the use of the inexact estimator, and the discussion hereafter will be directed to this estimator.

The decision on what reference vectors $\mathbf{c}$ are to be used is also crucial to the network performance. These vectors can be selected by prior knowledge, random picking or can be generated systematically. Many methods are available for choosing appropriate basis centers. The most intuitive rule is to locate points carrying important information about the spatial distribution trend of the data set. For example, the centroids from cluster analysis operated on the training inputs can be used as the centers for RBFNN learning.

In this paper, we chose the basis centers using a knowledge-driven approach, which was based on a previously geological study providing regional information on the structural, sedimentary and diagenetic facies zones[11]. In each of these zones, we then chose a location indicating a major geological environment. This information contains useful geological knowledge which cannot be captured by most numerical methods (e.g., cluster analysis).

RBF Constant

The RBF constant σ is another information parameter in RBFNN learning. Each center may use its own constant. The same constant can also be used for all centers. The latter case is often preferred for practical use of RBFNN, because it has only one parameter to be optimized. In this paper, we determine the optimum constant by examining the RBFNN learning behavior for different constants; the lower the total error, the better the constant.

The use of only one RBF constant in the model essentially assumes an isotropic model, which may be too simplistic in nature. However, the incorporation of more elements (reflecting the true geological anisotropy) in the input vector can significantly relax the isotropic assumption[4]. This paper uses a set of geological hand drawings to provide such anisotropic information.

RBFNN Learning

In most error minimization algorithms, a stopping criterion is required to terminate the iterative loop. In neural network, a popular technique known as

"early-stopping" is used to terminate iteration. It is done by using a validation data set with patterns not in the training set. The validation set is fed into the network, and the network stops at the minimum error on the validation set rather than the training set. This avoids the over-fitting of the data and gives better generalization[7]. The trained network can then be used for prediction.

In reservoir modeling, however, we often do not have enough data points to be used for the validation purpose. We therefore propose the use of the correlation range of the variogram to define the optimum point for early-stopping. We first calculate the variograms of the residuals at certain intervals of iterations, followed by fitting a model to the data. The correlation range is then defined at various iterations. In spatial analysis, the larger the range, the stronger the correlation between the residuals, and vice versa. This means that, if the network learns well, the range would be close to zero (i.e., no spatial correlation between the residuals). This is equivalent to claiming that all the essential features of the data set have been extracted by neural learning. What is left would be white noise. Thus, the optimum point is defined as the iteration number where the range reduces no further. This is essentially the same as early-stopping, but we use a spatial performance criterion rather than simple error measures such as the use of error squares, average error and error skewness[3].

Combining Neural Networks with Kriging

Most neural network estimators are inexact interpolators. In Kanevski *et al.*[12], the authors proposed a combined use of a multilayer perceptron[7] (an inexact estimator) and kriging to interpolate radionuclides for an environmental problem. This is the so-called neural network residual kriging (NNRK), which is an exact interpolator. The basic idea is to treat each estimate $z^*_{NNRK}(\mathbf{x})$ as the sum of two components:

$$z^*_{NNRK}(\mathbf{x}) = z^*_{NN}(\mathbf{x}) + \varepsilon(\mathbf{x}) \tag{3}$$

where $z^*_{NN}(\mathbf{x})$ is the neural network estimate indicating the non-linear component (large-scale non-stationary features), and $\varepsilon(\mathbf{x})$ represents the stochastic error component (small-scale stationary features), which is practically impossible to be modeled by a neural network. Based on the implication from Equation (3), the authors treated the residuals as $\varepsilon(\mathbf{x})$ and generated a $\varepsilon(\mathbf{x})$-map using kriging. The residuals were subsequently added to the neural network estimates, and hence $z(\mathbf{x}^0) = z^*_{NNRK}(\mathbf{x}^0)$. In Kanevski *et al.*[13,14], the authors also extended the methodology and modeled the residuals by cokriging and simulated annealing respectively.

In Wang *et al.*[3], the authors used a similar approach with a different neural network paradigm, the inexact RBFNN estimator or $z^*_{NN}(\mathbf{x}) = z^*(\mathbf{x})$, and extended the methodology to construct stochastic residual maps (and hence equally probable property maps) via sequential Gaussian simulation in a multivariate environment. This is referred to as the neural network residual simulation (NNRS). NNRS inherits all the advantages of NNRK with the capability to perform stochastic simulation. The integrated paradigm also allows the conversion of a complex multivariate problem to a simple univariate one, because we need to model only the variogram of the porosity residuals after neural learning. This offers a significant advantage over conventional cokriging, which requires tedious modeling of cross-covariances when many types of secondary (soft) data are used.

CASE STUDY

Oilfield Characteristics

The A'nan Oilfield is located in a compound basin formed in the Mesozoic and Cenozoic Era in north China. The average porosity ranges from 7% to 14% in the cored wells, and there are complex geological controlling factors on the porosity distribution in its oil-bearing stratum. In this paper, we used the above methodology to model the 2D porosity distribution in this oilfield. We set up

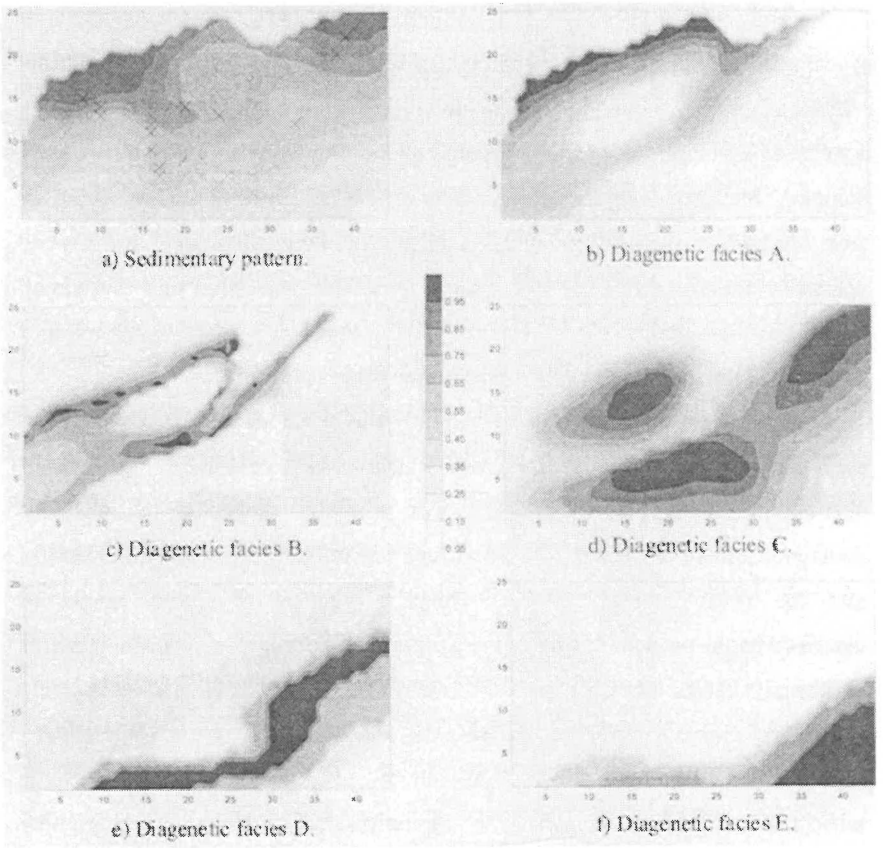


FIG. 2. The geological templates showing the distribution of sedimentary and diagenetic facies. a) Grayscale represents the relative thickness of the coarse sediments (dark: thick; light: thin). The north border of map a) is the location of Arshan fault. Conditioning points are shown as "x". b)-f) Grayscale represents the certainty of each diagenetic facies distribution (dark: certain; light: uncertain). Each unit in the axis represents 300 m.

1,100 (44 by 25) gridblocks around the case area. Fig. 2 shows the geological templates of the field, which are drawn by expert geologists according to their geological knowledge. Fig. 2a also shows the locations of the conditioning points, which are denoted by "x" (41 in total). Average porosity values are available at these points.

Along the north border of the oilfield, there is a large-scale tensional and syndepositional normal fault stretching approximately from east to west, and the oilfield is located in the southern dropping side of the fault (Fig. 2a). With the activity of this fault, a reverse dragging anticlinorium was formed in the south dropping side, and several secondary faults were also formed in the anticlinorium area. Meanwhile, the syndepositional subaqueous intermediate and basic fissure volcanic eruptions occurred along the large-scale fault. The sedimentary system of the oil-bearing stratum is fan-delta-lacustrine system. Fig. 2a shows the relative thickness of the coarse sediments in the sedimentary sequence.

The diagenetic events that occurred in the burial diagenetic process were very complex. Five different diagenetic facies were identified[11]. The spatial distribution of the diagenetic facies was primarily controlled by the flow directions of the compacted pore water, and transitional facies do exist laterally after the burial diagenetic process. Hence, the facies distribution cannot be described by the probability concept (e.g., facies proportions) as typically done for sedimentary facies. Therefore, we used the possibility concepts[15] (related closely to the fuzzy logic concepts) to derive a set of maps (Figs. 2b-f) expressing the degree of membership belonging to each of the diagenetic facies across the oilfield. We simply call the values the "certainty values." The higher the certainty value of a diagenetic facies, the more likely the point will behave as that facies. Note that the major difference between probability and possibility concepts is that, in possibility theory, the sum of all the certainty values at a given location does not necessary equal to one.

In this oilfield, the distribution of porosity is much more complex than in normal sedimentary petroleum reservoirs because of the complexities of the structure, ancient volcanic eruption, special sediments and burial diagenesis. The main oil-bearing space is the secondary dissolved pores, so the distribution of secondary dissolution facies (the diagenetic facies "C" shown in Fig. 2d) was very important for porosity distribution in this area. From geological study, it is clear that the porosity of the reservoir should increase with the certainty value of the diagenetic facies "C." Other facies also relate to the porosity distribution. For

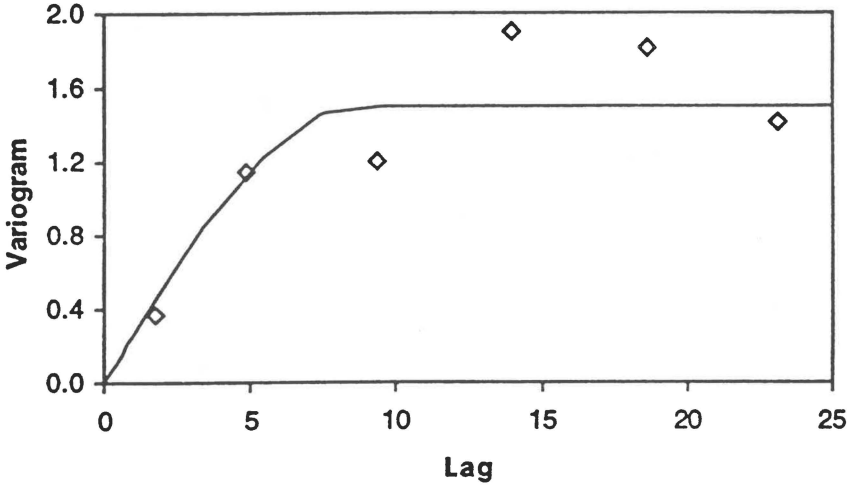


FIG. 3. Porosity variogram fitted with a spherical model.

example, the porosity should be small when the diagenetic facies "E" (Fig. 2f) is likely to exist. Therefore, it is important to incorporate such information into the modeling algorithm so that the geological framework could be honored.

Porosity Variogram and Ordinary Kriging

Fig. 3 shows a variogram of porosity data at the conditioning points. A spherical model was fitted to the variograms. The range and sill values were about 8.5 and 1.5, respectively. Ordinary kriging was performed using the fitted variogram model.

Fig. 4 shows the resulting porosity map. The results were not satisfactory. This conclusion was drawn purely based on qualitative comparison with the conceptual geological model developed in our previous studies[11]. For example, there was a large amount of syndepositional volcaniclastics along the large-scale normal fault, and the thick volcaniclastics were harmful to the formation of the oil-bearing zones. Hence, the porosity along the fault should be very low and

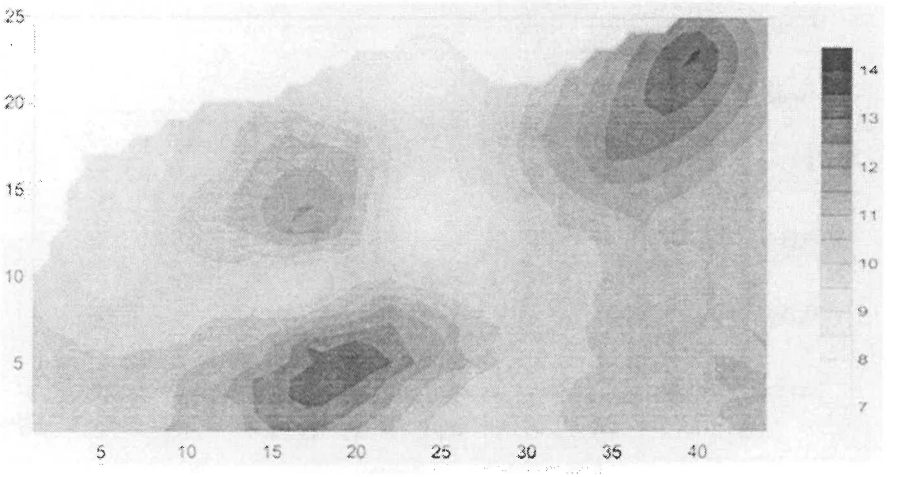


FIG. 4. Porosity map from ordinary kriging.

reduced quickly towards the north. From Fig. 4, we can see that the porosity values are still very high near the fault, and hence the results were considered unacceptable.

In earth sciences, the reservoir system is always open, and hence no objective truth is available for the verification of any reservoir models[16]. History matching is theoretically unsound to be used as means to verify the reservoir models[16]. For example, if two totally different reservoir models could produce the same outputs (including resulting statistical measures), there is no way to evaluate the reliability of the models.

Our comparison approach (currently qualitative) strongly emphasizes whether the results are geologically interpretable[2]. We recommend that the most effective and reliable way to improve the model is to directly incorporate reservoir geology (e.g., geological rules) into the modeling algorithm, such as the use of RBFNN as proposed in this paper. However, we should not ignore the fact that this approach may support incorrect geological intuitions.

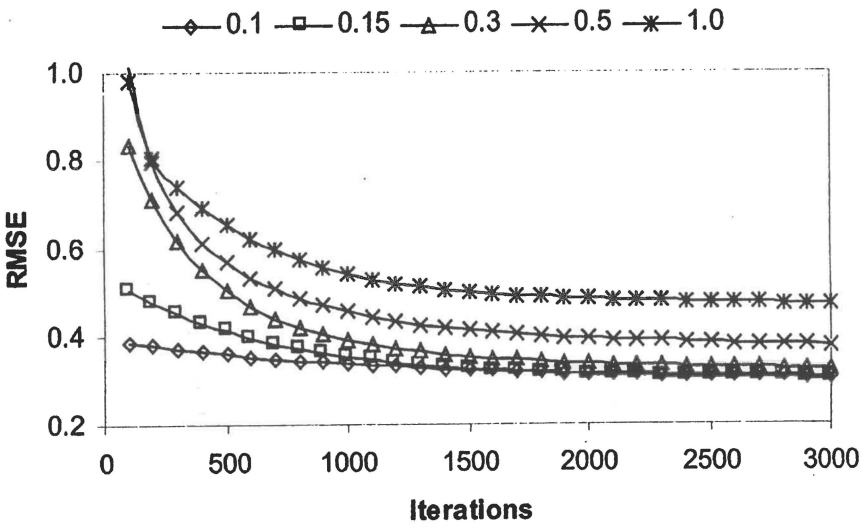


FIG. 5. Learning curves for different RBF constants.

RBFNN Setup

According to the distribution pattern of structural, sedimentary and diagenetic facies, we first identified nine important zones with distinct reservoir environments in this oilfield, and hence nine reference vectors were used as the radial basis centers of the Gaussian type. The vector has eight dimensions: x-coordinate, y-coordinate, the sedimentary factor (Fig. 2a) and the five diagenetic facies factors (Figs. 2b-f). All the elements were normalized into the range (0,1).

In this case study, 41 porosity values across the field were used as the conditioning points for training the network. We used a standard gradient descent algorithm to minimize the residuals. Different basis constants were used. The learning curves (root-mean-square error, RMSE, versus iterations) are shown in Fig. 5. The lowest error was obtained at $\sigma = 0.15$ after 3,000 iterations. This became the optimum constant for the study.

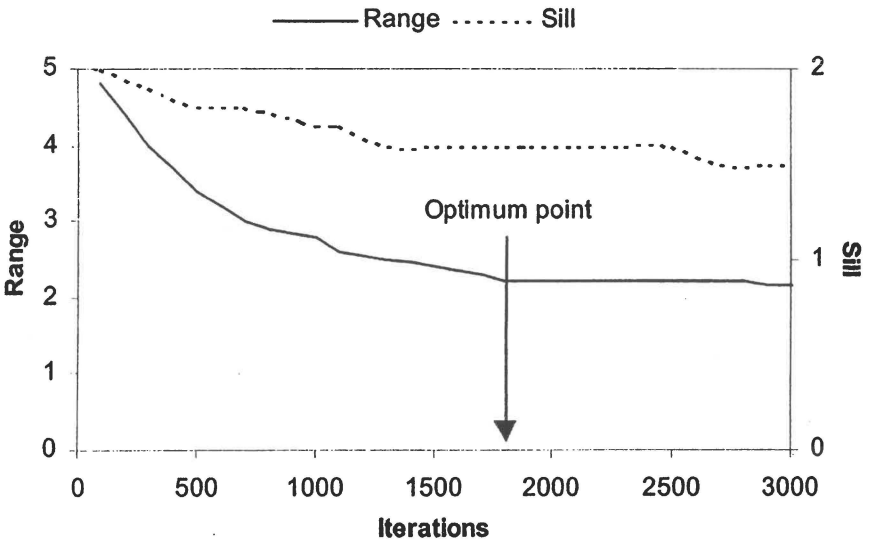


FIG. 6. Variogram parameters to define the optimum termination point.

Residual Variogram Analysis

The objectives of the variogram analysis are two-fold: 1) determine the optimum point to terminate iteration and 2) perform residual kriging and simulation. In order to locate the optimum point, we first trained the RBFNN with $\sigma = 0.15$. This was followed by calculating the variograms of the residuals at every 100 iterations. The data were fitted with a spherical model.

Fig. 6 shows the sill and correlation range obtained from the model as a function of iterations. The sill reduced with iterations because the estimates were closer to the actual values and hence the residual variances reduced. The range reduced from about 5 to 2 (60% reduction) with iterations. This indicated that the correlation of the residuals became weak (smaller range) as the amount of features extracted from the RBFNN increased. The optimum termination point was found to be about 1,800 iterations as the reduction of the correlation range became insignificant, that is, prolonged training does not lead to significant reduction of the parameter. Therefore, training was terminated at 1,800 iterations.

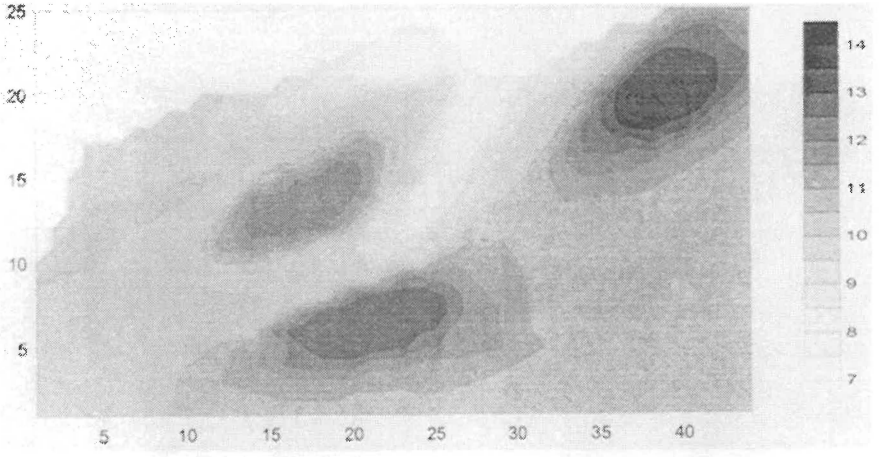


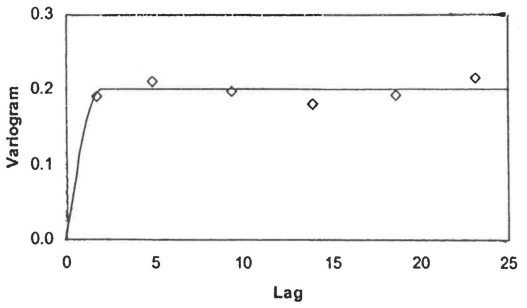
FIG. 7. Porosity map from RBFNN.

Fig. 7 shows the resulting map from RBFNN after 1,800 iterations. The results reflected the geological trends as shown in Fig. 2 because RBFNN incorporated the soft geological templates. Hence the results were considered superior to ordinary kriging (Fig. 4). The interpolations, however, were not exact and only one deterministic image was produced.

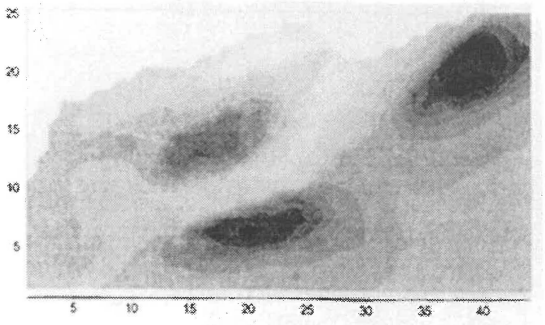
Residual Kriging and Simulation

To restore the conditioning data, residual kriging was performed using the residual variogram obtained at 1,800 iterations (Fig. 8a). The correlation range of the residual variogram was about 2. Most spatial features were learned by RBFNN and only a small amount of information was left unlearned.

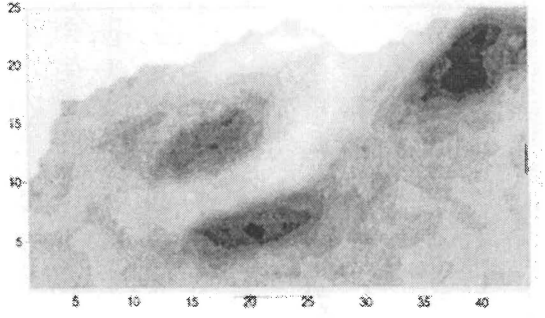
After residual kriging, the interpolated residuals were added to the RBFNN estimates (Fig. 7). The NNRK results are shown in Fig. 8b. The same variogram was also used to simulate the residuals using sequential Gaussian simulation. The residuals were also added to the RBFNN estimates. Two NNRS realizations were generated and are shown in Figs. 8c-d. All the porosity values



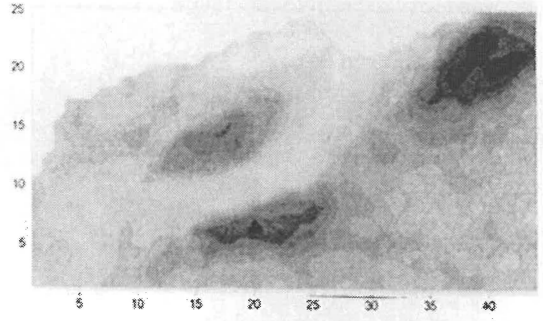
a) Residual variogram at 1,800 iterations.



b) NNRK estimates.



c) NNRS estimates: 1st realization.



d) NNRS estimates: 2nd realization.

FIG. 8. NNRK and NNRS results using the residual variogram at 1,800 iterations.

were restored at the conditioning points, and the final results also inherited the geological trends in the reservoir.

CONCLUSIONS

This paper introduces a combined used of radial basis function neural networks (RBFNNs) and geostatistics for stochastic reservoir modeling. RBFNNs are first used to estimate the large-scale variations (major trends), followed by estimating/simulating the small-scale variations (residuals) from kriging and sequential Gaussian simulation. Variogram analysis also provides valuable information, particularly the correlation range, to evaluate the performance of neural networks.

The integrated technique is demonstrated in the A'nan Oilfield in north China, where porosity is interpolated using the sedimentary and diagenetic patterns. According to the geological knowledge available, the results from the integrated technique are superior to ordinary kriging alone. It is primarily due to the fact that the integrated technique is able to incorporate the multivariate geological templates (hand drawings from expert geologists) into the model effectively, and hence the results inherit the expert knowledge about the spatial continuity of the reservoir properties obtained from conventional geological study. The technique is also fast and straightforward and does not require any tedious modeling of cross-covariances.

Future work will investigate advanced techniques to further optimize the network performance and extend the methodology to 3D modeling.

ACKNOWLEDGMENTS

The authors would like to thank the Australia Petroleum Cooperative Research Centre (APCRC) for providing financial support for this work. The comments from the two anonymous reviewers are also acknowledged.

REFERENCES

1. Holden, L., Hauge, R., Skare, Ø., and Skorstad, A.: "Modeling of Fluvial Reservoirs with Object Models," Mathematical Geology 30, 473-496 (1998).

2. Tamhane, D., Wang, L. and Wong, P.M.: "The Role of Geology in Stochastic Reservoir Modeling: The Future Trends," SPE 54307, SPE Asia Pacific Oil and Gas Conference and Exhibition, Jakarta, Apr. 20-22, in press (1999).

3. Wang, L., Wong, P.M. and Shibli, S.A.R.: "Modelling Porosity Distribution in the Anan Oilfield: Use of Geological Quantification, Neural Networks and Geostatistics," SPE 48884, presented at the 1998 SPE International Oil & Gas Conference and Exhibition, Beijing, China, 2-6 November, vol. 1, 509-515 (1998).

4. Wong, P.M., Tamhane, D. and Wang, L.: "A Neural Network Approach to Knowledge-Based Well Interpolation: A Case Study of a Fluvial Sandstone Reservoir," Journal of Petroleum Geology 20, 363-372 (1997).

5. Caers, J. and Journel, A. G.: "Stochastic Reservoir Simulation Using Neural Networks Trained on Outcrop Data," SPE 49026, presented at the 1998 SPE Annual Technical Conference and Exhibition, New Orleans, USA, 27-30 September, 321-336 (1998).

6. Deutsch, C.V. and Journel, A.G.: GSLIB: Geostatistical Software Library and User's Guide, Oxford University Press (1992).

7. Bishop, C. M.: Neural Networks for Pattern Recognition, Oxford University Press (1995).

8. Wong, P.M. and Shibli, S.A.R.: "Use of Interpolation Neural Networks for Permeability Estimation from Well Logs," Log Analyst 39, no. 6, 18-26 (1998).

9. Wong, P.M., Shibli, S.A.R. and Maignan, M.: "A Neural Network Approach for Spatial Mapping," Proceedings of GIS AM/FM ASIA & Geoinformatics, vol. 2, 649-656 (1997).

10. Shibli, S.A.R., Wong, P.M., Huang, Y. and Maignan, M.: "Applications of AI-based Techniques for Spatial Prediction," Proceedings of IAMG, part 2, 675-680 (1997).

11. Wang, L., Zhao, C., Liu, M., Jiang, Z.: "A study on the Structural Sedimentary Diagenetic Zone in an oilfield," Journal of Petroleum Exploration and Development (in Chinese) 19, 26-31 (1992).

12. Kanevski, M., Arutyunyan, R., Bolshov, L., Demyanov, V. and Maignan, M.: "Artificial Neural Networks and Spatial Estimations of Chernobyl Fallout," Geoinformatics 7, 5-11 (1995).

13. Kanevski, M., Demyanov, V. and Maignan, M.: "Mapping of Soil Contamination by using Artificial Neural Networks and Multivariate Geostatistics," International Conference on Artificial Neural Networks, LNCS 1327, 1125-1130 (1997).

14. Kanevski, M., Demyanov, V. and Maignan, M.: "Spatial Estimations and Simulations of Environmental Data by using Geostatistics and Artificial Neural Networks," Proceedings of IAMG, Part 2, 533-538 (1998).

15. Fang, J.H.: "Fuzzy Logic & Geology," Geotimes, October, 23-26 (1997).

16. Oreskes, N., Shrader-Frechette, K. and Belitz, K.: "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences," Science 263, 641-646 (1994).