Engineering Journal of the University of Qatar, Vol. 12, 1999, PP. 125-149

A COMPARISON BETWEEN ARTIFICIAL NEURAL NETWORK AND A GEOSTATISTICAL TECHNIQUE IN THE ESTIMATION OF REGIONALIZED VARIABLES

E. E. Tawo* and Saleh M. Al-Alawi**
*Department of Petroleum & Mining Engineering
**Department of Electrical & Electronic Engineering
Sultan Qaboos University, Al-Khod 123
Sultanate of Oman

ABSTRACT

In all mining estimation techniques data play a very important role. Insufficient data would mean the quality of the estimate is unreliable. Due to the importance attached to a good and representative data set, the application of **Geostatistics** to the mining industry puts data before any model. Obtaining a good data set which is usually from drillhole samples is an expensive experiment besides the economic constraint usually placed on the amount of data which can be collected particularly at an early stage of a mineral development.

The advent of a **novel** technique in **'Artificial Neural Network'** (ANN) and its application to the minerals industry is compared with a tested geostatistical technique. Both techniques are presented and tested on a Bauxite deposit. The effectiveness of ANN as a cost saving technique is appraised.

KEY WORDS:, **Geostatistics, Kriging, Artificial neural network, Variograms**

INTRODUCTION

The need for appropriate techniques to estimate regionalized variables in the minerals industry cannot be over emphasised. Developing and using an appropriate technique will enhance confidence levels about a deposit but some doubt will always exist until it has been fully exploited.

Once a mineralization has been found, there is often the need to delineate its extent as well as mapping the attribute $z(x)$ distributed over a given area $\{x \in A\}$. The value $z^*(x)$ may be estimated from samples obtained by preliminary drilling and possibly, trenching, amongst other techniques, depending on the type of

mineralization. While a drilling program could be designed for a full appraisal of the mineralization, there is always scepticism and apprehension as to the intensity of the program, yet this work leads to the eventual decision about whether to proceed or not with the development of the ore body. The scenario described above is an expensive experiment in view of drilling cost, but carries with it a lot of expectatioins. This cost is even further increased if the diamond drill hole (DDH) diameter has to be increased to improve the quality of the information so that the uncertainties about the orebody are reduced. The confidence interval associated with increased hole diameter (volume Support) is higher than that for small diameter holes (point support). (Tercan & Dowd)

It is the intention of this study to address this concern arising from the cost involved, by exploring the use of alternative approaches (Al-Alawi, S. & Tawo, E.E) to reducing the drilling volume, especially in the early stages of exploration. It is, however, emphasised that, it is not an attempt to replace the commonly used type of information, which is exact data, $z(x_\alpha)$, $\alpha = 1,\ldots,n$, whose estimate $z^*(x)$ is usually some linear function of these input data:

$$z^*(x) = \phi(z(x_\alpha)), \alpha = 1,\ldots,n$$

although other categories of data (soft data) have been examined (Tawo, E. E.; Journel, A.: Alabert, F). The geostatistical technique used in this study is familiar, and well tested in mineral evaluation when compared with artificial neural network. There are, however, more sophisticated geostatistical techniques available.

On the other hand, in recent years, considerable attention has been given to the ongoing development of forecasting models based on artificial intelligence techniques such as artificial neural networks (ANN) and expert systems. These techniques have been successfully applied to a wide range of engineering applications by many authors reporting higher accuracy compared to classical methods (Park et al., 1991; Kalam et al., 1995; Al-Alawi et al., 1995).

Artificial Neural Networks (ANNs) are massive parallel information architectures composed of many simple processing elements interconnected to achieve certain collective computational structures. ANNs possess features which are particularly attractive for data analysis. These features include graceful degradation, robust recall with fragmented and noisy data, speed inherent to parallel distributed architectures, generalisation, and the capability of modelling non-linear systems. Due to these features, ANNs has generated increasing acceptance for their

application in a wide range of fields. In engineering, this new technology has demonstrated a remarkable success as a result of numerous applications in electrical, civil, mechanical, petroleum and other engineering disciplines. Artificial neural networks, however, are relatively new to the field of Mining Engineering, and have been sparsely demonstrated in this area. This novel approach, is believed, to provide great benefits if applied in this discipline.

To the knowledge of the authors, no comparison with non-classical methods has so far been made and hence the need for the current work. A linear geostatistical technique, as one of the non-classical techniques, is applied to a real data set along side the ANN technique. The performances of both techniques are judged against the actual data set as a means of cross-validation, and their results are in turn compared and analysed.

A summary of both techniques is presented omitting the more rigorous aspects which could be found in standardized texts of the respect subject areas.

A linear geostatistical technique is compared with artificial neural network technique for the prediction of sample values at pre-defined locations. The information gained from trained network, shows it to be an appreciable and cost effective tool where there is a constraint in obtaining further real data, albeit at a lower confidence level. The geostatistical approach addresses this to an extent with a conceptual approach to errors through the study of a structural function underlying the regionalized variables and Kriging variances after estimation. Kriging is performed on a data set for a bauxite deposit with a random stratified grid. Results from a geostatistical approach are compared with predictions from an artificial neural network system. Both methods are compared from the point of view of relative performance with reference to actual values.

The Data Set

A bauxite data set is used in this study. It comprises x and y co-ordinates of the drill hole locations, thickness of the mineralization and assays for the first mineral (% Al_2O_3) alumina, and the second mineral (% SiO_2) silica respectively. The drill hole locations with an approximate sampling interval of 100ft, are shown in Fig. 1.

This data set is organized in such a way as to reveal the type of distribution, especially as the choice of the mathematical method to be used for block estimation is usually a function of the frequency distribution of the sample values. The geology

of the deposit also influences this choice. These are represented in the histograms in Figures 2.

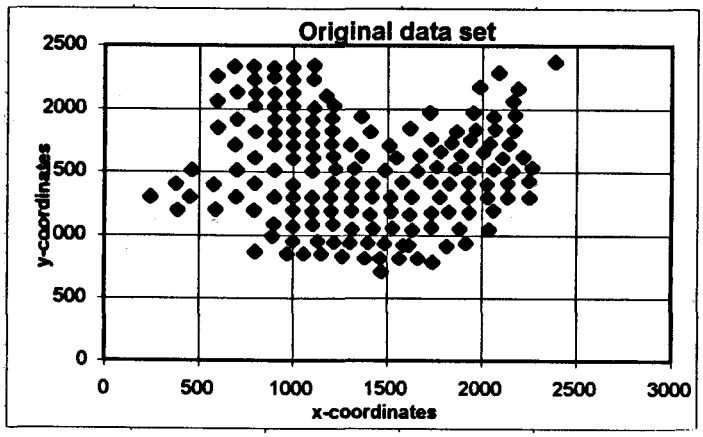


Fig. 1. Distribution of bore hole locations

However, since only a small proportion of the data values can be considered as having extreme values, their impact on estimation is minimal, and hence no assumptions are made about the distribution. The statistics of the raw data are presented in Table 1.

Table 1 : Statistics of the Raw Data

Variable	Mean	Variance
Thickness	123.58	6554.56
Alumina	248.71	58660.88
Silica	1138.31	1026845.53

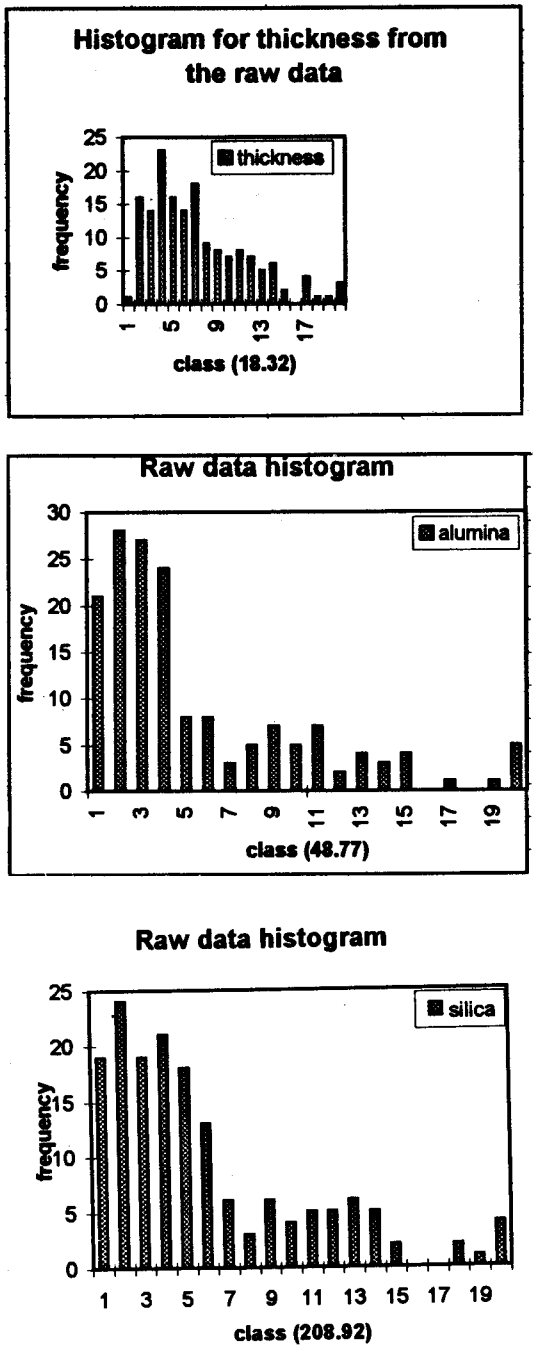


Fig. 2. Data distribution

It is evident from the histograms of the raw data that it is not distributed normally. This is typical of most deposits where sample values are not symmetrical. The sample values in the case study are positively skewed. The variances of the data values are seen to be highest for silica. This is possibly due to an increase in the relatively 'erratic' values occurring at the tail end of the skewed distribution. These features are further characterized in the structural analysis of the individual variables.

GEOSTATISTICS

Theoretical Considerations

The theory of geostatistics has already been well documented by a number of authors (Journel and Huijbregts 1978, Matheron 1971). Important aspects that emphasise the criteria for a good estimation of regionalized variables are summarized below.

Random Functions

Many earth-science variables exhibit major characteristics which are **random behaviour and a certain level of continuity.** While it is appreciated that the variable is random, knowledge of the value at a point x gives some information about other values in its neighbourhood, this phenomenon is simply described as 'continuity'. The information, however, may be insufficient to enable accurate prediction of the new values in the neighbourhood, and the variable would exhibit a random behaviour. This randomness is accounted for by considering the value **z(x)** of the variable **z** at location **x** as a realization, or the outcome **(e)**, of a random variable **Z(x).** The element of continuity is given by the variogram $\gamma(h)$ which associates to each distance **h,** a measure of the difference between values a distance **h** apart.

$$2\gamma(h) = \text{variance } (Z(x+h) - Z(x))$$

An experimental variogram is first calculated for the data, and then fitted by a simple mathematical function. This function then represents the degree of continuity of the variable under study.

The Spherical Scheme

About 90% of known orebodies have an intrinsic scheme(Royle, A.G.) This is a scheme characterizing a transition phenomenon, i.e. one in which the semi-

variogram reaches a finite value as h increases indefinitely. The value of h at which $\gamma(h)$ reaches a finite value is called the range of influence of the phenomenon and is denoted by the letter a.

In Matheron's spherical scheme

$$\gamma(h) = C_0 + C\{\frac{3}{2}\frac{h}{a} - \frac{1}{2}(\frac{h}{a})^3\} \text{ for } h < a$$

$$= C_0 + C \quad \text{for } \geq a$$

other schemes that represent transition phenomena are the exponential type in which

$$\gamma(h) = C(1 - e^{-h/a})$$

ANN Architectural Design

As illustrated in Figure 3, the network architecture is composed of many simple processing elements that are organised into a sequence of layers. These are the input, hidden, and the output layers. The neurons in the input layer receive two input signals representing the co-ordinates (X_i) and (Y_i) in a given grid. Therefore, two neurones are used in the neural network input layer.

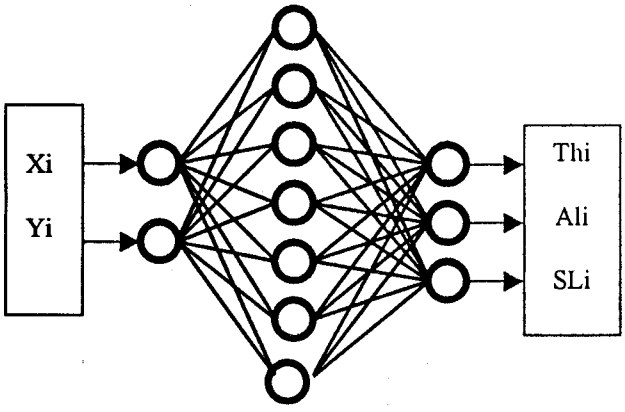


Fig. 3. The proposed ANN architecture

The output layer, on the right-hand side, consists of three output neurons representing the thickness of mineralization of the deposit, the percentage content of alumina, and the percentage content of silica. Between the input and output layers, there is generally one or more hidden layers. Determining the number of hidden layers and the appropriate number of neurons in each hidden layer is not an exact science. Research in this area [5,6] has proved that one or two hidden layers with an adequate number of neurons is sufficient to model any solution surface of practical interest. It has been found that a network with 6 to 8 nodes in the hidden layer would be a good choice. The network containing 7 hidden nodes, however, yielded the smallest error over the region of interest.

Based on these results, one hidden layer containing seven neurons was used to develop the ANN architecture.

Network Training

The multilayer feedforward networks developed in this work were trained using the backpropagation (BP) paradigm developed in reference [8]. The BP algorithm uses a supervised training technique. In this technique, the interlayer connection weights and the processing element thresholds are first initialized to small random values. The network is then presented with a set of training patterns, each consisting of an example of the problem to be solved (the input) and the desired solution to this problem (the output). These training patterns are presented repeatedly to the ANN model and weights are adjusted by small amounts that are dictated by the general delta rule [8]. This adjustment is performed after each iteration when the network's computed output is different from the desired output. This process continues until weights converge to the desired error level or the output reaches an acceptable level. The system of equations that provides a generalized description of how the learning process is performed by the BP algorithm is shown in Simpson [9].

For the present work, the training process was performed using the NeuroShell^{TM} simulator. After several adjustments to the network parameters and 12 hours and 43 minuets of training time, the network converged to a threshold of 0.0001. Predictions from this trained model were in agreement with the actual data, thereby producing an R^2 value of 0.72 for TH_i, 0.78 for AL_2O_3 and 0.82 for S_iO_2. These results indicate that 72 percent of the variability of thickness, 78 percent of alumina, and 82 percent of silica can be explained by the ANN model developed in this work.

Having trained the network successfully, the next step was to test the trained network, using the test data set, in order to judge its performance.

Network Testing and Validation

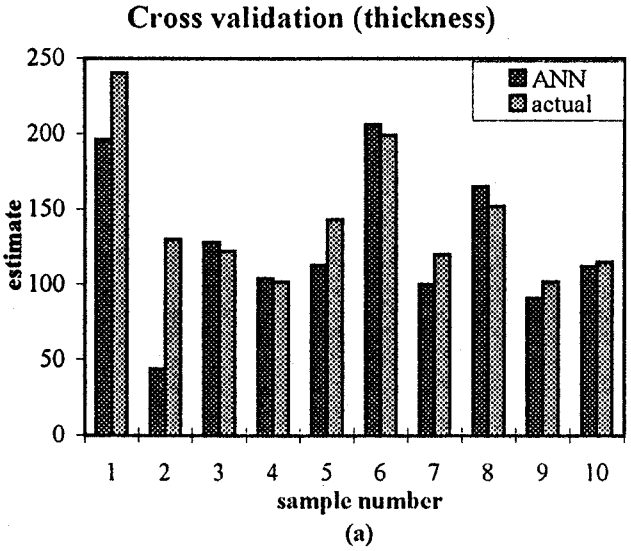
The generalization capability of the models was tested by presenting 35 patterns that were excluded from the data set prior to network training. Figures 4 (a), (b), and (c) provide the results and illustrate the relationship between actual results and estimates obtained from the ANN-based model for thickness, alumina and silica contents for ten randomly selected sample locations.

To validate these results, three of the more common techniques that are generally used to determine the accuracy and performance of the model were used. These techniques are: mean absolute deviation (MAD), mean squared error (MSE), and mean absolute percentage error (MAPE). MAD is the average of the absolute differences between the estimated values and the observed or experimental values. MSE is the average of the squared differences between estimated and observed values, and MAPE is the average of the absolute differences between the estimated and observed values expressed as a percentage of the observed values.

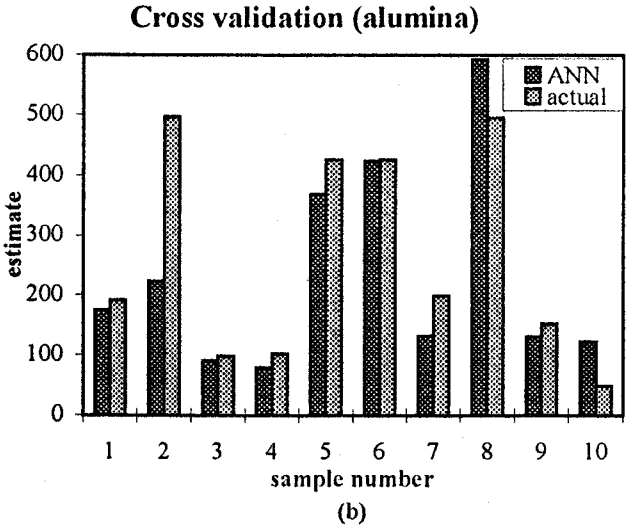
Statistical analyses of the results, shown in Table 2, indicate that the R^2 values for the testing set were 0.876 for TH_i, 0.755 for AL_2O_3, and 0.792 for S_iO_2. The results show that approximately 79 percent of the variation in the dependent variable estimates can be explained by the independent variables selected and the data set used. These results demonstrate that the ANN-based model developed in this work can predict the grades at unsampled locations with acceptable accuracy for the mineral under discussion. Verification of the ANN-based model's results is attempted using a geostatistical model with the same data and the results compared.

Table 2. Statistical Analysis of ANN Models

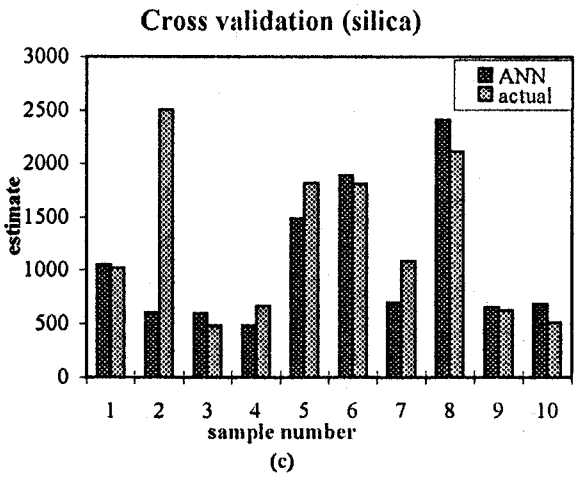
	TH_i	AL_2O_3	SiO_2
MSE	807.20	12746.30	178645.500
MAD	21.22	58.59	253.135
MAPE	25.82	33.72	35.740
SD	15.00	41.43	178.990
R^2	0.876	0.755	0.792


Cross validation (thickness)
ANN
actual
estimate
250
200
150
100
50
0
1 2 3 4 5 6 7 8 9 10
sample number

(a)

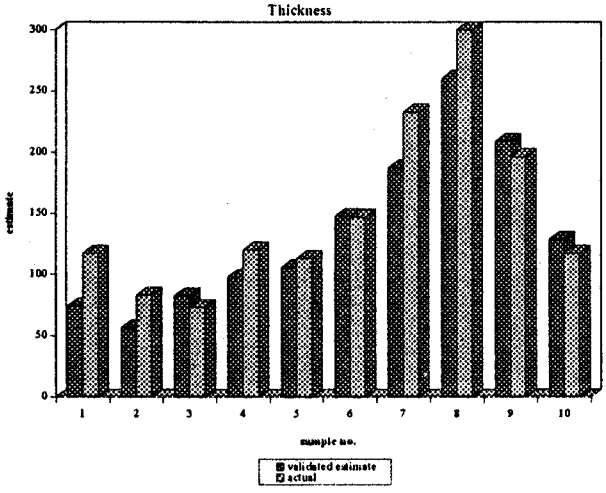

Cross validation (alumina)
ANN
actual
estimate
600
500
400
300
200
100
0
1 2 3 4 5 6 7 8 9 10
sample number

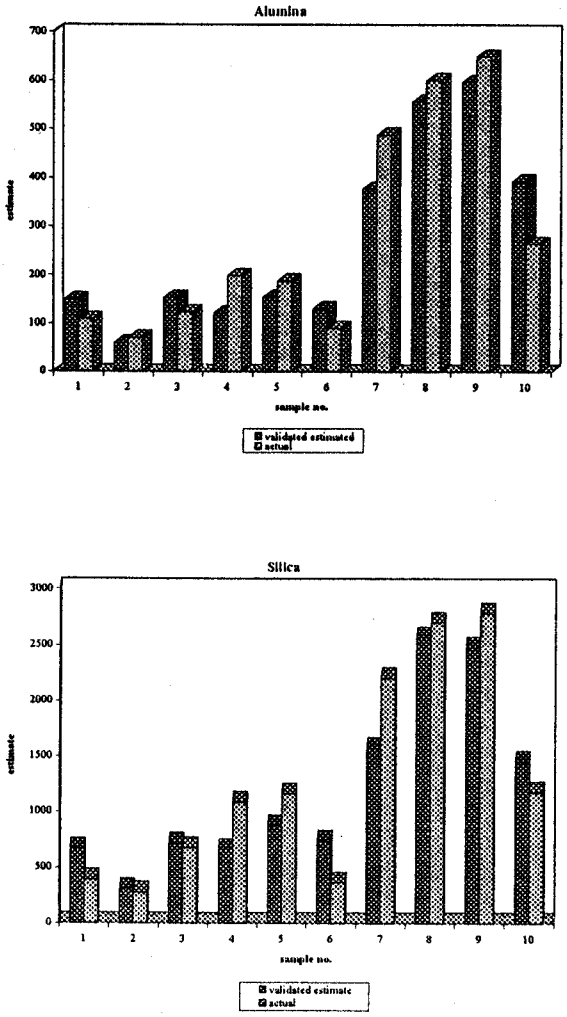
(b)



(c)

ANN cross-validated estimates





GEOSTATS cross-validated estimates

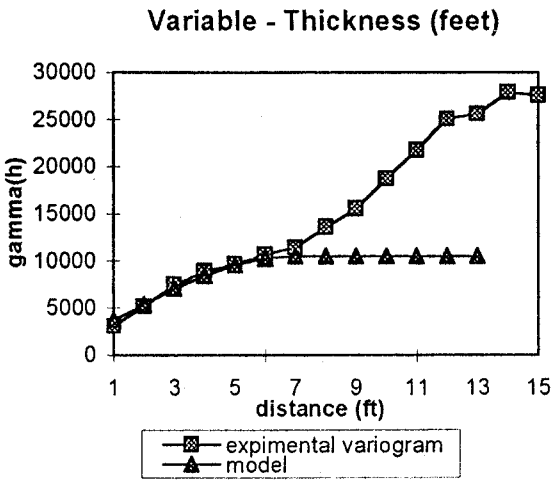
Fig. 4. Validated estimates by ANN and geostatistics

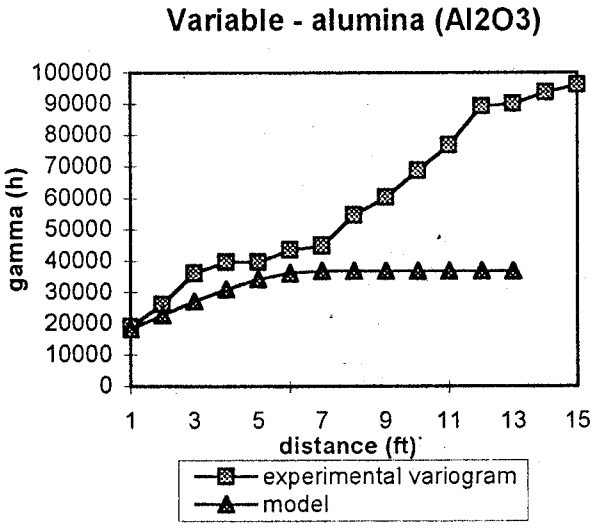
Variography and Structural Analysis

A geostatistical approach involved the computation of experimental variograms for the three regionalized variables that were used in the ANN approach.

Modelling

A spherical scheme model was selected for the modelling process as most known orebodies exhibit an intrinsic scheme model. The computed experimental variograms and their models are shown in Figure 5. In the modelling process, a necessary condition, was that the mean squared error of the estimation between the true values and the estimated values $(z^* - z)^2$ was less or approximately equal to the





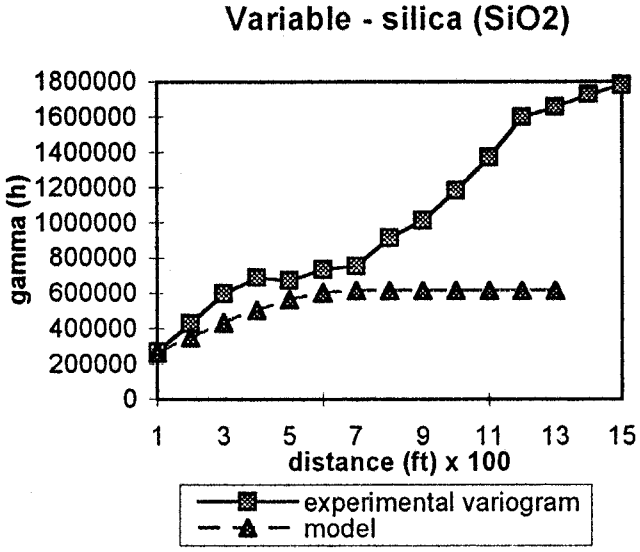


Fig. 5 : Variogram models

$$(z^* - z)^2 \leq \sigma_k^2 < 1.1(z^* - z)^2$$

mean Kriging variance (σ_k^2), and the Kriging variance is in turn less than 1.1 times the mean squared error during point Kriging. A relative measure of the goodness of fit for the model is given by the value of epsilon ε,

$$\varepsilon = \frac{C_o}{C}$$

The closer the value of epsilon is to zero the better is the model. It can be seen from the table below that the variable thickness has the lowest value of epsilon. The three variables showed a consistent range in the modelling process, although the variability in silica was found to be very prominent as observed from the raw data sample statistics. The summary of the modelled parameters are shown in Table 3.

Table 3: Modelled Variogram Parameters

Variable	C_o	C	Range	C_o / C
Thicknes	430.0	3170.0	700.0	0.135
Alumina	11800.0	25200.0	700.0	0.468
Silica	14200.0	47300.0	700.0	0.300

Cross Validation

To test the effectiveness of the estimation procedure, the modelled parameters were put through a cross validation process, as a means of justifying the Kriging technique chosen to perform the evaluation. It also serves as a test for the fit of the semi-variogram model to the data. The technique of cross validation is well documented in many texts and publications. (I.Clark, M.David (1979), Parker et al (1979))

Without loosing focus on the comparison between ANN and a geostatistical approach, a summary of the technique of cross validation is included here. It is the removal of one sample from the data set, and from the remaining samples in the data set, surrounding samples are used to estimate the value at this supposedly 'unsampled' location.

Suppose the total number of samples in the data set is n, the removal of one sample leaves (n-1) samples. The nearby samples to the removed value within the (n-1) samples in the data set are used to obtain an estimate $z^*(x_j)$. The difference between the true grade and the estimated grade at the sample location is the actual error e_r, while the 'theoretical' or 'expected error' is measured by the Kriging variance during estimation. The procedure is repeated for the entire data set, ideally, the sum of the differences is expected to be zero, thus serving as a proof that the original semi-variogram model fits the data in question. This proof is illustrated in the comparison between actual values and geostatistically estimated values based on the models for ten (10) randomly selected locations shown in Figure 6. Their performance for the three variables are shown in histogram form in Figures 4 (a), (b), & (c).

A similar test approach by ANN to further test the effectiveness of the validated model is illustrated with a prediction for thickness, alumina and silica grades at fifty-eight (58) randomly selected locations on the grid. Co-ordinates of the specific points on the grid were selected as shown by Figure 6. An estimate was then obtained from the model based on the given co-ordinates. Examples of these estimates are shown in Table 4.

COMPARISON BETWEEN BOTH TECHNIQUES

For purposes of effective comparison, both the techniques of geostatistics and the artificial neural network were compared. To test their performance, a data sub-

set from the original data set at thirty-seven locations were carefully but randomly selected. The selection was based on a spread within the domain of mineralization. The results given by both techniques were compared in turn individually and severally against the actual or 'true' values at the same locations.

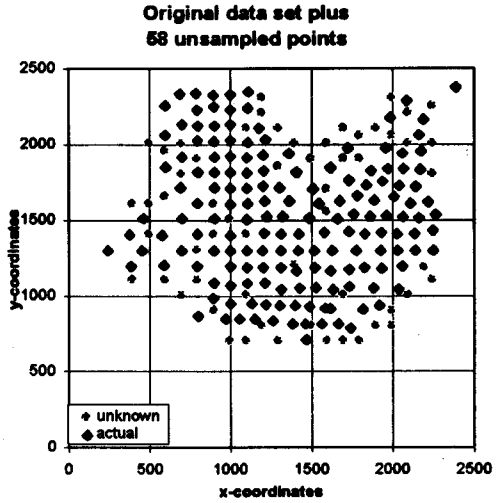


Figure 6. Additional locations selected on the grid

Table 4. Example of Estimates for Locations Selected

Co-ordinates		TH_i	AL_2O_3	SiO_2
X	Y			
400.00	1000.0	57.44	43.04	223.16
400.00	1500.0	74.40	45.39	264.72
500.00	1300.0	55.29	44.04	222.47
500.00	2000.0	76.51	45.88	270.64
600.00	1500.0	47.73	44.14	207.36
600.00	1950.0	69.93	46.13	259.20
700.00	1600.0	36.55	44.46	184.55
800.00	1300.0	25.46	42.37	152.27
900.00	900.0	25.40	26.77	102.89
1000.0	700.0	4.00	3.00	41.00
1100.0	1400.0	74.79	119.06	569.64
1500.0	2200.0	113.24	128.65	763.68
1800.0	1900.0	195.38	174.27	1210.1

For purposes of clarity, only one of the cases containing the entire thirty seven points is shown in Figure 7. The rest are represented on only ten randomly selected locations within the thirty-seven as shown in Figure 8. The summary statistics for both techniques on the same data set are presented in the following tables for comparison.

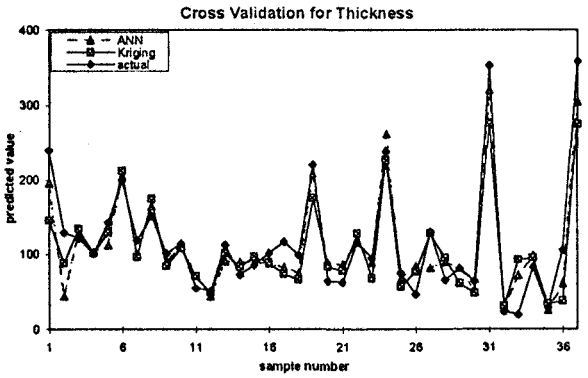


Fig. 7. Estimates by ANN and Kriging

The summary statistics concentrate on measures of location, spread and shape. These measures were compared to the estimates and the residuals and analysed statistically.

Estimates (thickness)

	Actual	ANN	Kriging
n	37.00	37.00	37 .00
m	117.892	111.135	107.585
σ	78.713	70.077	60.759
cv	0.667	0.631	0.565
min	20.00	26.00	30.72
Q_1	66.00	75.00	70.65
M	102.00	90.00	89.50
Q_3	130.00	113.00	128.91
max.	359.00	321.00	276.55

Estimates (alumina)

	Actual	ANN	Kriging
n	37	37	37
m	226.054	204.281	218.069
σ	218.836	160.178	165.248
cv	0.968	0.784	0.758
min	3	41	39.4
Q_1	102	119	120.54
M	146	144	147
Q_3	278	212	288.9
max.	1104	593	714.64

Estimates (silica)

	Actual	ANN	Kriging
n	37	37	37
m	1035.08	969.351	998.588
σ	923.945	782.36	712.648
cv	0.893	0.807	0.714
min	41	149	192
Q_1	509	595	563
M	677	711	714
Q_3	1211	953	1182
max.	4632	3595	3154

Residuals (thickness)

	ANN	Kriging
n	37	37
m	0.181	0.278
σ	15.002	18.266
IQR	38	58.26
MAE	0.258	0.281
MSE	807.16	1242.02

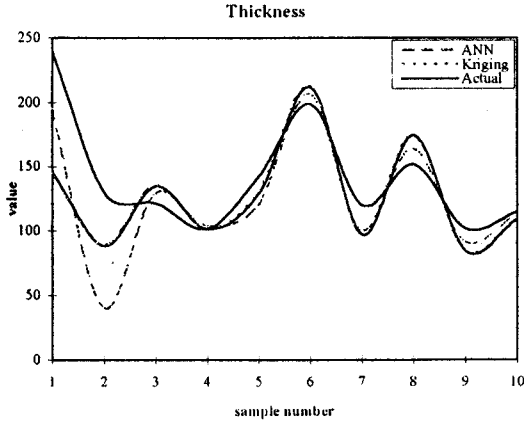
Residuals (alumina)

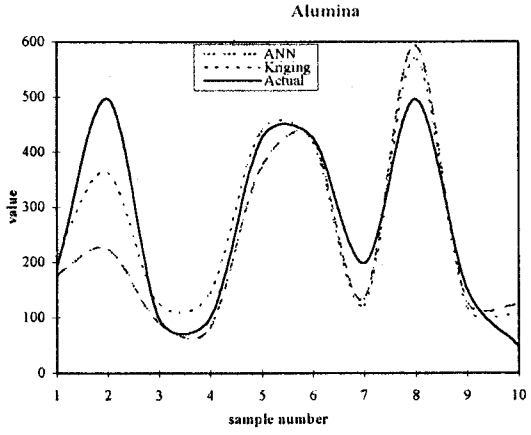
	ANN	Kriging
n	37	37
m	0.588	0.215
σ	41.88	44.40
IQR	93	168.36
MAE	0.341	0.335
MSE	12771.77	10258.01

Residuals (silica)

	ANN	Kriging
n	37	37
m	1.776	0.986
σ	178.99	187.48
IQR	358	619
MAE	0.357	0.302
MSE	178645.5	173246.3

For the same set of results, a comparison between estimates from Kriging and the ANN for ten randomly selected locations are presented in Figure 8.





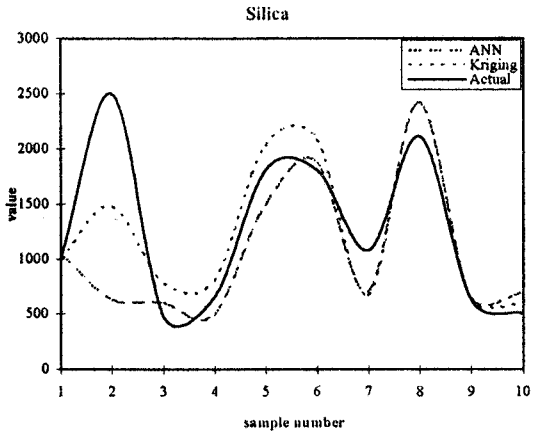


Fig. 8. Comparison between estimates from Kriging and the ANN for randomly selected locations

STATISTICAL ANALYSES

Residuals

Generally, a technique that gives a residual mean (true values minus the estimated values) closer to zero with less spread is preferred over another technique. Summary statistics of the randomly selected locations are shown in Table 6. ANN gave residuals of the mean as 0.181, 0.588 and 1.776, while Ordinary Kriging gave corresponding values of 0.278, 0.215, and 0.986 for thickness and the concentrations of alumina and silica. Apart from ANN having a lower residual mean for thickness which is closer to zero, Ordinary Kriging gave lower residual means for alumina and silica estimates which are in turn closer to zero.

A residual distribution of the mean of 0.278, represents a difference of only 8% , implying that Kriging produced biased estimates for thickness with a margin of 8% under estimation. For the other two variables, Kriging gave lower residuals of the mean, with values of 0.215 and 0.986 compared to ANN values of 0.588 and 1.776 for alumina and silica. These values represent only 10% and 6% under estimation for alumina and silica by ANN. A similar pattern of results is also seen in the comparison of standard deviations, mean absolute errors(MAE) and the mean squared errors(MSE)

Estimates

Comparing the distribution of the estimates by ANN and Ordinary Kriging to the distribution of the true values, the means of the estimates by Kriging are in general closer to the means of the actual values compared to the ANN means. Lower variances and standard deviations also observed point to the fact that the distribution of Ordinary Kriging estimates have less spread than the distribution of the ANN estimates and the true value estimates. This is in consonance with the smoothing effect of Kriging. A procedure that uses few nearby sample values will produce estimates that are less smoothed than one that combines many nearby values such as Ordinary Kriging.

Lower and Upper Quartiles

The data set is split into quartiles and arranged in increasing order, a quarter of the data falls below the lower quarter or first quartile (Q1), and a quarter falls above the upper or third quartile (Q3).

Ordinary Kriging estimates gave Q1 and Q3 values that are consistently closer to those of the true values than the ANN estimates. This confirmed that the distribution of the Ordinary Kriging estimates is closer to that of the true values than the ANN estimates.

Based on the statistics presented, the differences in the results obtained by both techniques are negligible. This emphasizes the effectiveness of the estimation techniques, especially where there is a large amount of data. Ordinary Kriging is, however, seen to be marginally better.

To further test the effectiveness of both techniques for the purposes of cost saving, another comparison was performed at unsampled locations and their relative performances assessed.

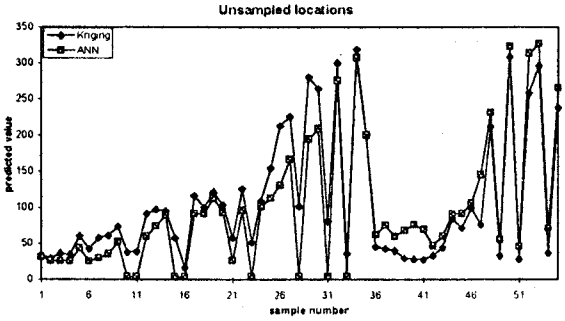
Estimation of Unsampled Locations

Point estimation using Ordinary Kriging and ANN for locations that were not sampled was performed with the view to determining the relative performance of each technique. Most importantly the suitability of ANN as an effective tool in predicting values at unsampled locations was also determined with the ultimate objective of using ANN as a cost saving tool in a drilling operation. However, the

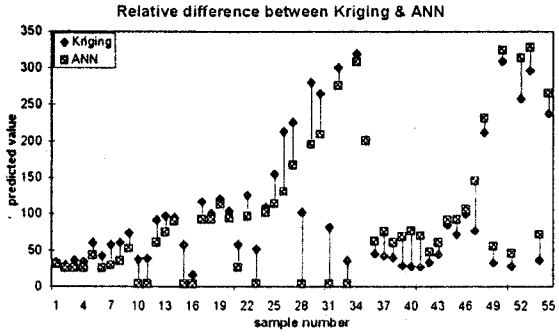
true grades at these points are not known. A grid showing 58 randomly selected unsampled locations together with 163 locations whose true values are known is shown in Figure 6.

Estimates obtained using both techniques at these locations are shown in Figure 9 for thickness, one of the three variables. Although the comparison here is relative, both techniques behaved in line with their previous performance when the estimates and residuals were compared (Fig.9a). A significant difference between both techniques is, noticeable at locations where ANN gave very low predictions (Fig.9b). This is essentially due to insufficient data being available, resulting in such locations being assigned minimum values that were set during the ANN network training.

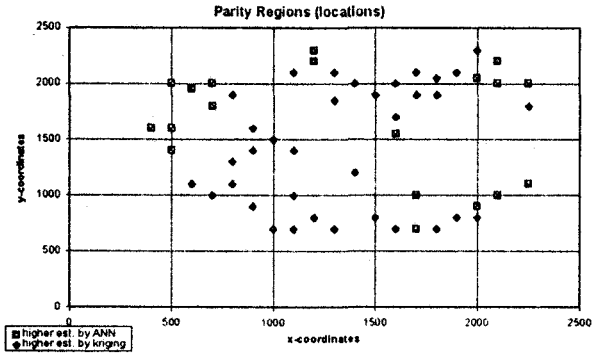
This effect is further illustrated by examining specific locations (parity regions) having with very low predicted values. Figure 9c shows that ANN predictions were generally higher than the Ordinary Kriging estimates at the periphery of the deposit.



(a)



(b)



(c)

Fig. 9 Predictions by both techniques

This observation was found to be consistent for all the estimated variables, while the Ordinary Kriging estimates are seen to be higher within the deposit. These, however, are not any higher than the estimates at known locations, especially when compared to the analyses between true and predicted grades and the residuals.

SUMMARY AND CONCLUSION

In this study, artificial neural network and geostatistical based models were developed and used in the estimation of point grades within a defined grid system. This consisted of 163 sample locations, their x and y co-ordinates, thicknesses and their assays.

A comparison of both techniques showed a very close agreement in the results obtained, provided there is a significant volume of data. Differences in the results were generally in magnitudes of 8%, 10% and 6% for the three estimated variables, thickness, percentage alumina, and percentage silica, when compared with the actual grades. Based on the performance of both techniques on known sample locations, the models were extended to the prediction of point grades at unsampled locations for the same variables. Results at this stage when compared showed an increase in disparity between both techniques as data become sparse, with the geostatistical model having an edge on the ANN model. This suggests that the ANN technique is more data sensitive than the geostatistical technique which is seen to be more robust.

This disparity is particularly significant at the edge of the deposit where data are generally less concentrated. Like most estimation algorithms that are linear and distribution dependent, their performance at these locations become reduced.

It can therefore be concluded that, ANN can be used as an effective tool in mineral resource evaluation at the production stages as a complimentary tool to a geostatistical technique by virtue of its speed, and ease of operation.

ANN may have further uses at the exploration stage such as being used to select an appropriate drilling grid or closing in on a drilling grid pattern.

REFERENCES

[1] **Park, D. C., El-Sharkawi, M. and Marks II, R. J., 1991**. Electric Load Forecasting Using an Artificial Neural Network. IEEE Trans. PAS, PAS-6, pp. 442-449.

[2] **Kalam, M. Z., Al-Alawi, S. M. and Al-Mukheini, M. 1995**. The Application of Artificial Neural Networks to Reservoir Engineering, Proceedings of IATMI International Symposium on Production Optimization, Bandung, Indonesia, pp.1-8.

[3] **Islam, S. M., Al-Alawi, S. M. and Ellithy, K. A. 1995**. Forecasting Monthly Electrical Load and Energy for A Fast Growing Utility Using an Artificial Neural Network. Electric Power Systems Research Journal, 34, pp.1-9.

[4] **Al-Alawi, S. M., Ellithy, K. A. and Choudhry, M. A., October 1995**. Tuning of Power System Stabilizer Over a Wide Range of Load Models Using an Artificial Neural Network. IEEE North American Power Symposium (NAPS).

[5] **Hecht-Nielsen, R., 1989**. Theory of the Backpropagation Neural Network, Proc., Int. Joint Conf. on Neural Networks. IEEE, Washington, D. C., Vol. 1, pp. 593-605.

[6] **Lapedes, A. and Farber, R., 1988**. How Neural Network Works, Neural Information Processing Systems, American Institute of Physics, pp. 442-456.

[7] **Carpenter, W. C. and Hoffman, M. E., March 1995**. Training Backpropagation Neural Networks, AI Expert, Vol. 10, No. 3, pp. 30-33.

[8] **Rumelhart, David E. and McClelland, James L., 1966**. Parallel Distributed Processing: Exploration in the Microstructure of Cognition, Vol. 1, Foundations, Cambridge, Massachusetts, MIT Press.

[9] **Simpson, Patrick K., 1990**. Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations. First Edition, Elmsford, New York: Pergamon Press, Inc.

[10] **David, M, 1977**. Geostatistical Ore Reserve Estimation. Elservier, Amsterdam, p. 364.

[11] **Issaks and Srivastava, 1989**. An Introduction to Applied Geostatistics. Oxford University Press, p. 561.

[12] **Journel, A.G. and Huijbregts, C., 1978**. Mining Geostatistics, Academic Press, London.

[13] **Rendu, J.M., 1984**. Geostatistical Modelling and Geological Controls. Application of Computers and Mathematics in the Minerals Industries. Institution of Mining and Metallurgy, London. pp467-476.

[14] **Rendu, J.M., 1978**. An Introduction to Geostatistical Methods of Mineral Evaluation. South African Institute of Mining and Metallurgy, Johannesburg.