Ore Grade Estimation with Modular Neural Network Systems – A Case Study

I K Kapageridis[+] and B Denby

[+]Sponsored by the State Scholarships Foundation of Greece

ABSTRACT: This paper introduces a neural network approach to the problem of ore grade estimation. The system under consideration consists of three neural network modules each responsible for a different area of the deposit, depending on the sampling density. Octant and quadrant search is used as a way of presenting input patterns to the modules. Both radial basis function networks and multi-layered perceptrons are used as the building blocks of these modules. An iron ore deposit provides the training and testing data for both the neural network system and kriging, and the results from the two approaches are compared.

Keywords: ore grade estimation, neural networks, radial basis functions, function approximation.

INTRODUCTION

The problem of grade estimation is quite complicated due to the very complex and, sometimes, not very well understood processes of orebody deposition. Most of the approaches to date, including the most important – geostatistics, have been based on certain assumptions about the spatial distribution of ore grades within the orebody. Efforts to limit the effects of these assumptions have led to far more complicated methods requiring a large amount of knowledge in order to be effectively applied.

This paper presents a new method for grade estimation based on modular neural network (MNN) systems. It is not the first neural network based approach to this problem. In the past, Wu and Zhou (1993), Clarici et al (1993) and Burnett (1995) have applied neural network techniques to grade and reserve estimation with some success. The system considered in this paper introduces some major changes in the overall network architecture and particularly on the way samples are presented to the system.

As in the other neural network approaches, the proposed system treats the relationship between the location of a sample and its grade as a complex function, only this time the function is broken down to simpler mappings of grades and distances from sample points surrounding each point considered and its own grade. The surrounding points are selected through octant and quadrant search leading to a completely direction oriented system following the paradigm of geostatistical variography. However, there is no need for calculating experimental variograms and therefore this approach requires considerably less knowledge. Radial Basis Function (RBF) networks learn the directional behaviour of grade data out of the given training samples. At the same time, a Multi-Layered Perceptron (MLP) learns the distribution of grades over the complete training set and is used for estimating grades at points in the borders of the deposit area where there are no surrounding points to provide any directional information.

FUNCTION APPROXIMATION USING RADIAL BASIS FUNCTION NETWORKS

Function approximation is one of the areas where RBF networks have been applied with success (Broomhead and Lowe, 1988; Lee and Kil, 1988; Moody and Darken, 1989a, 1989b; Leonard, Kramer and Ungar, 1992a, 1992b; Park and Sandberg, 1993). The model of RBF networks is motivated by the characteristics found in many parts of biologic nervous systems and also by work on interpolation with radial basis functions (Powell, 1987). Poggio and Girosi (1990) provide a more in-depth discussion on the approximation properties of RBF networks. The kind of input-output mapping from a set of examples, as in the case of grade estimation from limited number of samples, can be regarded as synthesising an approximation of a multi-dimensional function or, in the words of Girosi and Poggio (1990), solving the problem of hypersurface reconstruction.

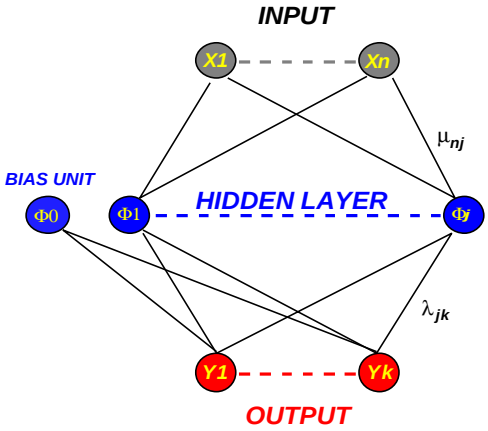

FIGURE 1: Basic structure of radial basis function network. The value of the bias unit in the hidden layer is set to 1.

The structure of an RBF network is that of a *feedforward*, single hidden layer, fully interconnected network with locally tuned hidden units (Figure 1). All hidden units simultaneously receive the n-dimensional real-valued input vector x (Hassoun, 1995). A non-linear basis function f_j is centred around each hidden unit weight vector m_j which also has an adaptable range of influence s_j (also called the *width of the receptive field* in the input space for this unit). The output of the hidden unit j, h_j is given as a radial function of the distance between each pattern vector and each hidden unit weight vector, $h_j = f_j(\|x-m_j\|/s_j)$ (Lowe, 1995). The output of the RBF network is a scalar product between the vector of hidden unit outputs and the weight vector attached to output unit k, l_k, as $y_k(x) = \Sigma\, h_j(x)\, l_k$. Key issues to the performance and accuracy are the choice of f_j , the number of basis functions used, their location, and their width.

Following the original formulation of Broomhead and Lowe (1988), there is a set of input/output pairs of input/target patterns representing data from an unknown smooth surface. During learning, the network centres basis functions on samples randomly selected or uses clustering for the choice of centres. The width of the receptive field of these functions is adjusted to improve performance, and the number of centres also changes to improve the overall approximation. The local fit performed by RBF networks leads to good generalisation from each training example. It should

be noted though, that the input data is never sufficient to reconstruct uniquely the mapping in regions where data are not available.

MODULAR NEURAL NETWORK STRUCTURE FOR GRADE ESTIMATION

The system under development consists of three individual estimation modules. There is also a data pre-processing module that receives the original X-Y-Grade values from the available samples. This module performs octant and quadrant search on these samples and builds training and validation sets for two of the estimation modules. In essence, the pre-processing module partitions the data set into three parts (Figure 2). During octant and quadrant search, the deposit area around each point considered is divided into sectors, e.g. in the case of octant search the sectors will be: WNW (West-North-West), NNW, NNE, ENE, ESE, SSE, SSW, and WSW. Clearly, the points qualifying from the octant search (i.e. have 8 surrounding points, one in every sector) do qualify for quadrant search as well (i.e. have 4 surrounding points).

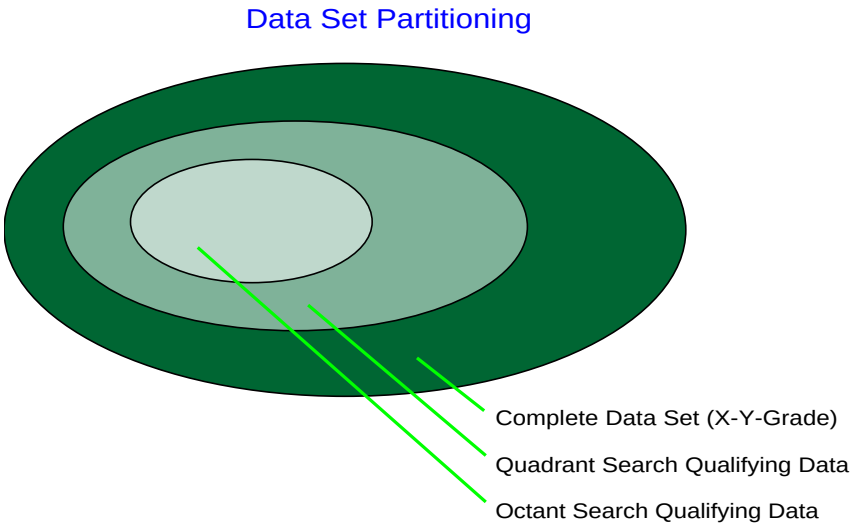

FIGURE 2: Partitioning of the data set according to the number of neighbour points available to each point under consideration – areas shown are not representative of the actual percentages.

Three data sets are built, one for each estimation module. Each record in the octant and quadrant training sets contains grade and distance from the surrounding points as the input fields and the grade at the estimation point as the output field for the network(s). For example, a record in the quadrant training set would look like this:

NW_grade, NW_distance, NE_grade, NE_distance, SE_grade, SE_distance, SW_grade, SW_distance, Target_grade

The third training data set is actually the complete data set of samples as entered in the pre-processing module.

Figure 3 shows a schematic diagram of the overall structure of the modular neural network system. The two RBF network modules, octant and quadrant, contain 9 and 5 radial basis function networks respectively. In the octant module, there are 8 RBF networks each responsible for one of the eight sectors, and each having two

inputs (grade, distance) and one output (target grade). The outputs of the eight RBF networks are directed to a final RBF network that gives the final estimate. In other words, the final RBF network learns to combine the individual estimates in order to provide a more accurate estimate. The quadrant module follows the same formulation with the difference of having only four RBF networks for each of the four sectors (NW, NE, SE, SW).

The parameters of these individual RBF networks vary from sector to sector. This is due to the varying complexity of the mapping that each one is trying to achieve. In the octant module, the number of centres, i.e. basis functions, varies from 4 to 10, while in the quadrant module this varies from 10 to 17. The number of centres is determined during training based on the minimum error achieved on the validation set. The non-linear function used is the multi-quadratic: $f(d) = (d^2 + c^2)^{1/2}$, where d is the distance from the centre and c is a smoothing parameter. The initial location of centres is on selected samples from the input set. The main difference between the sector RBF networks and the final RBF networks, apart from the number of inputs, is in the error distance measuring technique. In the sector RBF networks, *Euclidean* distance is used, while in the final networks the *City Block distance* (*Minkowski metric* of power one) is used. This was found to improve performance during training and validation.

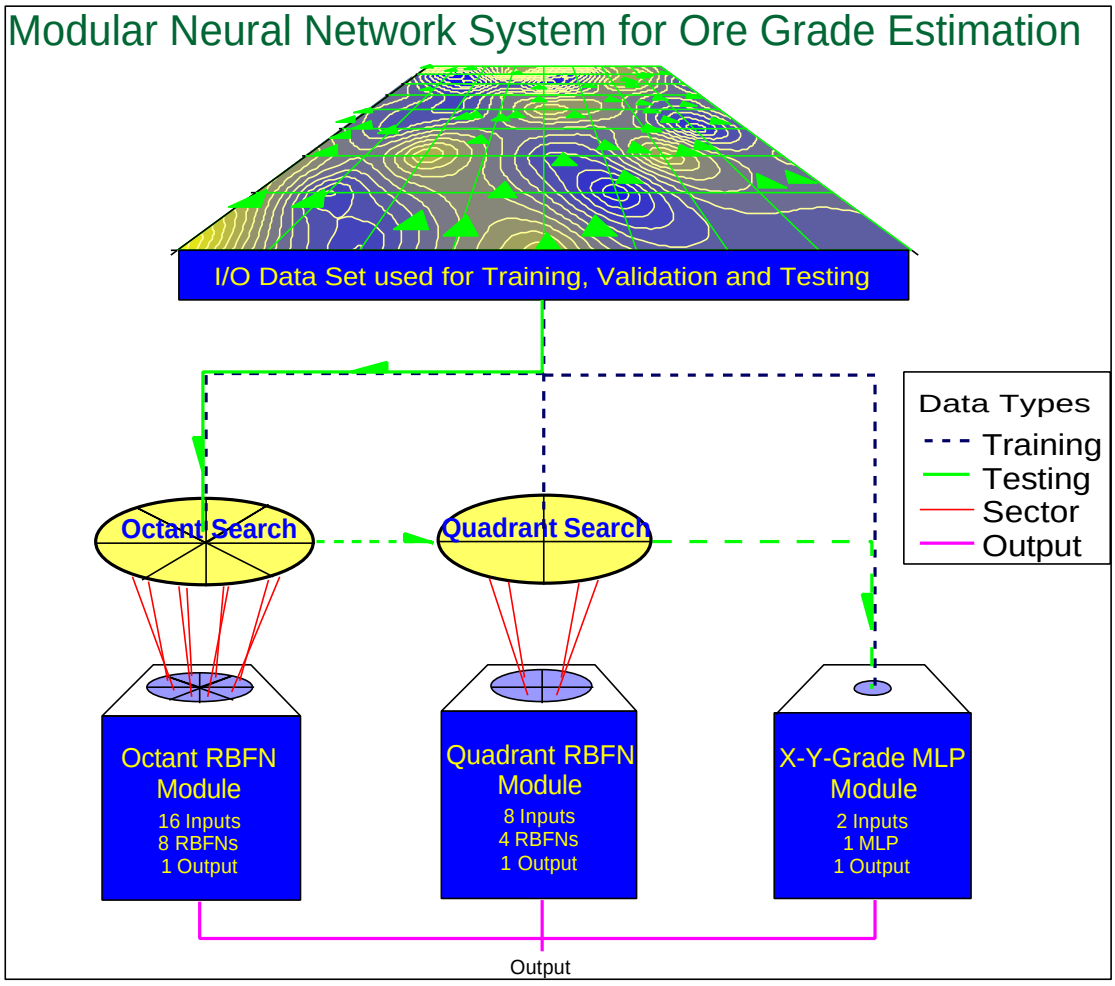

FIGURE 3: General structure of the developed modular system. Details of each module are not shown.

The X-Y-Grade module is based on a multi-layered perceptron with two inputs, fourteen hidden layer units and one output. Training is performed using the steepest descent algorithm. The bipolar sigmoid activation function (tanh) is used in the hidden and output units. The network weights are updated after every complete pass of patterns.

APPLICATION OF THE MODULAR NEURAL NETWORK SYSTEM – AN IRON ORE DEPOSIT CASE STUDY

The system described above was trained and tested on a simulated iron ore deposit taken from a geostatistical study (Clark, 1979). Table 1 shows the 50 samples used for training and validating the system. From a geostatistical point of view, the simulated deposit has a range of 100m, a sill of 25% Fe, and no nugget effect. The semi-variogram is spherical. These parameters were used to obtain results from kriging.

TABLE 1: Samples taken from a simulated iron ore deposit (Clark, 1979).

Easting	*Northing*	*%Fe*	*Easting*	*Northing*	*%Fe*
0	170	34.3	5	195	33.9
10	40	35.5	20	105	32.5
15	135	28.6	25	155	29.6
55	145	29.4	50	40	30.6
125	20	41.5	155	15	40.4
175	50	36.8	145	125	30.1
120	180	33.4	130	185	35.3
160	175	36.0	175	185	41.1
240	185	30.2	220	90	28.5
260	115	33.2	205	0	40.1
235	15	33.7	265	65	24.4
365	60	34.3	390	65	31.6
285	110	35.3	325	105	39.5
345	115	31.0	310	150	34.8
335	170	27.4	385	165	29.9
325	195	33.9	325	220	37.8
350	235	37.6	375	215	29.8
290	230	39.9	200	230	37.4
10	390	27.2	55	375	27.4
85	380	34.2	395	245	36.5
50	270	30.2	165	355	40.8
200	280	30.4	270	285	32.9
400	355	39.9	365	340	40.0
360	335	40.0	330	320	44.1
335	310	40.6	330	290	41.4

After training with the above data set, each of the sector RBF networks has learned the relation between the grade and distance of the neighbouring point in its sector and the target grade. Figure 4 shows the activation graph of the WSW sector RBF network of the octant module. Clearly, the activation increases with increasing neighbour grade (WSW G) and decreases with increasing neighbour distance (WSW D). There are however areas where this changes slightly showing the ability of RBF networks to map with more detail due to the local fit performed. Also the speed advantage of RBF networks over MLPs was once more confirmed. RBF networks required a few seconds to train while the MLP required a couple of minutes.

In order to test the system and to provide a comparison with kriging, 41 points on a diagonal grid were chosen to provide with some target grades. This test set was

again partitioned, only this time the sub-sets did not contain one another. In other words, the three modules were tested on different points and there was no overlapping of the test areas, as happened during training.

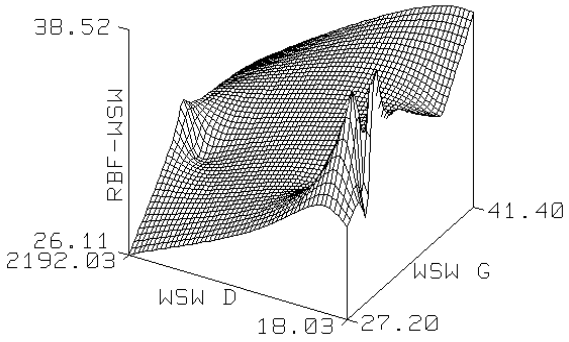

FIGURE 4: Activation graph of WSW sector radial basis function network, showing the learned relation to the grade (WSW G) of and distance (WSW D) from neighbour points in this sector.

Figure 5 shows the test data fit performed by both the MNN system and kriging. It is quite obvious that both methods tend to underestimate in high grade areas. The reason for this, at least in the case of the MNN, is because these areas are close to the borders of the deposit where the MLP is providing the estimates. The MLP module seems to give estimates close to the average grade. From the 41 test points, the MNN system is more accurate on 22 of them while kriging is doing better on the rest. The absolute error levels were the same with a very small advantage of MNN over kriging (0.2%).

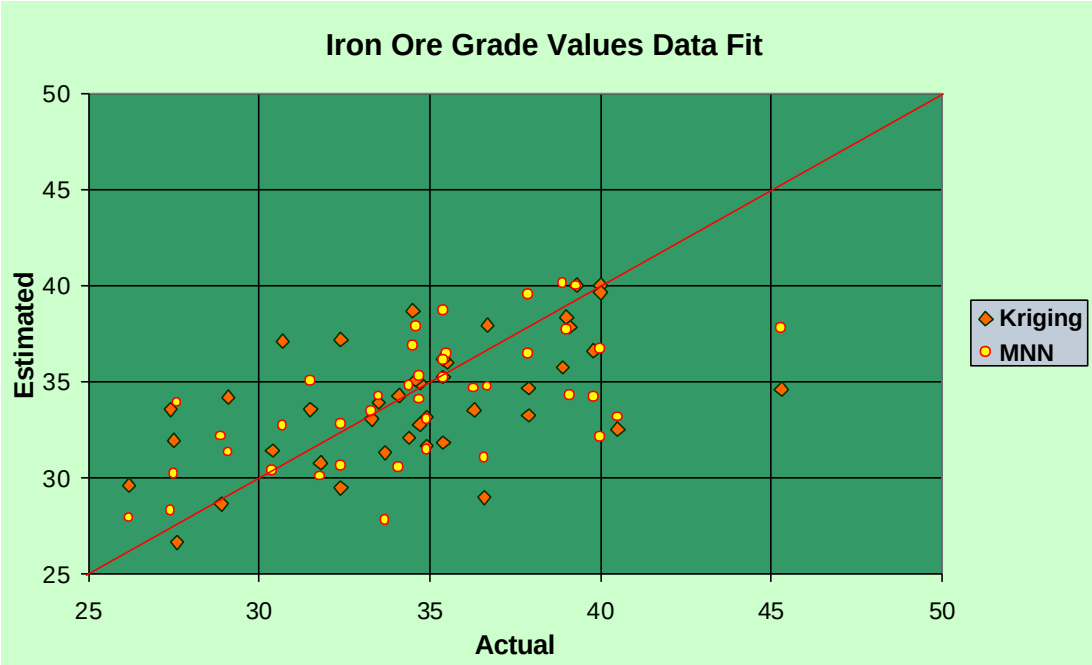

FIGURE 5: Data fit comparison graph for the MNN system and kriging.

Figure 6 shows contour maps of the actual and estimated grades as well as error levels for the two methods. Kriging and the MNN seem to perform better in different

regions except for a part in the south-west of the deposit where they both perform badly. Lack of enough training samples is the main reason for the high error level areas produced. The MNN system seems to map better the low grade area on the north-west region while kriging did better on the south-east.

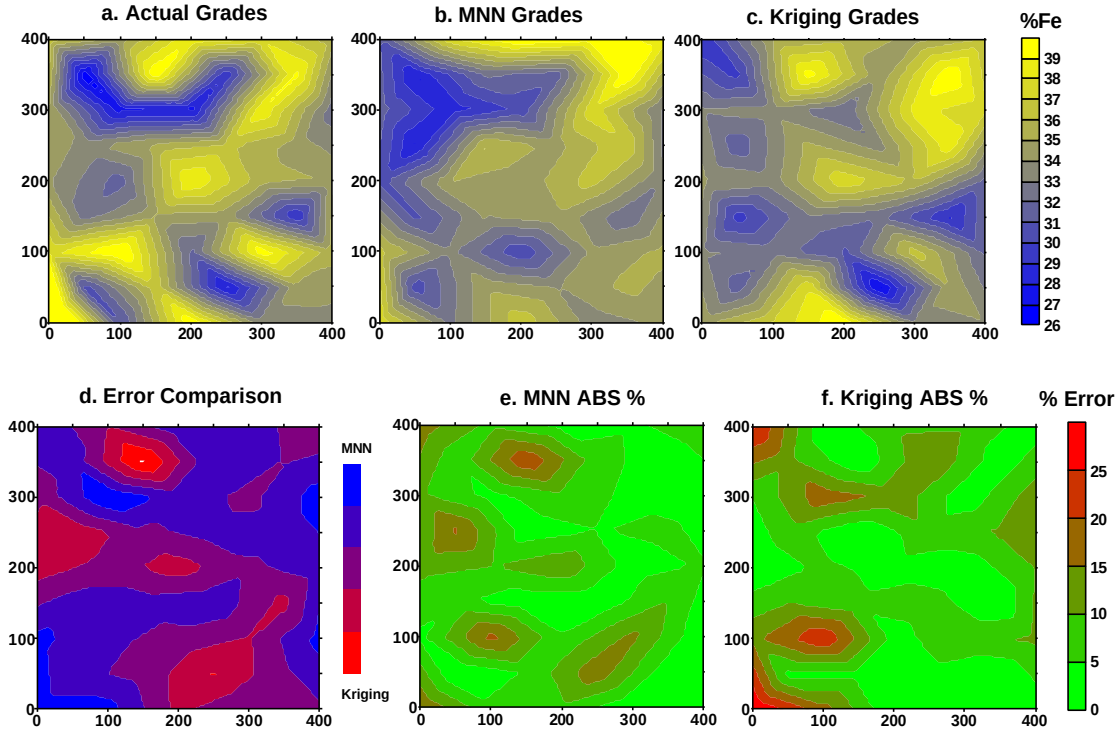

FIGURE 6: Top - Contour maps of actual grades (a), modular neural network estimates (b), and kriging estimates (c), Bottom –Contour map of absolute error difference showing where the MNN system (blue) performs better than kriging (red) (d), MNN absolute % error (e), and kriging absolute % error (f).

CONCLUSIONS AND FUTURE CONSIDERATIONS

This paper has shown how a modular neural network system developed for ore grade estimation can be applied on a randomly sampled deposit. The system was described and its main components were analysed. A basic discussion on the theory behind the neural networks used was also given. The results from both the MNN system in development and kriging were given showing that neural networks can actually provide comparable results with geostatistics while requiring far less knowledge to be effectively applied.

Further research to be carried out in the AIMS Research Unit will include the integration of the *Validity Index network* (VI-net) as proposed by Leonard et al (1992a, 1992b). The VI-net is an extension of RBF networks that calculates the reliability and confidence of its output and indicates local regions of poor fit and extrapolation. The VI-net will be an addition to the current architecture. Further tests on different deposits will also be carried out for further optimisation of the current system.

REFERENCES

Broomhead, D., & Lowe, D., (1988), Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2, 321-355.

Burnett, C., (1995), Application of Neural Networks to Mineral Reserve Estimation, Ph.D Thesis, Dept of Mineral Resources Engineering, University of Nottingham.

Clarici, E., Owen D., Durucan, S., and Ravenscroft, P., 1993, *'Recoverable Reserve Estimation Using a Neural Network'*, 24th International Symposium on the Application of Computers and Operations Research in the Mineral Industries (APCOM), Edited by Elbrond, Jorgen & Tang, Xiaoli, Montreal, Quebec, Canada.

Clark, I., (1979), Practical Geostatistics. Elsevier Applied Science Publishers, London.

Girosi, F., & Poggio, T., (1990), Networks and the best approximation property. *Biological Cybernetics*, 63, 169-176, Springer-Verlag.

Hassoun, M., (1995), Fundamentals of artificial neural networks. MIT Press, London.

Lee, S., & Kil, R., (1988), Multilayer feedforward potential function networks, in *Proceedings of the IEEE Second International Conference on Neural Networks* (San Diego, 1988), vol. I, pp. 161-171, IEEE, New York.

Leonard, J., Kramer, M., and Ungar, L., (1992), A neural network architecture that computes its own reliability. *Computers Chem. Eng.*, Vol. 16, No. 9, pp. 819-835, Pergamon Press.

Leonard, J., Kramer, M., and Ungar, L., (1992), Using radial basis functions to approximate a function and its error bounds. *IEEE Transactions on Neural Networks*, Vol. 3, No. 4, July 1992, IEEE.

Lowe, D., (1995), Radial Basis Function Networks, *in The Handbook of Brain Theory and Neural Networks*, ed. Arbib, M., MIT Press.

Moody, J., & Darken, C., (1989), Fast Learning in Networks of Locally-Tuned Processing Units, *Neural Computation*, 1, 281-294, MIT Press.

Park, J., & Sandberg, I., (1993), Approximation and Radial-Basis-Function Networks, *Neural Computation*, 5, 305-316, MIT Press.

Poggio, T., & Girosi, F., (1990), Networks for Approximation and Learning, *Proceedings of the IEEE*, Vol. 78, No. 9, September 1990, IEEE.

Powell, M., (1987), Radial Basis Functions for Multivariate Interpolation: A Review, in *Algorithms for The Approximation of Functions and Data*, Mason, J., and Cox, M., Clarendon Press, Oxford.

Wu, X., & Zhou, Y., (1993), Reserve Estimation Using Neural Network Techniques, *Computers & Geosciences*, Vol. 19, No. 4, pp. 567-575, Pergamon Press.

AUTHORS DETAILS

Mr. Ioannis K. Kapageridis
AIMS Research Unit,
Department of Mineral Resources Engineering,
University of Nottingham,
University Park,
Nottingham,
NG7 2RD United Kingdom
e-mail: enxik@unix.ccc.nottingham.ac.uk

Dr. Bryan Denby
Head of School of Chemical, Environmental, and Mining Engineering,
University of Nottingham,
University Park,
Nottingham,
NG7 2RD United Kingdom
e-mail: bryan.denby@nottingham.ac.uk