ELSEVIER

Computers & Geosciences 31 (2005) 704–717

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

Input space configuration effects in neural network-based grade estimation

I.K. Kapageridis*

Department of Geotechnology and Environmental Engineering, School of Technological Applications, Technological Education Institute of West Macedonia, Koila, Kozani GR 501 00, Greece

Received 9 June 2004; received in revised form 3 January 2005; accepted 3 January 2005

Abstract

The way input data are presented to Artificial Neural Networks is one of the most important parameters controlling their performance during the development and application stages. The choice of dimensions that form the input space of a network (dimensionality) is very important and must be investigated as to its effects on the performance of artificial neural network systems applied to grade estimation.

The study of these effects was achieved by configuring the available data in order to form different multidimensional input spaces and testing on different datasets. The results obtained from the numerous tests in this study lead to a better understanding of the behaviour of artificial neural network systems when facing different input space configurations using the same data, and aid the choice of dimensions that will allow better representation of samples for their development for grade estimation.

Keywords: Grade estimation; Spatial interpolation; Neural networks; Radial basis functions; Dimensionality

1. Introduction

The main scope of the study presented in this paper, was to investigate the effects of the choice of input variables to the performance of Artificial Neural Network (ANN) systems in grade estimation. The choice and total number of inputs form a problem known in the language of artificial neural networks as *dimensionality*.

The input variables, in a spatial interpolation problem such as grade estimation, define the way the estimated value is approached. In other words, as ANNs construct the projection from the input vector space to the output vector during training, it is clear that the input parameters define to a large extent this projection and the whole approach to this problem. For example, if we choose the n coordinates of the samples as input parameters then grade is considered a function of sample coordinates and the whole estimation problem becomes a simple grade surface fitting in the n-dimensional coordinate space.

The various ANN approaches to the problem of grade estimation provided in the literature are very interesting as they use different input parameters as well as different ANN architectures and learning algorithms. They also present differences in the quantity and quality of input data used for ANN development. Examples of ANN application to grade/reserves estimation are given by Caiti and Parisini (1991), Clarici et al. (1993), Wu and Zhou (1993), Burnett (1995), Kapageridis and Denby (1998a, b), Cortez et al. (1998), Kapageridis (1999) and Yama and Lineberry (1999).

*Tel.: +30 24670 85236; fax: +30 24610 41819.
E-mail address: ikapa@airlab.teikoz.gr.

0098-3004/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2005.01.001

1.1. Grade estimation and neural networks

Grade estimation is probably one of the most important stages in reserve calculations.

Over the past 30 years, geostatistics became the most established methodology for grade estimation. Since 1962 and the first introduction of G. Matheron, the field of geostatistics has sustained significant development in the mining industry, constantly producing better methods for solving the most complex cases of grade distribution (David, 1977; Journel and Huijbregts, 1978; Clark, 1979; Krige, 1981; Isaaks and Srivastava, 1989). However, the expert knowledge required to effectively apply these methods created a dependency between the reliability of the derived results and the skills and expertise of the person applying them.

An alternative approach that is considered particularly in the last decade is the application of ANN systems to grade estimation. ANN systems typically approach grade variance and distribution as complex functions in space, approached by their various components. These components usually consist of ANN architectures such as Radial Basis Function networks (RBF) or Multi-Layer Perceptrons (MLP). After the development of the ANN system with exploration data follows the estimation of grade in unknown locations. Estimation is usually performed on the basis of a block or grid model. Generally, the aims of an ANN system applied to grade estimation are the following:

1. fast and reliable grade estimation,
2. minimising the required assumptions on grade distribution,
3. minimising expert knowledge requirements, and
4. making the quality of the estimates independent of the skills and knowledge of the person doing the estimation.

For the purposes of this study we will concentrate on one type of ANN, the RBF Networks, as they were found to be outperforming other architectures when applied to the problem of grade estimation (Kapageridis, 1999).

2. Radial basis function networks in grade estimation

2.1. General

RBF networks consist of three layers (input, hidden and output layer) that are fully interconnected (Fig. 1). The input layer connects the RBF network to the input vector space. The unique hidden layer performs a non-linear transformation between the input and hidden space. In grade estimation, as in most RBF network applications, the hidden space is multidimensional. The output layer is linear, providing the network's response to the presented vectors (input signals) at the input layer. These input vectors need to be normalised to ensure proper operation of the RBF network.

Each processing unit of the hidden layer has a non-linear function that forms an arbitrary 'basis' for the input vectors as they are projected to the hidden layer space. These functions are called *radial basis* functions. The hidden layer processing units have a limited receptive field, i.e. receive input vectors only from a

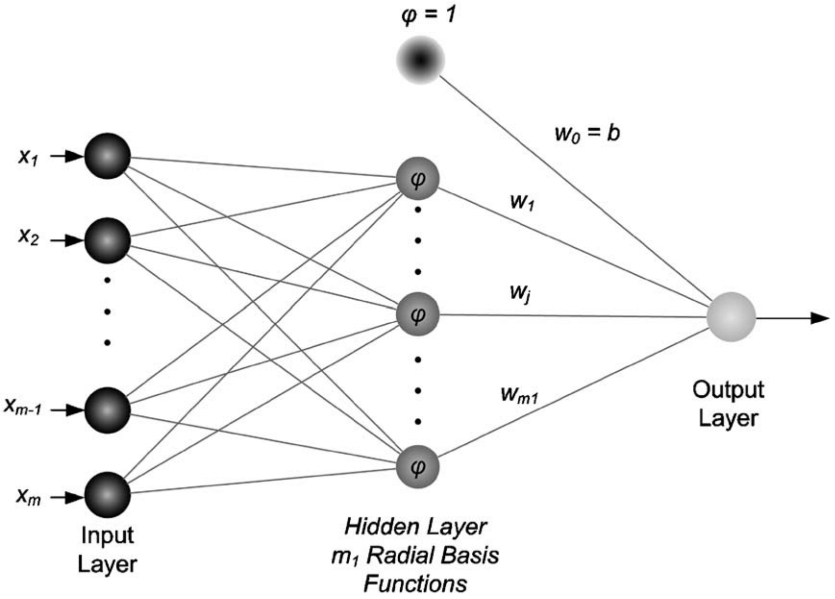

Fig. 1. RBF network architecture.

limited part of the input space. This characteristic of RBF networks makes them a favourable architecture for grade estimation. In general, the operation of an RBF network is as follows:

- All processing elements of the hidden layer receive the n-dimensional input vector after it is normalised.
- The centre of the radial non-linear basis function φ_j is positioned on the vector of weight μ_j of each hidden unit that also has an adjustable receptive field σ_j.
- The output of hidden unit j, h_j, is given as a radial function of the distance between the input vector and the unit's weight vector

  $$h_j = \varphi_j(||x - \mu_j||/\sigma_j) \tag{1}$$

  (Haykin, 1999).
- The network output is the product of output vectors of the hidden layer and the weight vector of each unit k, λ_k

  $$y_k(x) = \sum h_j(x)\lambda_k. \tag{2}$$

During training, the RBF network locates the basis functions on random samples or uses *clustering* for the choice of these locations. It is also quite common to select these locations using unsupervised learning. In this case, the hidden layer during function location acts as a Kohonen layer (self-organised). Training proceeds with the adjustment of the functions' receptive fields in order to minimise the output error. The overall network performance is dependent on:

- the choice of radial basis function,
- the number of functions,
- their centre location in the input vector space, and
- their receptive field.

RBF networks have been successfully used for function approximation (Broomhead and Lowe, 1988; Moody and Darken, 1989; Girosi and Poggio, 1990; Poggio and Girosi, 1990; Park and Sandberg, 1993). The model of these networks has been inspired by the characteristics of many biological neural network systems as well as studies on interpolation with radial basis functions (Powell, 1987) (Table 1).

2.2. Input dimensionality

A very critical issue in the use of RBF networks as function approximators is the dimension of the input space and its effect on the intrinsic complexity of the approximating function(s). It is generally accepted that this complexity increases with the input dimensionality. The space of approximating functions attainable with

Table 1
Results summary from ANN grade estimation case studies showing available data (training/validation/testing), number of RBF functions and estimation error for different input space configurations

Input space configuration	Coordinate space			Triangulation/tetrahedral model			Quadrant sectors			Octant sectors/3D tetrahedral sectors		
Study	Data	RBF	%	Data	RBF	%	Data	RBF	%	Data	RBF	%
Jura (Cd)	180/90/89	31	24.4	165/83/82	81	39.6	161/80/80	9	43.3	130/65/65	126	49.2
Copper-David	200/100/100	90	17.01	180/90/90	99	17.76	162/81/81	23	12.71	162/81/81	136	14.32
Iron ore	37/23/31	18	9.49	31/19/27	26	10.44	30/20/20	21	6.93	24/20/10	8	10.01
Potash (KCl)	1429/715/714	100	24.7	1397/699/698	15	24.96	1398/699/698	29	23.02	1293/647/646	19	24.52
Phosphate	948/474/473	25	17.75	828/414/414	138	17.76				697/349/348	147	17.31
Zinc	3328/1664/1664	197	40.46	3304/1652/1652	126	33.56				2787/1394/1393	135	37.19
Copper (3D)	8327/4164/4163	146	36.4	7500/3500/3500	123	28				7350/3240/3240	119	24.5
Marl (CaO)	1947/973/325	143	7.83	1222/611/610	54	3.77				552/276/276	61	3.48

RBF networks becomes increasingly constrained as the input dimensionality is increased (Haykin, 1999).

Increased dimensionality also has a great effect on the computational overhead caused during training of the RBF network. The dimension of the input space has a direct control over the RBF network architecture—the number of input nodes, the number of required RBFs, and consequently, the number of linear weights between hidden and output layer. Therefore, any increase in the input dimensionality causes an increase in computer memory and processing power requirements, and an almost certain increase in development time due to the increased number of unknown network parameters that have to be fixed during training. One of the most common ways of addressing high input dimensionality for a given problem is to identify and ignore inputs that do not contribute considerably to the output or to try to combine inputs that present a high correlation. Another way of reducing input dimensionality, which is not always applicable though, is to try and break a complex problem into a number of lower dimensionality problems that can be more effectively addressed using RBF networks.

3. Input space configurations

Grade and reserve estimation usually involves the prediction of various parameters that characterise an ore deposit. Input data normally come in the form of samples at known locations in three-dimensional (3D) space. The majority of ANN systems developed for such predictive tasks are based on the relationship between the predicted parameters and sample locations. The most common practice during development of training datasets for an ANN is to produce pairs of input–output, the input being the sample location and the output being the required value of the predicted parameter. In other words, most of the ANN systems treat the unknown parameters estimation as a problem of function approximation in the samples coordinate space. Based on a number of examples in the literature and an earlier research programme (Kapageridis, 1999), the following ANN input spaces are considered.

3.1. For 2D samples

In many cases the available samples are unique along the Z coordinate axis, as the parameter to be estimated is considered to be non-varying along this axis. For example, in a stratigraphic potash deposit, the KCl grade shows small variation on the Z-axis. In such cases, samples are usually located in space according to the seam that they come from and their X–Y coordinates. Estimation is consequently performed on a grid model basis. The following input spaces can be used with two-dimensional (2D) samples for grid model based estimation.

3.1.1. X–Y coordinate space

This is probably the simplest space where the sample coordinates are taken as the only input parameters to the ANN while the estimation parameter value is taken as the only output. The ANN is, in this case, trying to create a projection for the X–Y coordinate space to the estimation parameter space. The low dimensionality of this problem leads inevitably to architectures with many hidden units and sometimes in multiple hidden levels. In the case of an RBF network this means a large number of radial basis functions.

In order to achieve the required projection, the ANN must have the appropriate number of weights between the input and hidden level. It is therefore necessary, given the low number of input nodes (2) to have sufficient number of hidden units to form the necessary number of weights. The number of required hidden units is initially unknown and can only be found using various architecture optimisation techniques such as Genetic Algorithms. As this is the simplest configuration, it is also the most common and there are plenty of examples of this approach in the literature.

3.1.2. Triangulation neighbour samples space

Based on the triangulation method of interpolation, the triangulation neighbour samples space is a very simple approach to the presentation of samples to an ANN system for grade estimation. The triangle method has been used in the past as a simple geometrical method for grade estimation prior to the extended use of computers that we experience nowadays. According to this method, three samples are selected that surround the point of estimation and form a triangle as shown in Fig. 2A. Each point within the triangle is considered to have a grade equal to the average grade of the samples that form the triangle.

In the case of using such an input space for the training of ANN systems, the grades of the samples that form the triangle are taken as the input values while the grade of the training sample is taken as the only output. The triangles are formed excluding the training sample from the triangulation procedure, which is based on the Delaunay algorithm. Each training sample is contained in a unique triangle that forms the input space. An interesting variation of this input space is derived by taking the distances of the neighbour samples as inputs to the ANN. The low dimensionality of this space leads to ANNs with many hidden units, especially when the approximated grade surface is complex (Table 2).

3.1.3. Quadrant/octant search neighbour samples space

In this configuration, the space around the estimation point is divided to four or eight sectors with a 90° or 45°

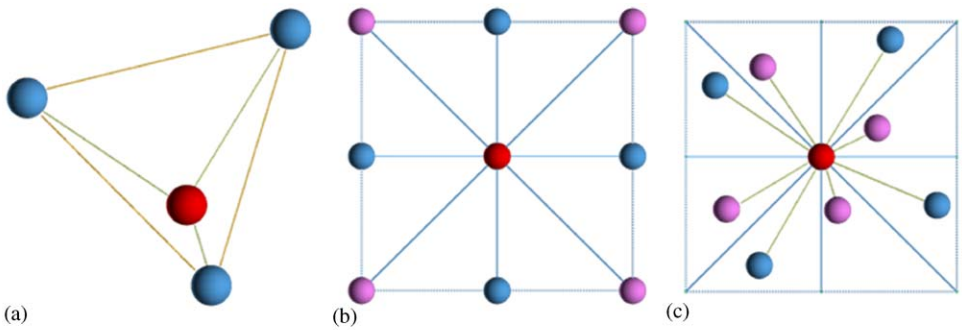

Fig. 2. (a) Triangulation neighbour samples space configuration. (b) Quadrant/octant search neighbour samples space configuration (fixed grid). (c) Quadrant/octant search neighbour samples space configuration (flexible).

Table 2
Characteristics of considered input space configurations relative to the number of coordinates of the original samples

Input space configuration	Number of dimensions	Dimensions	Samples coordinate space dimensions
Coordinates *X*–*Y*	2	Coordinates *X*–*Y*	2
Coordinates *X*–*Y*–*Z*	3	Coordinates *X*, *Y* and *Z*	3
Coordinates *X*–*Y*–*Z* and volume	4	Coordinates *X*, *Y*, *Z* and sample volume	3
Triangulation neighbour samples	6	Grade and distance of triangulation neighbour samples	2
Tetrahedral model neighbour samples	8	Grade and distance of tetrahedral neighbour samples	3
Quadrant sectors samples (single ANN)	8	Grade and distance of all quadrant sectors samples	2
Quadrant/octant sectors samples (multiple ANNs)	2	Grade and distance of one quadrant or octant sector sample	2
Octant sectors samples (single ANN)	16	Grade and distance of all octant sectors samples	2
3D Tetrahedral sectors samples (single ANN)	3–144	Grade, distance and volume of all tetrahedral sectors samples	3
3D Tetrahedral sectors samples (multiple ANNs)	3	Grade, distance and volume of one tetrahedral sector sample	3

angle, respectively. The nearest neighbour sample is selected from each sector (Fig. 2B). Depending on the sampling scheme, samples can be arranged on a regular grid or on random locations. In the case of a regular grid, only the grades of the four or eight neighbour samples are taken as ANN inputs. It is worth noticing that when eight samples are selected, the four samples on the diagonals are at a greater distance to the estimation point from the other four.

As expected by the fixed position of the neighbour samples, results using this configuration are usually very good. However, the area of application for this configuration is very limited as grade estimation is very rarely performed using samples on such geometrically specific locations. In most cases, samples are located in arbitrary positions, and even though they can still be located inside angular sectors, their distance from the estimation point is non-constant. Therefore, it is necessary to use this distance as yet another input to the network (Fig. 2C).

This input space configuration gives more freedom and can be used almost with any sampling scheme. However, this comes at a price as it leads to a space with up to 16 dimensions (eight samples = eight grades + eight distances = 16 dimensions), which also increases the minimum quantity of data required to successfully develop the ANN. The architectures shown in Fig. 3 require a sample in all sectors. If one of the sectors is empty, this will make the complete presentation of data to the ANN impossible. This can happen due to the sampling scheme or when we consider estimation points close to the sampling area limits.

3.2. For 3D samples

The addition of one more coordinate for the location of samples in 3D space leads to more complex input space configurations and network architectures. Additionally, approaching grade estimation in three dimensions often leads to the need for treating samples as volumes and not points. Therefore, sample volume and estimation volume become parameters in the 3D space that have to be included in the development and application process of the ANN system. Estimation is performed on block models and not grid models. Grid models are not capable of dealing with the concept of estimation volume in contrast with block models which are built to represent volumes in 3D space. The following input space configurations are considered for ANN-based grade estimation with block models.

3.2.1. *X–Y–Z coordinate space*

This space is an extension to the *X–Y* coordinate space in three dimensions. The addition of the third coordinate *Z* to the ANN input enables its development with data from deposits with grade varying in all three dimensions due to the inherent complexity of the deposition or any structural deformations (faults, folds, etc.). Depending on whether samples and estimates are considered as points or volumes, it is possible to add the sample volume as an extra input to the ANN, increasing the number of input space dimensions to four (Fig. 4). This allows the use of samples from different sampling methods and with different support for ANN development without the need for sample compositing to constant volume. This way the ANN learns the relationship between volume and grade as well as the relationship between sample location and grade.

3.2.2. *Tetrahedral neighbour samples space*

This is probably the most uncommon case of an input space and one that has never been used before in a similar application. It is based on the construction of tetrahedra from the available samples. The concept is very similar to the triangulation neighbour samples space discussed before and is, in essence, an extension of the Delaunay triangulation algorithm in three dimensions for the construction of a tetrahedral model.

All volumetric geometries can be represented in 3D using a set of tetrahedra. Tetrahedral modelling uses the tetrahedron as the basic unit for representing volumetric geometry. A tetrahedral model is composed of indexed 3D tetrahedra, in contrast to the set of connected flat triangles forming a standard triangulation.

Each training or estimation point is contained in a tetrahedron built from surrounding samples (Fig. 5). The grades of these samples, their distance from the training/estimation point and their volume are used as inputs to the ANN. The non-fixed orientation of each tetrahedron leads to the construction of an ANN model between the four sets of input values (grade, distance and volume) and the training/estimation point grade that does not consider the location of the samples and of the training/estimation point. This model also does not consider the direction of the vector between these points and therefore leads to the assumption that grade variance is independent of location and direction in space.

3.2.3. *3D sectors neighbour samples space*

Extending the quadrant/octant space to three dimensions leads to the configuration shown in Fig. 6. The space around each estimation point is divided into tetrahedral sectors. The closest sample from each sector

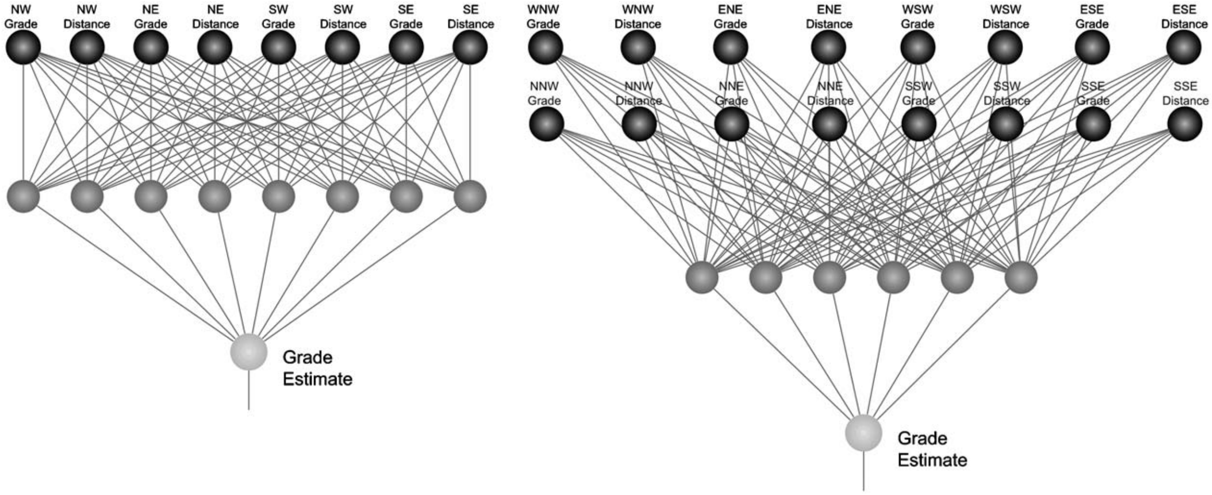

Fig. 3. Possible ANN architectures with four (left) or eight (right) input samples. Neighbour sample grade and distance are taken as inputs.

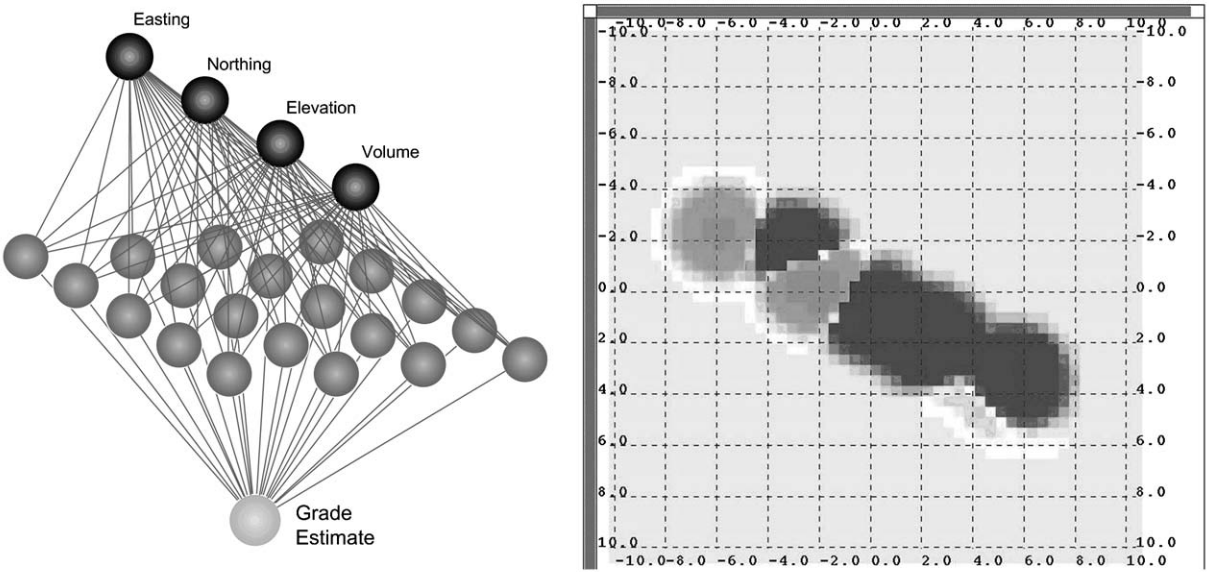

Fig. 4. Left: Possible ANN architecture for grade modelling in sample coordinates space. Right: output map example from ANN trained in *X*–*Y*–*Z* coordinate sample space.

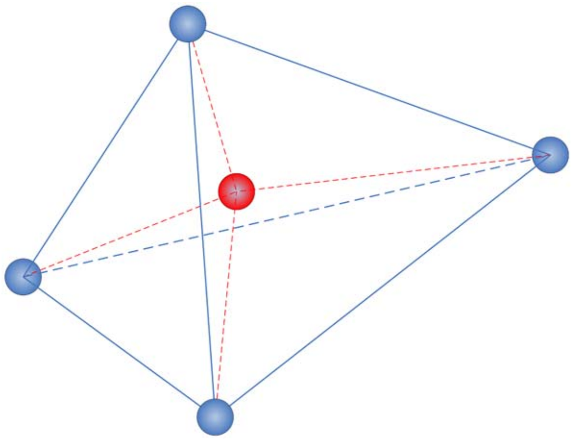

Fig. 5. Estimation point (centre) and selected neighbour samples (corners) using a tetrahedral model built from available samples.

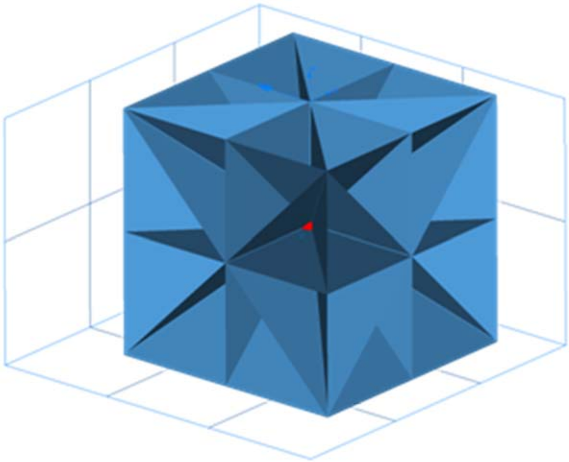

Fig. 6. 3D space breakdown into tetrahedral sectors around an estimation point.

is selected and its grade, distance from the estimation point and volume used as inputs to the ANN system. In very rare cases there will be enough samples in all directions to fill all 48 sectors formed. This means that prior to the ANN development it is necessary to decide which sectors to use or discard depending on the sampling scheme and some analysis of sample pair directions.

This configuration requires the use of more than one ANN due to the high dimensionality. If we assume the extreme case where there are enough samples to fill all 48 sectors, each one giving 3 inputs, this leads to a total of 144 inputs. In such a case, tens of thousands of samples would be required to properly train the ANN. Therefore, it is necessary to break down the problem to smaller ones that can be more efficiently handled by a single ANN and a common number of samples.

The simplest way of such breakdown is to use one ANN for each sector. Depending on the directional analysis of the available sample pairs, the number of required ANNs will equal the number of sectors with adequate sample pairs. Thus, each ANN is trained on

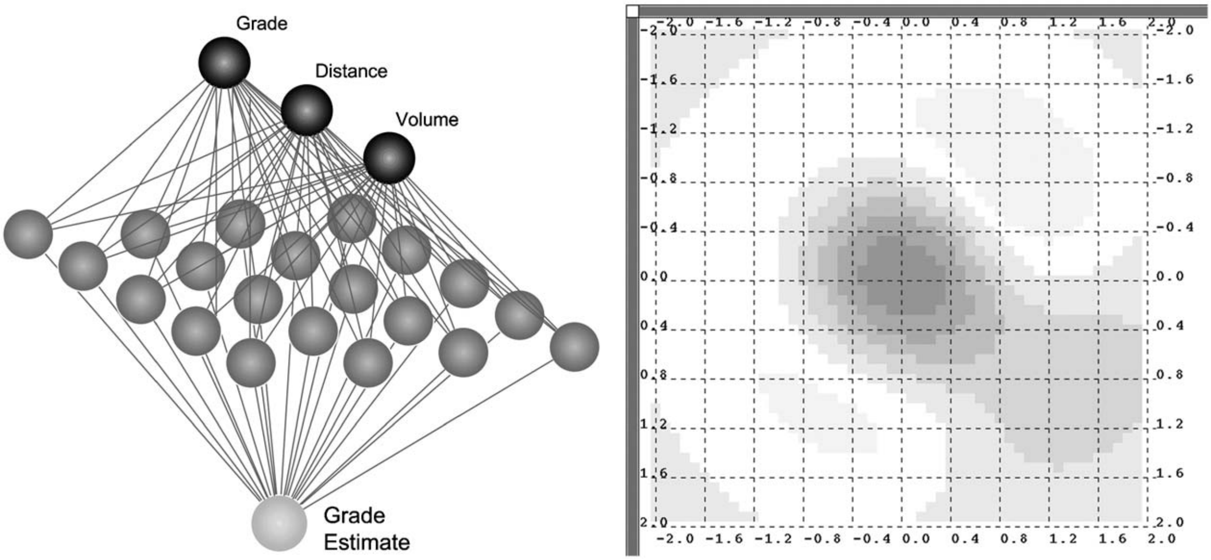

Fig. 7. Left: possible ANN architecture for grade modelling based on neighbour samples. Right: output map from network trained using neighbour sample grade and distance as inputs.

sample pairs that lie approximately in the same direction. In essence, each ANN constructs a model of grade variance for a particular direction. An example of such a network is shown in Fig. 7. This network has only three inputs and thus requires a much lower number of samples to train.

The use of multiple ANNs, one for each sector, leads to multiple estimates for each estimation point (their number depending on the estimation point location relative to the sampling area). These estimates need to be averaged to a final estimate for each estimation point. One way to achieve this is using another ANN that will take as inputs the individual estimates and will produce the final estimate as an output (Kapageridis, 1999).

4. Case studies

The case studies were split into two groups based on the number of coordinate dimensions of the original samples: 2D and 3D. The 2D samples were considered as points in space with a grade value assigned to them and estimation was performed on a grid model basis, i.e. point estimates. In the case of 3D samples, their volume was also considered, while the estimates were produced on a block model basis. Eight case studies were performed, four from each group, and their results are summarised at the end of this paper. Two of these case studies are discussed briefly in the following paragraphs.

4.1. Example study on 2D data—potash (KCl)

The data used in this study come from a large stratigraphic potash deposit in the United Kingdom. A total of 1429 samples from underground drillholes were split into three groups, 50% for training, 25% for validation and 25% for testing. Training data are used to develop the networks and fix the free parameters of their architecture such as RBF centre locations and the weights between hidden and output layers. The validation set is not used for changing any of the free parameters but to find the set of parameters that generates the best possible validation error and ensure good generalisation levels. Testing data are kept completely out of the network development process and are used to calculate the various estimation error measures after network development is complete.

4.1.1. Test in the X–Y coordinate space

Initially an ANN was developed in the *X*–*Y* coordinate space. The mean absolute error of estimation was 24.7% using this input space configuration. The best validation error was achieved with 100 RBF units. As shown in the scatter diagram of the estimates (Fig. 8), this ANN tends to produce estimates close to the average, slightly overestimating low-grade samples and underestimating high grade ones.

4.1.2. Test in the triangulation neighbour sample space

The triangulation procedure produced 1397 training samples, 699 validation samples and 698 estimation testing samples. The best validation error was achieved

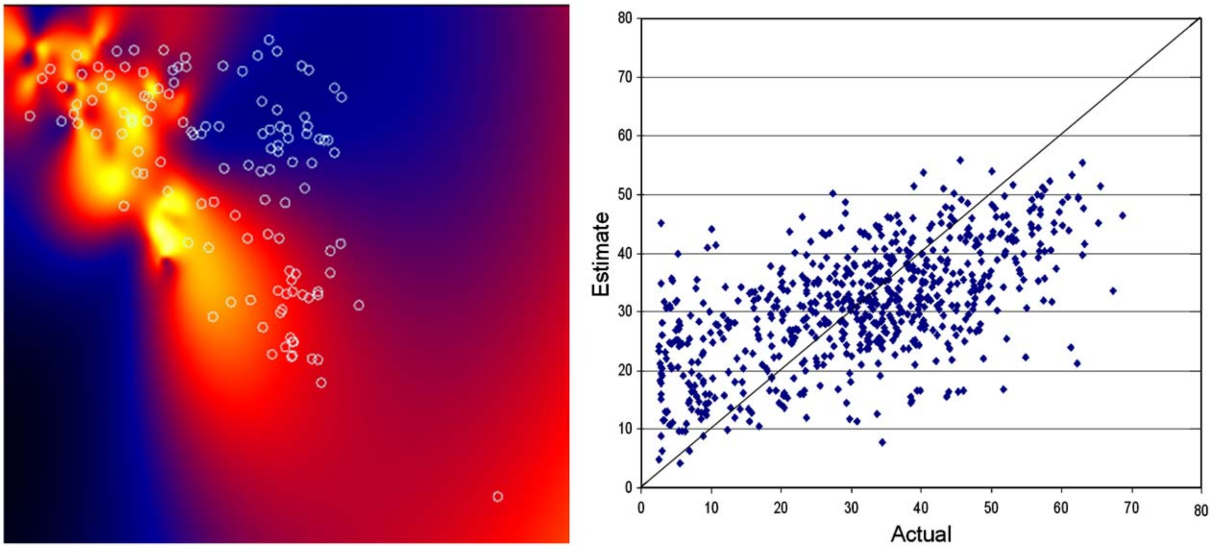

Fig. 8. Left: ANN output map and RBF locations after training using potash (XY). Right: scatter diagram of potash estimates (XY).

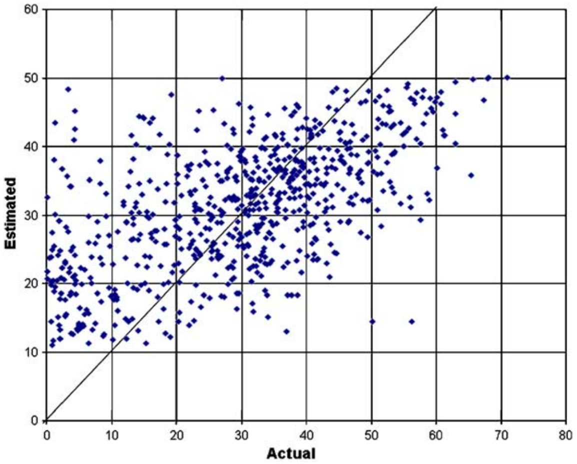

Fig. 9. Scatter diagram of KCl estimates in triangulation neighbour samples space.

with just 15 RBF units. The mean absolute estimation error was 24.96%, similar to the previous test. The scatter diagram of the estimates produced with the testing samples is shown in Fig. 9.

4.1.3. *Test in the quadrant/octant search neighbour samples space*

The potash data produced 1398/699/698 samples with the quadrant search and 1293/647/646 samples with the octant for training, validation and testing, respectively. The best validation error was achieved with 29 RBF units (0.78 RMS) using the quadrant data and with 19 (0.81 RMS) using the octant. The mean absolute estimation error was 23.02% and 24.52%, respectively. The estimation was slightly more balanced in the case of octant data as shown in the scatter diagrams below (Fig. 10).

4.2. *Example study on 3D data—marl*

The marl deposit considered in this study is located near Skopje (Former Yugoslavian Republic of Macedonia). The available samples come from 77 drillholes providing a total of 3245 samples—1947 are used for training, 973 for validation and 325 for testing. All samples were of equal length and therefore sample volume was not included in the input space configurations.

4.2.1. *Test in the X–Y–Z coordinate space*

The available data were used for the development of an ANN with three inputs (X–Y–Z). The best validation error (0.87 RMS) was achieved with 143 RBF units. The mean absolute estimation error was 7.83%. Fig. 11 shows a scatter diagram of the CaO estimates.

4.2.2. *Test in the tetrahedral neighbour samples space*

The tetrahedral model produced 1222 samples for training, 611 samples for validation and 610 samples for testing. The best validation error (0.65 RMS) was achieved with 54 RBF units. The mean absolute estimation error was 3.77%, significantly improved from the previous test, as can be seen in Fig. 12.

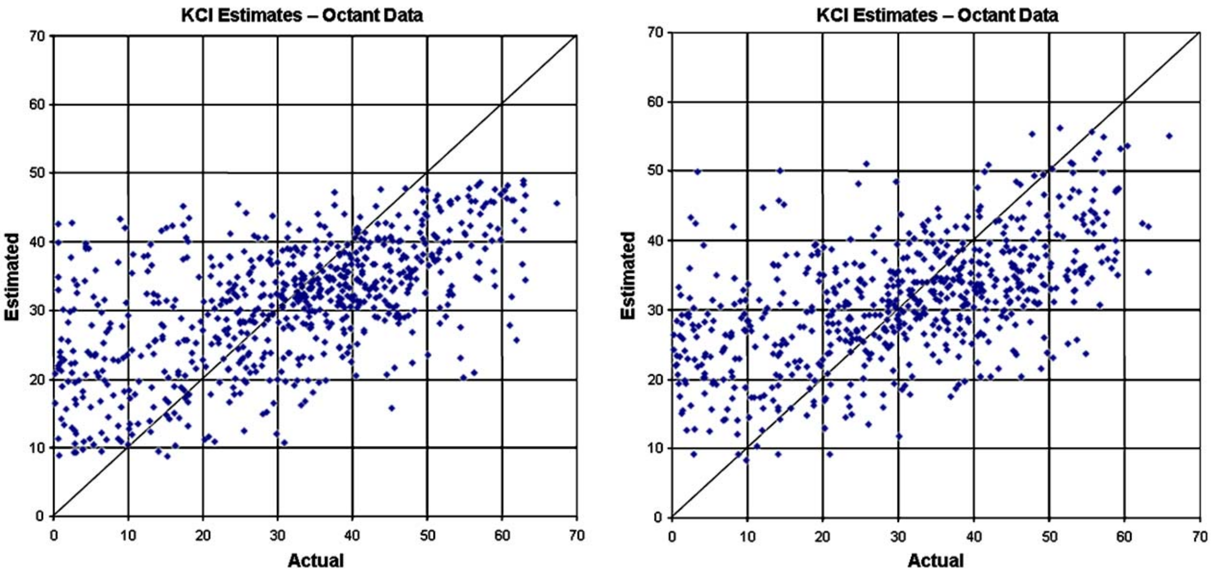

Fig. 10. Scatter diagrams of KCl estimation using data from quadrant and octant search.

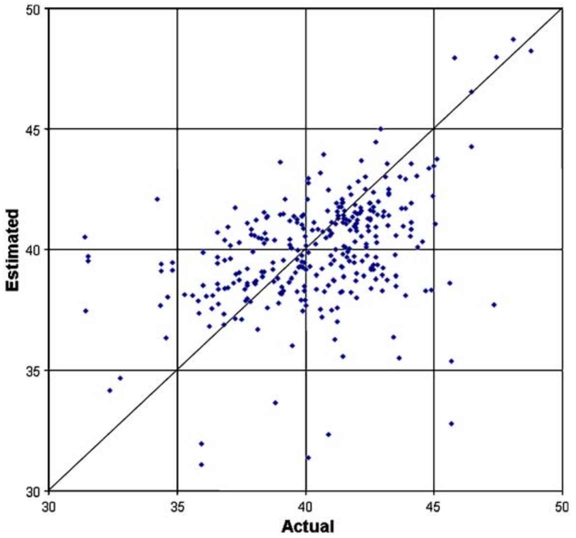

Fig. 11. Scatter diagram of CaO estimates in *X–Y–Z* coordinate space.

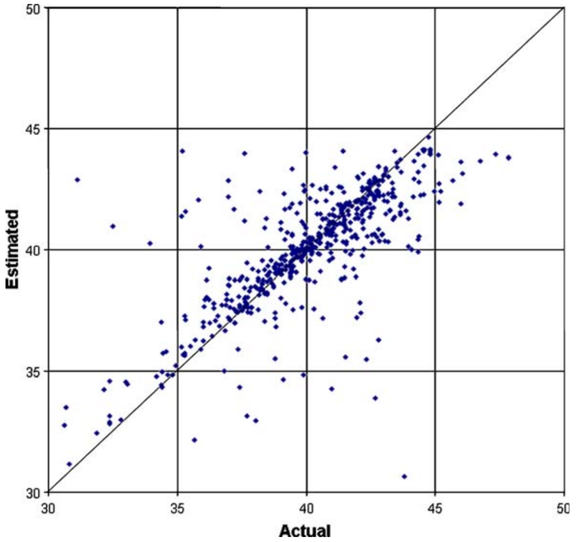

Fig. 12. Scatter diagram of CaO estimates in tetrahedral neighbour samples space.

4.2.3. Test in the 3D sectors neighbour samples space

The final test in this study was performed with 552 training samples, 276 validation samples and 276 estimation testing samples in the 3D sectors neighbour samples space. The fully developed ANN produced a validation error of 1.11 RMS with 61 RBF units. The mean absolute estimation error was further reduced to 3.48%. The quality of the produced estimates is shown by the scatter diagram in Fig. 13.

5. Summary of results—conclusions

The main purpose of this study was the investigation of the dimensionality effects on the performance of ANN systems when applied to grade estimation from exploration data. Dimensionality was defined as the number and type of ANN input parameters. Their application was defined as a three-stage process: training, validation and estimation testing. Accordingly, the effects of dimensionality to these development and application stages were examined for the particular problem of grade estimation.

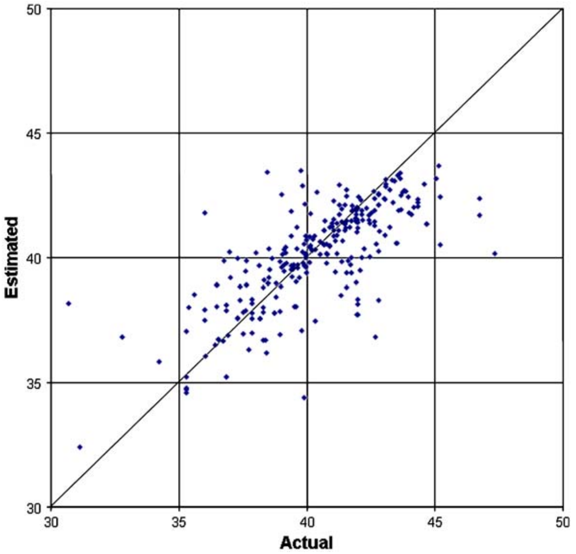

Fig. 13. Scatter diagram of CaO estimates in 3D sectors neighbour samples space.

The possible effects were investigated through a number of tests on different deposits with varying number of samples and different sampling schemes and qualities of assays. The conclusions given in this section are based on more than 24 case studies summarised in the following table. The measures used to investigate the effects of each input space configuration were the following:

- The validation error of a fully developed ANN, i.e. the error produced by the ANN on the samples used during training to guide the adjustment of function locations, the number of basis functions and their receptive field.
- The number of RBF functions required in every case to achieve the most accurate possible approach of the validation samples. The optimum number of RBF functions was found during development by increasing the number of functions from a minimum of 5 to a maximum of 200 and repeating training and validation of the network.
- The estimation error of the ANN on the testing samples that are not involved in the development of the ANN. This was also presented graphically through a scatter diagram of the estimated grades versus the actual testing sample grades.

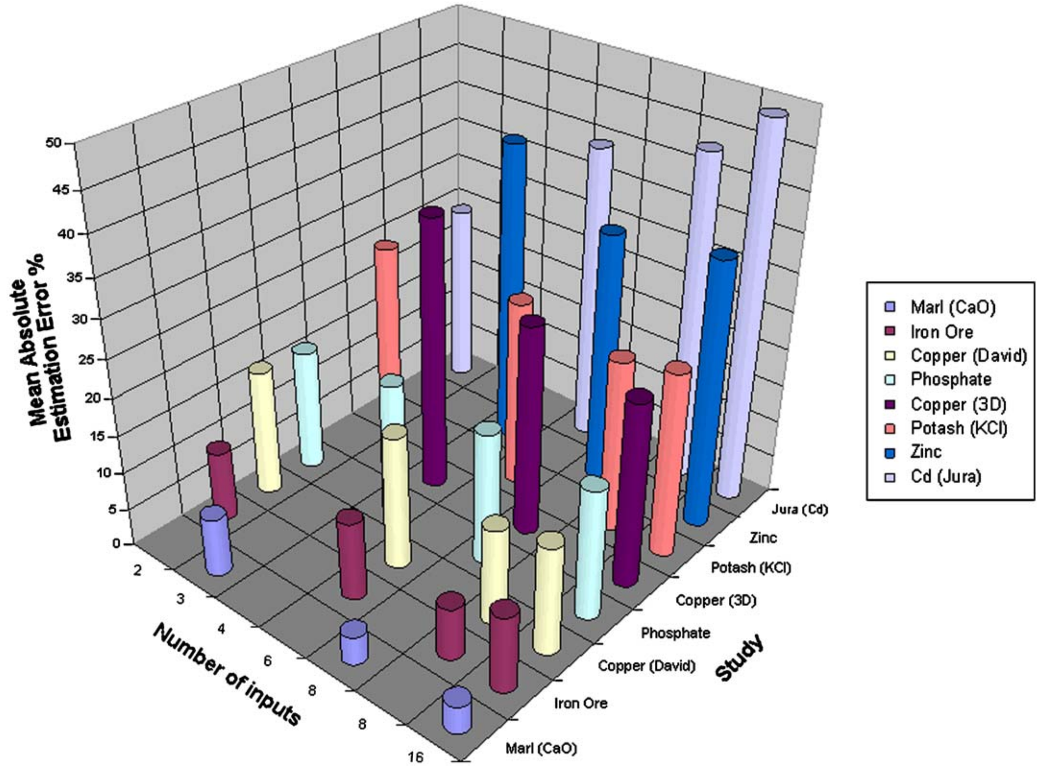

Fig. 14. Graph of mean absolute estimation error versus number of inputs.

Fig. 14 shows the estimation error in each study relative to the number of input spaces.

Generally, a decrease in the estimation error is observed as the number of input space dimensions increases. This is more obvious in cases where the ANN gives better results, while the exact opposite is observed in cases were the error is particularly high. The decrease in the estimation error is also more obvious in those studies where 3D data are used. The following diagram (Fig. 15) shows the number of RBF functions relative to the number of input space dimensions for each study.

The number of RBF functions provides less consistent results. In the Iron Ore, Potash, Marl and Copper (3D) studies the number of required functions decreases as the number of input space dimensions increases. This trend follows the logic that the more inputs we have in an ANN the more free parameters (weights) we have to adjust during training and the less number of processing elements we require in the ANN (in this case RBF functions). However, in the Jura (Cd), Phosphate and Copper (David) the exact opposite is observed: the number of RBFs increases with the input space dimensions. In these studies the estimation error was very high and in some cases constant regardless of the input space used. In the studies with low quantity of data (Jura, Copper-David, and Iron Ore) overfitting is evident in the results and could possibly be avoided with more data.

Examining the performance of the ANN systems in all input space configurations leads to the conclusion that there is no globally applicable configuration for all deposit and sampling scheme types. Each deposit and sampling scheme must be considered separately to find the best configuration applicable. The following figure (Fig. 16) gives an idea of the difficulty in the development of ANN systems in the considered input space configurations through the validation error in each study.

The dimensionality of the input space can be a problem in certain cases where the required mapping is particularly complex and the quantity of available data is low. Low quantity of data combined with a high number of inputs can also lead to bad estimation results as there is no sufficient information to effectively train all of the network's free parameters. In other cases, a dimensionality increase combined with a sufficient quantity of data leads to improvement of the estimates. The above conclusions must be considered relative to the

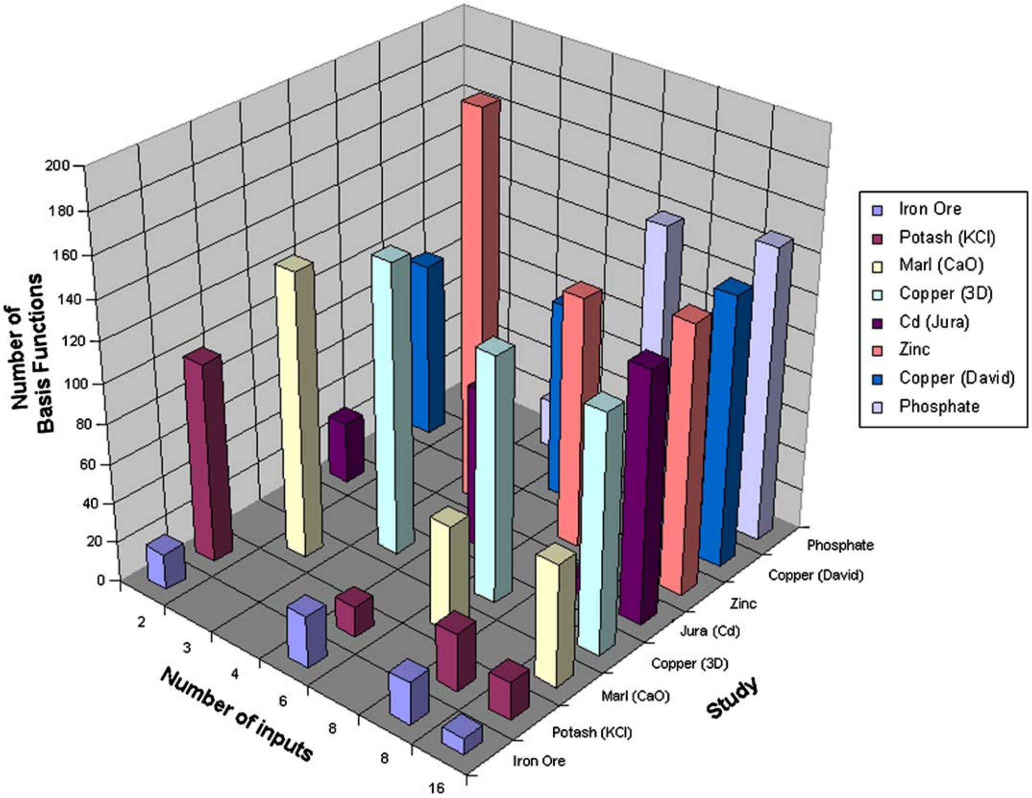

Fig. 15. Graph of number of RBF functions versus number of inputs.

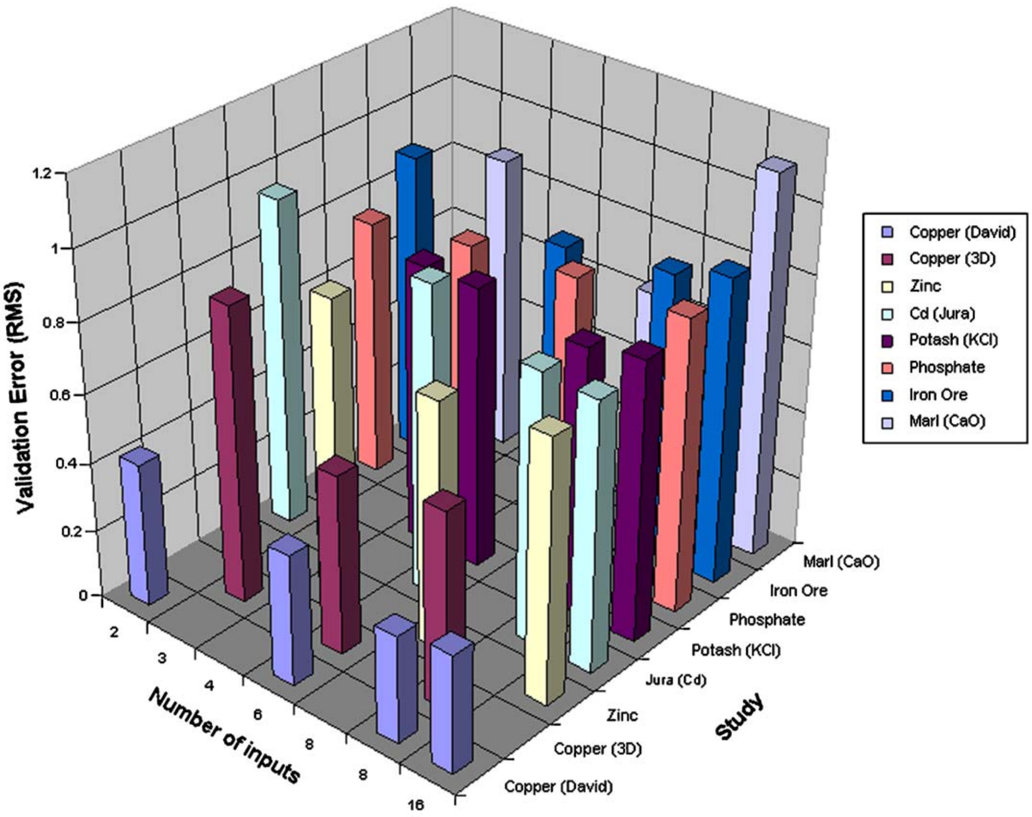

Fig. 16. Graph of validation error (RMS) versus number of inputs.

chosen ANN architecture, the learning algorithm used and the particular deposits examined.

References

Broomhead, D., Lowe, D., 1988. Multivariable functional interpolation and adaptive networks. Complex Systems 2, 321–355.

Burnett, C.C.H., 1995. Application of neural networks to mineral reserve estimation. Ph.D. Thesis, Department of Mineral Resources Engineering, University of Nottingham, Nottingham, 254pp.

Caiti, A., Parisini, T., 1991. Mapping of ocean sediments by networks of parallel interpolating units. In: Proceedings of the IEEE Conference on Neural Networks for Ocean Engineering, Washington, DC, USA, pp. 231–238.

Clarici, E., Owen, D., Durucan, S., Ravencroft, P., 1993. Recoverable reserve estimation using a neural network. In: Proceedings of the 24th International Symposium on the Application of Computers and Operations Research in the Minerals Industries (APCOM), Montreal, Quebec, pp. 145–152.

Clark, I., 1979. Practical Geostatistics. Elsevier, Amsterdam 129pp.

Cortez, L.P., Sousa, A.J., Durao, F.O., 1998. Mineral resources estimation using neural networks and geostatistical techniques. In: 27th Internation Symposium on Computer Applications in the Mineral Industries (APCOM 98). The Institution of Mining and Metallurgy, pp. 305–314.

David, M., 1977. Geostatistical Ore Reserve Estimation. Elsevier, Amsterdam 364pp.

Girosi, F., Poggio, T., 1990. Networks and the best approximation property. Biological Cybernetics 63, 169–176.

Haykin, S., 1999. Neural Networks–A Comprehensive Foundation. Prentice-Hall, New Jersey 842pp.

Isaaks, E.H., Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, New York 580pp.

Journel, A.G., Huijbregts, Ch.J., 1978. Mining Geostatistics. Academic Press, London 600pp.

Kapageridis, I.K., 1999. Application of artificial neural network systems to grade estimation from exploration data. Ph.D. Thesis, School of Chemical, Environmental and Mining Engineering, University of Nottingham, Nottingham, 260pp.

Kapageridis, I., Denby, B., 1998a. Ore grade estimation with modular neural network systems—a case study. In: Panagiotou, G. (Ed.), Proceedings of the First International Conference on Information Technology in the Minerals Industry (MineIT '97). AA Balkema, Rotterdam CDROM Paper Code: KI16.

Kapageridis, I., Denby, B., 1998b. Neural network modeling of ore grade spatial distribution. In: International Conference on Artificial Neural Networks 98 (ICANN 98). Springer, Berlin, pp. 209–214.

Krige, D.G., 1981. Log-normal–de Wijsian Geostatistics for Ore Evaluation. South African Institute of Mining and Metallurgy, Johannesburg 51pp.

Moody, J., Darken, C., 1989. Fast learning in networks of locally tuned processing units. Neural Computation 1, 281–294.

Park, J., Sandberg, I., 1993. Approximation and radial-basis-function networks. Neural Computation 5, 305–316.

Poggio, T., Girosi, F., 1990. Networks for approximation and learning. In: Proceedings of the IEEE, vol. 78, no. 9, IEEE Press, New York, pp. 1481–1497.

Powell, M., 1987. Radial basis functions for multivariate interpolation: a review. In: Mason, J., Cox, M. (Eds.), Algorithms for the Approximation of Functions and Data. Clarendon Press, Oxford, pp. 143–167.

Wu, X., Zhou, Y., 1993. Reserve estimation using neural network techniques. Computers & Geosciences 19 (4), 567–575.

Yama, B., Lineberry, G., 1999. Artificial neural network application for a predictive task in mining. Mining Engineering 2/99, 59–64.