IOP PUBLISHING JOURNAL OF GEOPHYSICS AND ENGINEERING

J. Geophys. Eng. **6** (2009) 29–34 doi:10.1088/1742-2132/6/1/003

Inversion of self-potential anomalies caused by 2D inclined sheets using neural networks

Hesham M El-Kaliouby and Mansour A Al-Garni

Department of Geophysics, King Abdulaziz University, PO 80206, Jeddah 21589, Saudi Arabia

E-mail: hkaliouby@yahoo.com and helkaliouby@kau.edu.sa

Received 12 July 2008
Accepted for publication 5 January 2009
Published 28 January 2009
Online at stacks.iop.org/JGE/6/29

Abstract

The modular neural network (MNN) inversion method has been used for inversion of self-potential (SP) data anomalies caused by 2D inclined sheets of infinite horizontal extent. The analysed parameters are the depth (h), the half-width (a), the inclination (α), the zero distance from the origin (x_o) and the polarization amplitude (k). The MNN inversion has been first tested on a synthetic example and then applied to two field examples from the Surda area of Rakha mines, India, and Kalava fault zone, India. The effect of random noise has been studied, and the technique showed satisfactory results. The inversion results show good agreement with the measured field data compared with other inversion techniques in use.

Keywords: self-potential, neural networks, SP, inclined sheet, 2D, inversion

Introduction

The self-potential (SP) method is based on measurements of natural electric fields at the Earth's surface or in boreholes (Corwin 1990, Mendonça 2008) generated by coupled-flux processes taking place in the shallow part of the subsurface (Sill 1983). A gradient in a primary potential of pressure, temperature or concentration potentials drives a primary flux of matter, heat or charge. This flux leads to charge separation which is balanced by a counteracting (coupled) flux of current whose divergence may generate a measurable SP anomaly (Mendonça 2008). The diversity of mechanisms which generate the primary flux, provides wide applications in engineering problems, groundwater investigations, subsurface temperature distributions and mineral investigations related to sulfides and graphite (Sundararajan *et al* 1998, Mendonça 2008). In mineral exploration, natural gradients in the ground redox potential generate SP anomalies that have helped locate sulfide ore bodies. Main factors creating gradients in the redox field are rainwater infiltration and oxygen diffusion from the atmosphere that generate an oxidizing environment in the shallow subsurface, in contrast to the reducing conditions found at depth. A conductive material connecting such contrasting shallow and deep regions conveys an upward flow of electrons. The upper portion of the conductor releases electrons and works as a battery cathode pole. Its lower portion retrieves electrons from the surrounding medium and works as a battery anode. The corresponding flow of groundwater ions in the earth surrounding the deposit generates the observed SP anomaly (Sato and Mooney 1960, Mendonça 2008).

Several methods available for the interpretation of SP anomalies generally assume the causative bodies to be of regular geometrical shapes which can be described with appropriate analytical formulae. The method of characteristic points (Paul 1965, Paul *et al* 1965, Rao *et al* 1970), logarithmic-curve matching (Meiser 1962, Murthy and Haricharan 1984) and the method of nomograms (Bhattacharya and Roy 1981, Murthy and Haricharan 1985) all involve many approximations. A least-squares approach has been used to estimate the depth of a causative target from residual SP anomaly caused by bodies of simple geometries (Abdelrahman and Sharafeldin 1997). This method requires a series of trials to minimize the error between the observed and calculated values. A few methods have been developed to determine the shape of a SP anomaly using the least-squares method (Abdelrahman *et al* 1997a) and the numerical gradient method (Abdelrahman *et al* 1997b). The derivative analysis method (Abdelrahman *et al* 1998) in addition to extended derivative analysis to higher derivatives is used to estimate not only the shape factor but also the depth of the source of SP anomaly. The spectral analysis method is trustworthy only for very long profiles (Sundararajan *et al* 1998). While the simplicity of use and the accuracy of results differ with the specific interpretation technique, they are subject to many constraints. These methods do not yield a precise location of origin of the source of the anomaly, which is a must for meaningful interpretation.

Sundararajan *et al* (1990) utilized the Hilbert transforms for the interpretation of SP anomalies for determining the centre or origin of the body based on the use of the Hilbert transform, which avoids many of the drawbacks listed above.

Geophysical inversion involves the estimation of the parameters of a postulated earth model from a set of observations. It may be viewed as an attempt to fit the response of an idealized subsurface earth model to a finite set of actual observations. Neural networks (NNs) have been implemented successfully within many areas of geophysics including the inversion of seismic data, well logging, electromagnetic, magnetotelluric, magnetic and gravity data (e.g. Poulton 2001, 2002, Van der Baan and Jutten 2000, Bescoby *et al* 2006, El-Kaliouby and Poulton 1999, Macias *et al* 2000, Zhang *et al* 2002, Spichak and Popova 2000, and many others).

In this paper, NNs are used in inverting the SP anomalies over a sheet-like causative target.

Theory

Neural networks

NNs can be considered as a class of universal approximators that are capable of approximating any function in terms of its variables. Hence, they may yield important contributions to finding solutions to a variety of geophysical applications (Macias *et al* 2000, Poulton 2001).

NN models can be more accurate than polynomial regression models used for approximating functions, allowing more dimensions than look-up table models, and allowing multiple outputs for a single model. Models using NNs are developed by providing sufficient training data (either simulated or measured data) from which they learn the underlying input/output mapping. Several valuable characteristics are offered by NNs.

- First, no prior knowledge about the input/output mapping is required for model development. Unknown relationships are inferred from the data provided for training. Therefore, with a NN, the fitted function is represented by the network and does not have to be explicitly defined.
- Second, NNs can generalize, meaning they can respond correctly to new data that have not been used for model development.
- Third, NNs have the ability to model highly nonlinear as well as linear input/output mapping. In fact, it has been shown that NNs are capable of forming an arbitrarily close approximation to any continuous nonlinear mapping (Zhang and Gupta 2000).
- Once properly trained, the NNs can perform the inversion in nearly no time.

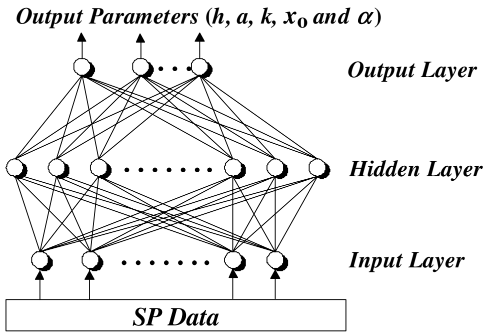

Figure 1. Typical neural network architecture showing the input and output parameters.

These capabilities make it possible for the NN to solve complex (large-scale) problems that are currently intractable.

Typically a neural network is given a training set of a group of examples from which it learns to approximate the mapping function described by the example patterns. The most commonly used training scenarios utilize *supervised* learning, during which the network is presented with input patterns together with the desired output patterns, the target output constituting the correct answer or correct classification of the input data. In *unsupervised* learning, the network is provided only with input patterns and it finds common features in groups of those patterns (Poulton 2001).

A neural network consists of a layered system of interacting nodes (figure 1); each one is a single processing element (PE) that acts on data to produce a result. Each node has also an extra input called the threshold input, which acts as a reference level or bias for the node. There is a minimum of three layers for nonlinear problems: an input layer, at least one hidden layer and an output layer. Data enter the network through the input layer; each node broadcasts a single data value over weighted connection to the hidden nodes, which process the input data and broadcast their results to the output layer. The output nodes also have distinct sets of weights and process input values to produce a result. This architecture is called feed-forward multi-layer perceptron (MLP) (figure 1).

The hidden node processes its input in two steps. First, it multiplies every input by its weight, sums the product and then passes the sum through a nonlinear transfer (e.g. hyperbolic tanh) function also called a threshold function to produce a result which is the activation of the PE. The activation is multiplied by the connection weights going to the next layer. The input signal is propagated through the network in this way until it reaches the output layer.

The most important and time-consuming step in model development is NN training. A NN learns the problem behaviour through this process. The NN would be taught with measured/simulated samples from a training set. From the training point of view, the NN performance is evaluated by computing the difference between the actual NN outputs and the desired outputs for all the training samples.

In this paper, the NN was designed to learn to extract the five parameters (depth to the centre of the sheet (h), half-width of the sheet (a), polarization amplitude (k), zero distance

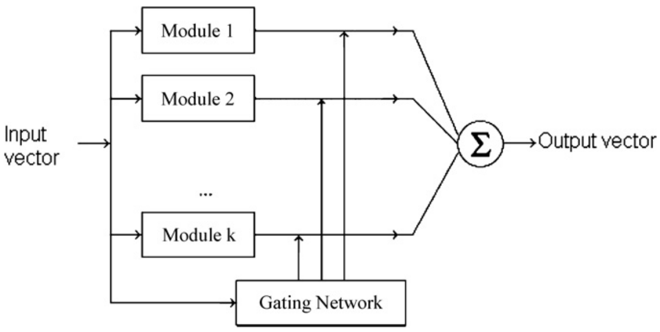

Figure 2. Block diagram of the general modular neural network architecture (Haykin 1994).

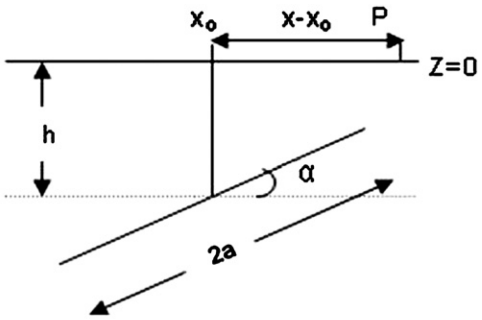

Figure 3. Inclined sheet geometry of infinite horizontal extent.

from the origin (x_o) and angle of inclination of the sheet (α), figure 3) from the input data (SP data with distance).

The input layer has as many input nodes as there are input samples (SP data). There are five output nodes in the output layer for the desired parameters (h, a, k, x_o and α).

The training process adjusts the weight parameters (w) in the network such that the error between the neural model predictions and the desired output $E(w)$ is minimized, where $E(w)$ is a nonlinear function of w. Due to the complexity of $E(w)$, iterative algorithms are used to explore the weight space. In iterative methods, we start with an initial guess of w and then iteratively update w as

$$w_{\text{new}} = w_{\text{old}} + \eta\, n$$

where w_{old} and w_{new} are the current and new vectors containing the values of the weights, n is the update direction and η is a positive step size regulating the extent to which w can be updated in that direction (Zhang and Gupta 2000).

If the training error remained high and flat for a large number of iterations, this means that the training process is trapped in a local minimum. We can perturb w, try a new initial guess and restart the training process.

Modular neural network (MNN)

The NN used was trained using the modular neural network (MNN) architecture, which was successfully used for different geophysical data (El-Kaliouby and Poulton 1999, El-Kaliouby 2001, Zhang *et al* 2002, Bhatt and Helle 2002). A modular neural network (MNN) as defined by Haykin (1994) is one in which the computation performed by the network can be decomposed into a group of modules (local experts) that operate on distinct inputs without communicating with each other (figure 2). The outputs of the modules are mediated by an integrated unit (gating network) that is not permitted to feed information back to the modules. In particular, the integrated unit both (1) decides how the output of the modules should be combined to form the final output of the system and (2) decides which modules should learn which training patterns. The use of a modular approach can be justified on biological grounds.

A MNN combines supervised and unsupervised learning paradigms in a unique way. The gating network learns to break a task into several parts, which is unsupervised learning, and each module is assigned to learn one part of the task, which is supervised learning. The modules compete with each other to learn each training pattern, controlled by the gating network, which performs the function of a mediator among the modules (Haykin 1994). Each module or local expert and the gating network receive the same input pattern from the training set. The gating network and the modules are trained simultaneously. The gating network determines which local expert produced the most accurate response to the training pattern and the connection weights in that module are allowed to be updated to increase the probability that this module will respond best to similar input patterns (Haykin 1994, Zhang *et al* 2002).

Formulation of the problem

The SP potential at any point P on the surface on a line perpendicular to the strike of a 2D inclined sheet of infinite horizontal extent (figure 3) is given by (Murthy and Haricharan 1985, Sundararajan *et al* 1998)

$$V(x) = k \ln \left\{ \frac{[(x - x_o) - a\cos\alpha]^2 + (h - a\sin\alpha)^2}{[(x - x_o) + a\cos\alpha]^2 + (h + a\sin\alpha)^2} \right\} \quad (1)$$

where the polarization parameter $k = \frac{I\rho}{2\pi}$, h is the depth to the centre of the sheet, α is the inclination, a is the half width, ρ is the resistivity of the medium, x_o is the zero distance from the origin and I is the current density (current per unit area) of the medium.

Results

Synthetic example

The synthetic example data are samples at 81 points of input data over a 160 m profile with a 2 m interval (figure 4). We used the MNN to invert the SP data using 81 points for the input layer; a similar number of nodes was also used in the hidden layer. Five local experts with seven processing elements have been used for the MNN. The hyperbolic tangent (Tanh) transfer function was used to modify activations in the hidden layer. We assumed the target to have the parameters $h = 10$ m, $a = 5$ m, $\alpha = 40^\circ$, $x_o = 10$ m and $k = 100$ mV, leading to the response shown in figure 4. We have used 6125 training models covering the ranges of the parameters. The

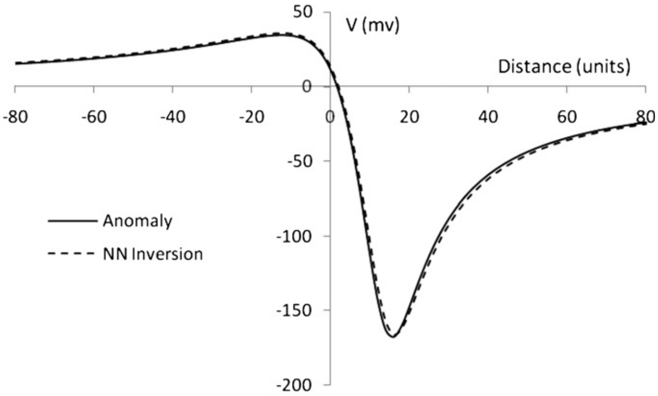

Figure 4. Synthetic SP anomaly profile over a sheet-like body and its NN inversion response. The sheet parameters are $h = 10$ m, $a = 5$ m, $\alpha = 40°$, $x_o = 10$ m and $k = 100$ mV.

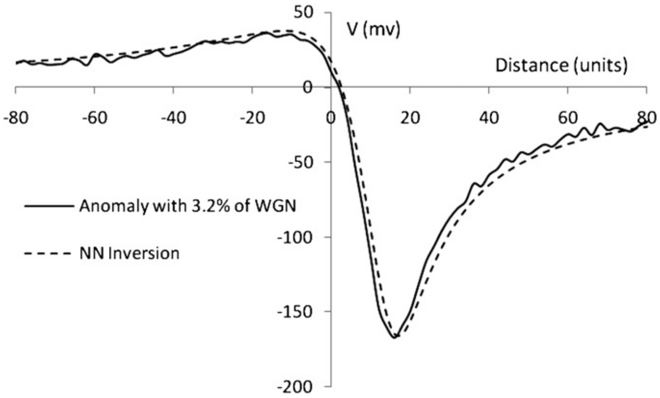

Figure 5. Synthetic SP anomaly profile over a sheet-like body with 3.2% of random WGN and its NN inversion response.

parameter ranges that were used for training the network are as follows:

- the depth h range (5–15 m), with five points in this range;
- the half-width a range (2–8 m), with five points in this range;
- the inclination α range (20°–60°), with seven points in this range;
- the polarization amplitude term k range (70–130 mV), with seven points in this range;
- the origin location x_o range (−15 m to 30 m), with five points in this range.

The choice of the centres of the parameter ranges and the expected ranges of the parameters is based on the measured field data behaviour. These expected ranges are less restricted than assuming an initial starting model in the local inversion methods that might be far from the true parameters and the solution might get trapped in a local minimum.

The choice of the parameter ranges is dependent mainly on the measured field (voltage) data response. Based on the master curves of the parameter responses, deep targets show a broad curve while shallow targets show a sharp narrow curve. The angle of inclination will affect the symmetry of the field curve anomaly while the zero distance from the origin range is selected around the anomaly minimum (negative trough). A coarse range is usually selected at the beginning with a small number of points for each parameter in order to explore the suitability of the selected range to fit the field data. If any of the parameters does not fall within the selected range, the NN can report that this parameter is out of range so that we can expand the range. In the learning process, we can examine the learning error of each parameter individually and the overall root mean square error (RMS) as well; if the learning error is accepted, we compare the misfit between the NN inversion response with the field data. If the misfit is reasonable, then the suggested range is suitable; otherwise we need to narrow/expand the parameter ranges based on the learning error of each parameter. We can also increase the resolution of each parameter by increasing the number of points per range. This process usually requires two-step (coarse to fine) training for the neural network which does not take much time. However, once the network is well trained, it can invert any field data that fall within the training range in almost no time.

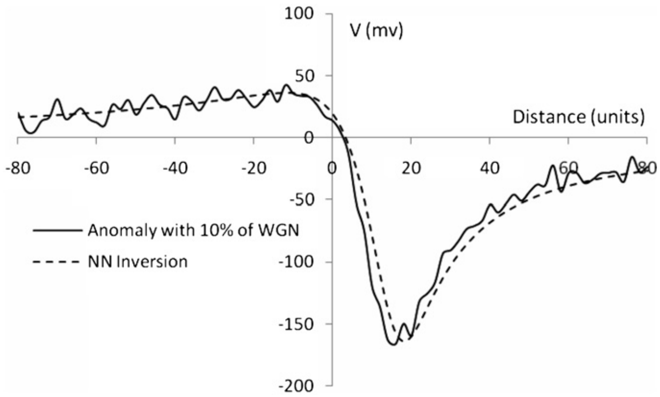

Figure 6. Synthetic SP anomaly profile over a sheet-like body with 10% of random WGN and its NN inversion response.

The result of the MNN inversion response is shown in figure 4 and the inversion parameters are tabulated in table 1.

Noise analysis

The effect of random noise has been studied by adding 3.2% (30 dB) and 10% (20 dB) of white Gaussian noise (WGN) to the SP anomaly (figures 5 and 6). It can be observed that the inverted noisy anomalies are still acceptable, confirming that the NN inversion provides satisfactory results up to 10% of random WGN (table 1).

Field examples

To investigate the applicability of the proposed technique, we used two field examples as follows.

Surda anomaly

Figure 7 shows a well-known SP anomaly (Murthy *et al* 2005) obtained from the Surda area of the Rakha mines, Singhbhum copper belt, Bihar, India. The Surda SP anomaly was sampled at 26 points of input data over a 250 m distance with a 10 m interval. We have used 16 807 training models covering the

Table 1. Theoretical example.

Parameters	h (m)	a (m)	α (degrees)	k (mV)	x_o (m)
Assumed values	10	5	40	100	10
NN inversion values without noise	10.23	5.42	38.65	96.67	10.70
NN inversion with 3.2% of WGN	10.32	5.54	37.89	96.35	10.71
NN inversion with 10% of WGN	10.85	5.49	38.66	97.3	11.93

Table 2. Interpreted SP parameters of Surda anomaly.

Parameters	h (m)	a (m)	α (degrees)	k (mV)	x_o (m)
Paul (1965)	21	40.20	20.01	–	–
Rao *et al* (1970)	30.48	34.87	10.01	–	–
Jagannadha Rao *et al* (1993)	29.88	29.40	45	–	15.00
Sundararajan *et al* (1998)	27.65	32.35	13.20	–	–
Murthy *et al* (2005)	26.52	19.81	57.63	–	15.84
NN inversion	27.78	19.51	50.96	130.86	5.86

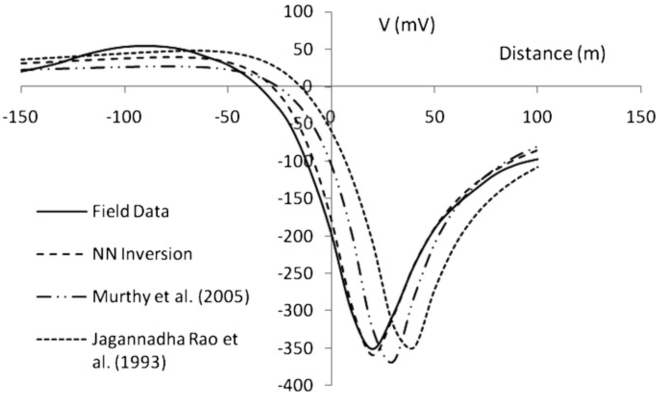

Figure 7. Self-potential anomaly profile over a sulfide mineralization zone in the Surda area of Rakha mines (after Murthy and Haricharan 1984) and its neural network inversion compared with Murthy *et al* (2005) and Jagannadha Rao *et al* (1993).

ranges of the parameters. The parameter ranges that were used for training the network are as follows:

- the depth h range (10–40 m), with seven points in this range;
- the half-width a range (10–30 m), with seven points in this range;
- the inclination α range (20°–50°), with seven points in this range;
- the polarization amplitude k range (90–180 mV), with seven points in this range;
- the origin x_o range (−20 to 40 m), with seven points in this range.

The MNN inversion response compared with the Murthy *et al* (2005) inversion response and the measured response is shown in figure 7, and the inversion parameters are tabulated in table 2. The present inversion response shows a better fit with the measured data compared with that of Murthy *et al* (2005).

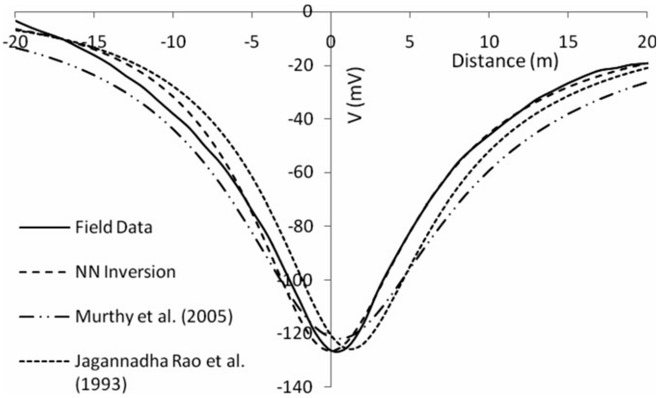

Figure 8. Self-potential anomaly profile over a sulfide body in the Kalava fault zone, Cuddapah Basin, India (after Rao *et al* (1982)), and its neural network inversion compared with Murthy *et al* (2005) and Jagannadha Rao *et al* (1993).

Table 3. Interpreted SP parameters of Kalava anomaly.

Parameters	h (m)	a (m)	α (degrees)	k (mV)	x_o (m)
Jagannadha Rao *et al* (1993)	7.59	3.75	80	–	0.4
Murthy *et al* (2005)	9.38	3.96	80.76	–	−0.4
NN inversion	7.2	3.15	78.72	68.29	−0.9

Kalava anomaly

Figure 8 shows the SP anomaly profile across a mineralization in the Kalava fault zone, Cuddapah basin, India (Rao *et al* 1982). The Kalava SP anomaly was sampled at 41 points over a 40 m distance with a 1 m interval. We have used 12 500 training models covering the ranges of the parameters. The parameter ranges that were used for training the network are as follows:

- the depth h range (5–15 m), with ten points in this range;
- the half-width a range (1–7 m), with ten points in this range;
- the inclination α range (80°–120°), with five points in this range;
- the polarization amplitude k range (50–100 mV), with five points in this range;
- the origin location x_o range (−5 to 5 m), with five points in this range.

The MNN inversion response compared with the inversion response of Rao *et al* (1982) and with the measured results is shown in figure 8, and the inversion parameters are tabulated in table 3. Again, the present inversion response shows a better fit with the measured data compared with that of Rao *et al* (1982).

Conclusion

Inversion of SP parameters of 2D inclined sheets of infinite horizontal extent has been studied. MNN inversion has been utilized for five SP parameters, where it has been tested first on synthetic data and then on two field examples. The hyperbolic tangent function (tanh) has been used as an activation function in MNN. The effect of noise has been studied and the results show that the technique gives satisfactory results even up to 10% noise. The inversion results of the two field examples showed good agreement compared with other inversion techniques in use. Therefore, the successful application of the NN inversion to synthetic and field data demonstrates the validity of this method.

References

Abdelrahman E M, Ammar A A, Hassanein H I and Hafez M A 1998 Derivatives analysis of SP anomalies *Geophysics* **63** 890–497

Abdelrahman E M, Ammar A A, Sharafeldin S M and Hassanein H I 1997a Shape and depth solutions from numerical horizontal self-potential gradients *Appl. Geophys.* **36** 31–43

Abdelrahman E M, El-Araby T M, Ammar A A and Hassanein H I 1997b A least-squares approach to shape determination from self-potential anomalies *Pure Appl. Geophys.* **150** 121–8

Abdelrahman E M and Sharafeldin S M 1997 A least squares approach to depth determination from residual self-potential anomalies caused by horizontal cylinders and spheres *Geophysics* **62** 44–8

Bescoby D J, Cawley G C and Chroston P N 2006 Enhanced interpretation of magnetic survey data from archaeological sites using artificial neural networks *Geophysics* **71** H45–53

Bhatt A and Helle H 2002 Committee neural network for porosity and permeability prediction from well logs *Geophys. Prospect.* **50** 645–60

Bhattacharya B B and Roy N 1981 A note on the use of a nomogram for self-potential anomalies *Geophys. Prospect.* **29** 102–7

Corwin R F 1990 The self-potential method for environmental and engineering applications *Geotechnical and Environmental Geophysics* ed S H Ward (Tulsa, OK: Society of Exploration Geophysicists) pp 127–45

El-Kaliouby H 2001 Extracting IP parameters from TEM data *Computational Neural Networks for Geophysical Data Processing* ed M M Poulton (Oxford: Pergamon) chapter 17

El-Kaliouby H and Poulton M 1999 Inversion of coincident loop TEM data for layered polarizable ground using neural networks *Society of Exploration Geophysicists (SEG) 69th Annual Meeting (Houston, TX, USA)*

Jagannadha Rao S, Rama Rao P and Radhakrishna Murthy I V 1993 Automatic inversion of self-potential anomalies of sheet-like bodies *Comput. Geosci.* **19** 61–73

Haykin S 1994 *Neural Networks: A Comprehensive Foundation* (London: Macmillan)

Macias C, Sen M and Stoffa P 2000 Artificial neural networks for parameter estimation in geophysics *Geophys. Prospect.* **48** 21–47

Meiser P 1962 A method for quantitative interpretation of self-potential anomalies *Geophys. Prospect.* **10** 203–18

Mendonça C A 2008 Forward and inverse self-potential modeling in mineral exploration *Geophysics* **73** F33–43

Murthy B V S and Haricharan P 1984 Self-potential anomaly over double line of poles—interpretation through log curves *Proc. Indian Acad. Sci. Earth Planet Sci.* **93** 437–45

Murthy B V S and Haricharan P 1985 Nomograms for the complete interpretation of spontaneous potential profiles over sheet like and cylindrical 2D structures *Geophysics* **50** 1127–35

Murthy I V R, Sudhakar K S and Rao P R 2005 A new method of interpreting self-potential anomalies of two-dimensional inclined sheets *Comput. Geosci.* **31** 661–5

Paul M K 1965 Direct interpretation of self-potential anomalies caused by inclined sheets of infinite extension *Geophysics* **30** 418–23

Paul M K, Data S and Banerjee B 1965 Interpretation of SP anomalies due to localized causative bodies *Pure Appl. Geophys.* **61** 95–100

Poulton M M 2001 *Computational Neural Networks for Geophysical Data Processing* (Oxford: Pergamon)

Poulton M M 2002 Neural networks as an intelligence amplification tool: a review of applications *Geophysics* **67** 979–93

Rao B S R, Murthy I V R and Reddy S J 1970 Interpretation of self-potential anomalies of some simple geometrical bodies *Pure Appl. Geophys.* **78** 60–77

Rao R M, Babu H and Sivakumar Sinha G 1982 A Fourier transform for the interpretation of self-potential anomalies due to two-dimensional inclined sheet of finite depth *Pure Appl. Geophys.* **120** 365–74

Sato M and Mooney H M 1960 The electrochemical mechanism of sulfide self-potentials *Geophysics* **25** 226–49

Sill R 1983 Self-potential modeling from primary flows *Geophysics* **48** 76–86

Spichak V and Popova I 2002 Artificial neural network inversion of magnetotelluric data in terms of three-dimensional earth macro-parameters *Geophys. J. Int.* **142** 15–26

Sundararajan N, Arun Kumar I, Mohan N L and Seshagiri Rao S V 1990 Use of Hilbert transform to interpret self potential anomalies due to two dimensional inclined sheets *Pure Appl. Geophys.* **133** 117–26

Sundararajan N, Srinivasa Rao P and Sunitha V 1998 An analytical method to interpret self-potential anomalies caused by 2D inclined sheets *Geophysics* **63** 1551–5

Van Der Baan M and Jutten C 2000 Neural networks in geophysical applications *Geophysics* **65** 1032–47

Zhang L, Poulton M M and Wang T 2002 Borehole electrical resistivity modeling using neural networks *Geophysics* **67** 1790–7

Zhang Q and Gupta K 2000 *Neural Networks for RF and Microwave Design* (Boston, MA: Artech House)