Journal of Applied Geophysics 118 (2015) 37–46

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

Capability of self-organizing map neural network in geophysical log data classification: Case study from the CCSD-MH

Ahmed Amara Konaté [a,*], Heping Pan [a,*], Sinan Fang [a], Shazia Asim [a], Yao Yevenyo Ziggah [b], Chengxiang Deng [a], Nasir Khan [a]

[a] Institute of Geophysics and Geomatics, China University of Geosciences (Wuhan), Lu Mo Road 388, 430074 Wuhan, Hubei Province, China
[b] Faculty of Information Engineering, China University of Geosciences (Wuhan), Lu Mo Road 388, 430074 Wuhan, Hubei Province, China

ARTICLE INFO

Article history:
Received 27 December 2014
Received in revised form 20 March 2015
Accepted 7 April 2015
Available online 11 April 2015

Keywords:
Well log interpretation
Lithology
Neural networks
Crystalline rocks
Cross validation technique
Geophysical statistics

ABSTRACT

Well log interpretation is one of the prime sources of information for deep lithology in drilling research. Because of the complex geological features of the crystalline metamorphic rocks, more complex nonlinear functional behaviors exist for well log interpretation purposes. Hence, establishing a prediction technology that can accurately interpret/classify well log data in terms of lithology is of major significance. This study, for the first time, explores the application of self-organizing map neural network (SOM) in the classification of metamorphic rocks from Chinese Continental Scientific Drilling Main Hole (CCSD-MH) log data. For this purpose, a total of 33,326 data points derived from resistivity, P-wave velocity, bulk density, photoelectric absorption capture cross section, gamma ray, potassium content and neutron logs were used as an input pattern to a SOM to classify lithology in five categories: orthogneiss, paragneiss, eclogite, amphibolite and ultramafic rocks. Comparison of SOM results to those of feed-forward neural network (FFNN) was also carried out. The cross-validation method was used to investigate the robustness of the two neural networks in terms of classification accuracy in the context of lithology clustering tasks by sampling rotation. Statistical tests such as student paired samples *t*-test was carried out to guide in classification decision of the CCSD-MH data. The results of this study have proven that SOM appears to be comparable to FFNN in classifying lithology using geophysical log data from crystalline rocks. This proposed SOM approach can serve as practical alternative technology to be used in drilling research.

1. Introduction

Lithology (rock type) identification is significant for many earth science disciplines (Chang et al., 2000). It is the geological description of rocks in wells at different depths and is important for the determination of petrophysical properties. The general method of acquiring information on lithology during drilling operation consists of interpretation of underground cores by experienced geologists. However, the recuperation of cores is very expensive and its interpretation is time-consuming. Furthermore, technically, core recovery is usually partial (less than 100%). This limitation in core recovery has been echoed as a serious difficulty for lithology interpretations or reconstructions by several authors such as Benaouda et al. (1999), Bhatt and Helle (2002), Chang et al. (2000), Chang et al. (2002) and Maiti and Tiwari (2010). As a result, it is desirable to find a cost effective and practical alternative method that offers similar or improved accuracy (Chang et al., 2000).

Conversely, lithology can also be obtained from geophysical well logs thereby providing indirect information of various physical properties of the formations penetrated. It is considered a relatively non-expensive and closely related to true conditions when compared with core analysis. This indirect geophysical information from borehole sidewall constitutes one of the most important contributions of geophysics in many geological and engineering areas such as petroleum, scientific research drilling, mining, hydrogeology, geotechnique, environment and many others.

Geophysical well logs can be categorized into electrofacies. These electrofacies are equivalent, but not equal to the lithology deduced from core description because electrofacies are based on the combination of logging tool signal responses to lithology, while lithofacies are inferred directly from the visual check of featured rocks (Bhatt and Helle, 2002). In line with this, electrofacies can be used as an alternative to predict lithology in uncored intervals as well as cored intervals in holes (Benaouda et al., 1999). Consequently, this study utilized only geophysical well log data obtained from the records of the Chinese Continental Scientific Drilling Main Hole (CCSD-MH). The CCSD was a joint project by the Chinese government and the international continental drilling program. Its major scientific aim was to investigate the composition, deep structure, tectonic evolution and geodynamic processes of the Sulu ultrahigh-pressure metamorphic belt that are not exposed, by means of drilling a hole into crystalline rocks (Xu et al., 2009).

* Corresponding authors.
E-mail addresses: konate77@yahoo.fr (A.A. Konaté), panpinge@163.com (H. Pan).

http://dx.doi.org/10.1016/j.jappgeo.2015.04.004
0926-9851/

Several deep drill holes explore the sedimentary rocks worldwide. So far, very few drill holes penetrate the crystalline rocks to a great depth, and most of them are under scientific drilling programs (He et al., 2006). Crystalline rocks cover a broad spectrum of igneous, metamorphic and some sedimentary rocks in which recrystallization process has been important to their formation. These occur in a range of continental and oceanic settings, where a number of boreholes have been and continue to be drilled in order to provide greater insight into exact, fundamental and globally significant knowledge of the composition, structure and processes of the earth's crust by means of geophysical surveys and geological mapping (Harvey et al., 2005).

Enormous investigations have been carried out by geophysicists based on crystalline rocks using well logging data (Bartetzko et al., 2005; Luo and Pan, 2010; Maiti and Tiwari, 2009; Pan et al., 2010; Pechnig et al., 1997, 2001, 2005; Pratsone et al., 1992). The general understanding gathered from these research uncovered that, the well logging data from crystalline rocks are mostly difficult to analyze because of their complicated geological characteristics and the difficulty in understanding and using the intensive information content in these data.

It is important to note that the logging tools were initially not designed for crystalline rocks (Bartetzko et al., 2005). Subsequently, in complex geological situations such as the study in context, several factors such as pore fluid, effective pressure, fluid saturation and pore shape among others extremely affect the borehole logging tool signals (Maiti et al., 2007; Pechnig et al., 1997). In this case, lithology classification from well log data in crystalline rocks is considered as a complex and nonlinear geophysics problem (Maiti and Tiwari, 2009). Besides, to date, there are no well-bounded interpretation/classification methods available for crystalline rocks (Pechnig et al., 2005) in Geophysics.

Geophysics is an area that incorporates geology, physics and mathematics in order to understand how the Earth works. With the appropriate mathematical and statistical methods, geophysicists can extract interpretation about the subsurface properties by utilizing either the surface or borehole measurements (Maiti et al., 2007). So, with this in mind, the statistical methods in this prevalent situation are the use of methods such as principal component analysis (Wolff and Pelissier-Combescure, 1982) and discriminant-factor analysis (Busch et al., 1987; Delfiner et al., 1987). It is worth mentioning that most of these statistical methods rarely take into consideration the complexity and nonlinearity of the well log data. That is why these statistical approaches of automatic data classification are not adequate and/or efficient enough (Briqueu et al., 2002).

Due to the complexity and nonlinearity in crystalline rocks and the limitations in statistical methods, Artificial Intelligence such as artificial neural networks (ANNs) are therefore used as an alternative classification technologies to mathematical modeling in crystalline rocks.

Several researchers have incorporated well logging data into lithology classification through supervised ANN scheme (input/output mapping) in borehole studies into continental crystalline rocks such as Maiti et al. (2007), Maiti and Tiwari (2009, 2010); including CCSD-MH data (Pan et al., 2010). These researchers, in their respective studies have mentioned promising performance by supervised scheme approach. However, in this study an unsupervised scheme of self-organizing map (SOM) is employed on CCSD-MH logging data for classification. To the best of the author's knowledge, this is the first time SOM is applied on the CCSD-MH logging data. Therefore, this study constitutes a good foundation for quality control of SOM classifier from future deep boreholes drilled in similar geological contexts.

2. Previous work

Detailed studies of lithology have been carried out by researchers on deep borehole logging data from the CCSD-MH. Among them, are Niu et al. (2004) using a cross-plotting technique to recognize lithology in the depth section 0–2000 m. Pan et al. (2005) also are basing on a radioactive logging interpretation to make estimates of lithology. Wang et al. (2005) also worked on rock type identification based on well logging analysis of ultrabasic rock section. In Xu et al. (2006), discriminant function analysis was applied successfully to identify the lithology. However, Jing et al. (2007) demonstrated that the spatial distance method can be used to identify eclogites. On the other hand, Luo and Pan (2010) used core-log correlation and cross-plotting methods, and the results allowed the authors to conclude that the lithology are mainly comprised of orthogneiss, paragneiss, eclogite, amphibolite, and ultramafic rocks. Moreover, a preliminary study using artificial neural network as a classifier was started by Pan et al. (2010). In their study, feed forward back propagation is used and it was found to be accurate in training results.

Based on the above development, the general understanding that can be inferred is as follows:

(i) There are very few studies reported that have applied Artificial Intelligence approach to the CCSD-MH well log database. The study of Pan et al. (2010) employed a supervised design well-known as Feed Forward Neural Network using a Back Propagation algorithm (FFNN-BP). However, the FFNN-BP algorithm has some drawbacks. For instance, a) The FFNN-BP need a target vector for each input vector used to make training feasible. In case of lithology classification tasks, the target vector is core description "lithofacies" which must be provided in the FFNN-BP training process. As stated earlier, technically the availability of the core data is often partial. Thus, FFNN-BP is limited in training; b) for the same cause, FFNN-BP may also be ineffective to supply a real-time response. This is because the geophysicists may have to wait until the core samples are gathered and interpreted by experienced geologists; c) FFNN-BP in training process, still depends on human interference that is, this algorithm leaves a lot of responsibility to the user for determining the number of hidden layer and number of node in hidden layer by sequential trial and error procedure; d) FFNN-BP is sensitive to the random weight value initialization and the local minima problem, which may lead the model to develop in an inaccurate direction. Therefore, in such situations, unsupervised neural classifiers might be more suitable tools to be utilized.

(ii) Finally, CCSD-MH log database are still not fully exploited. In view of this, it creates an opportunity to apply and evaluate some of the well-known methods of ANNs.

This current study suggests a better technique based on unsupervised neural networks, Self-organizing map (SOM) to interpret/classify CCSD-MH log data in terms of lithology. The main advantage of this approach with respect to others is that, it uses only well log signatures and tries to discover well log data having similar characteristics followed by classification.

This study constitutes an additional contribution of neural network as a well log interpreter/classifier in terms of lithology in the framework of research drilling in continental crust. This point here highlights the ability of SOM to classify geophysical well log data in the context of lithology as well as evaluate its performance to those obtained by feed forward neural network (FFNN) on CCSD-MH data. We demonstrate the significance for geophysicists to understand the implication of statistical resampling method on ANN modeling. This statement stresses the application of *k*-fold cross-validation technique, with the aim of quantifying the performance of these two neural classifiers in term of classification accuracy by sampling variation. Finally, statistical tests such as student paired samples *t* test was carried out to measure the sensitivity of the two neural networks in rock type classification tasks using CCSD-MH data.

3. Well logging data

In the CCSD project (Fig. 1), well logging was one of the most significant stages and crucial technologies. The logging engineers collected many modern high technologies and utilized more than 20 kinds of logging parameters using advanced ECLIPS5700 logging system to investigate the sidewall of the CCSD-MH continuously (Niu et al., 2004) in order to: (i) Analyze log response and restore the lithology profile; (ii) Correct the depth and restore the core orientation; (iii) Interpret the image geological feature; (iv) Research the rock's mechanics character, sonic anisotropy and formation stress; (v) Compare wave impedance with the VSP profile, analyzing the reflectors; and (vi) Interpret the magnetic susceptibility and temperature log.

CCSD project which commenced in June 2001 and achieved its target depth of 5158 m in April 2005 has provided continuous records of physical and chemical data of the crystal metamorphic rocks drilled. The petroleum industry logging methods were utilized to survey the entire section of the main hole continuously. General information on logging technology can be found in Pan et al. (2002) and Niu et al. (2004).

For this study, it has collected 33,326 data points from the interval depth of 100–5025 m of CCSD-MH. For each data point there are seven input variables. There are: CNL (neutron log), DEN (density log), PE (photoelectric absorption capture cross section), V_P (P-wave velocity), natural GR (gamma ray log), K (potassium content), and RD (resistivity of laterolog deep). The profile of these logs is depicted in Fig. 2. Besides, the data for this study involved different parameters that have different physical meanings and units. To make sure that each variable (= log curve) is treated similarly in the model, the data were normalized in [0, 1].

4. Methods

4.1. Self-organizing map (SOM)

The self-organizing map (SOM) or Kohonen networks is one of the utmost popular neural network technology, which was developed by professor Kohonen (2001). It is based on unsupervised learning that is to classify the training data without any target vector as compared to FFNN. SOM is frequently used to visualize a high dimensional space, since it is capable of mapping highly nonlinear high dimensional input space in lower dimensional spaces usually one or two dimensions,

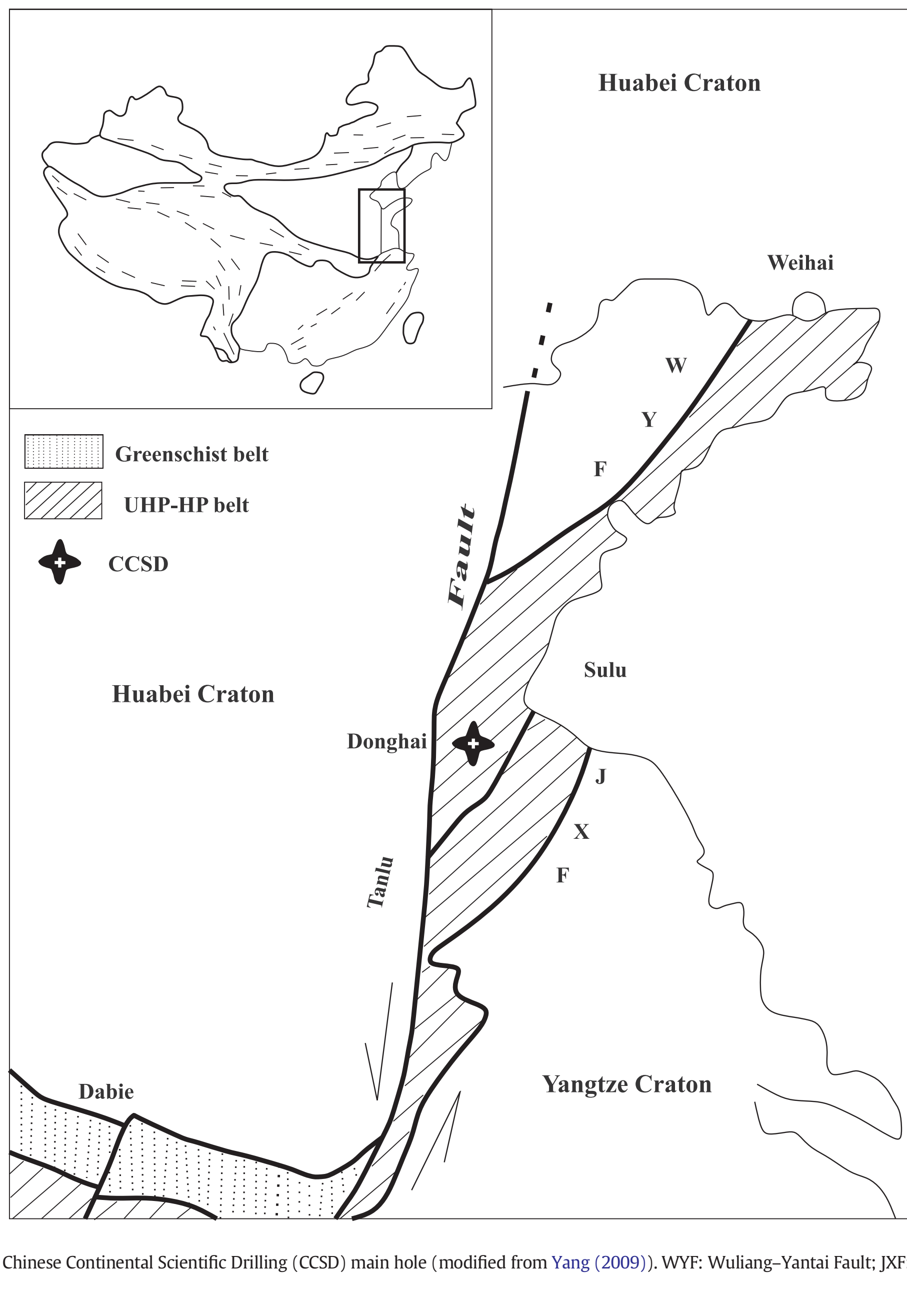

Fig. 1. The location of Chinese Continental Scientific Drilling (CCSD) main hole (modified from Yang (2009)). WYF: Wuliang–Yantai Fault; JXF: Jiashan–Xionshui Fault.

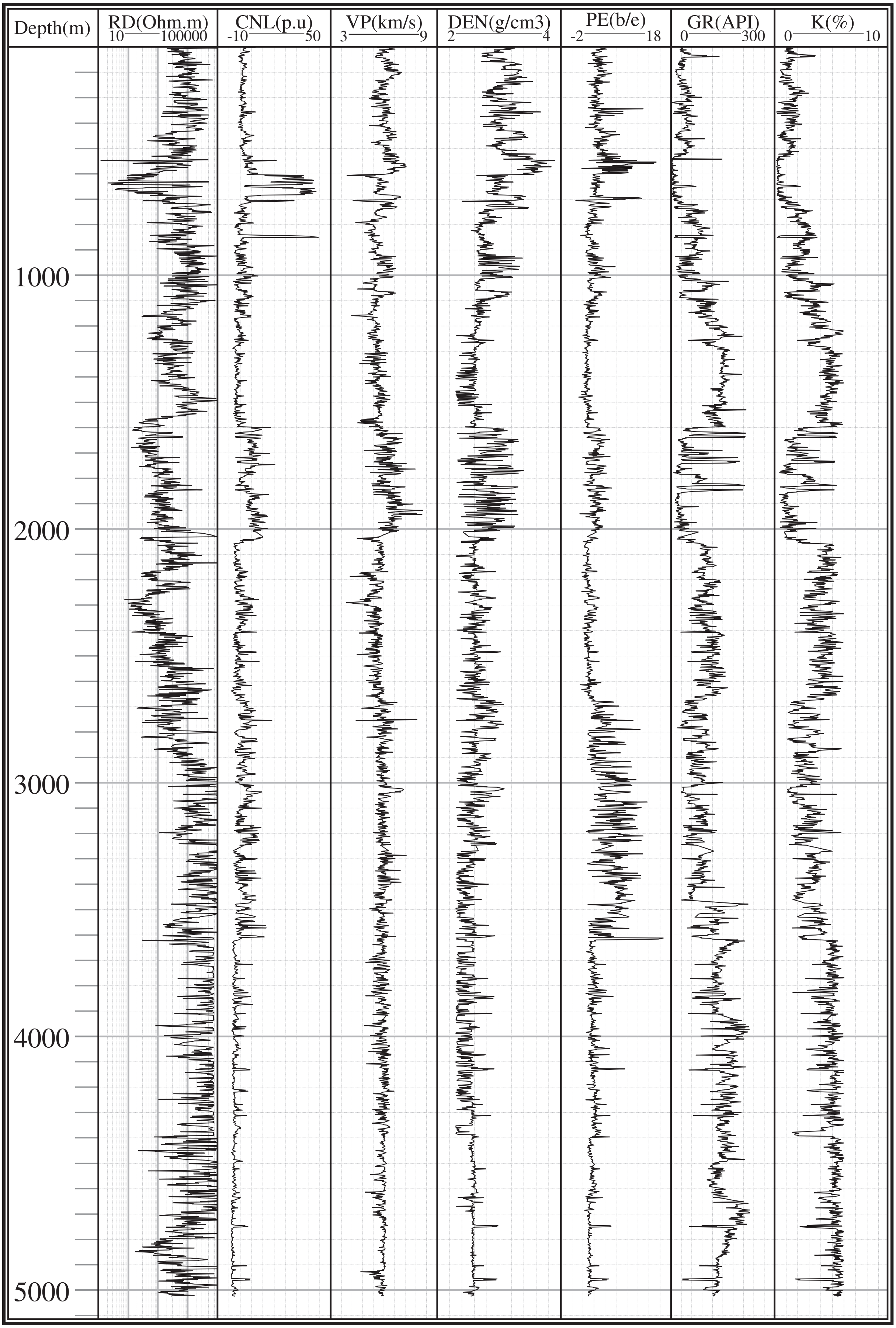

Fig. 2. CCSD-MH logs utilized in this study.

preserving topological neighborhood constraint (Oja, 2002). In this sense, SOM overcame the shortcomings of the cross-plotting technique, in which visualization and interpretation become seriously complex when the dimensional space is more than three. Nevertheless the main application of SOM is for data classification as used in this study to find and classify structure having similar characteristics in large, complex and nonlinear geophysical log data.

The applications of SOM in logging studies have been carried out by many scholars to demonstrate its effectiveness in lithology classification (Baldwin et al., 1990; Briqueu et al., 2002; Chang et al., 2002; Ouadfeul and Aliouane, 2012; Zhang et al., 1999). All of this research has further shown the abilities of SOM to solve different problems using well log data. Based on the aforementioned assertions, SOM was therefore considered as the best potential candidate in unsupervised learning to investigate possible application in crystalline rocks.

The design of the SOM utilized in this study is depicted in Fig. 3. It follows from Fig. 3, the SOM consists of two layers of neurons that is input neurons ($X_1, X_2, \dots X_n$) and output neurons ($Y_1, Y_2, \dots Y_m$). Input neurons fully connected with weights, W_{ij}, to output neurons. n designed the number of inputs neurons (length of training vectors) and m denotes the maximum number of clusters to be formed (number of categories).

It is important to recall that for this study, each data point (input vector) has seven input variables (RD, CNL, Vp, DEN, PE, GR, and K) in the training data. Each of the data points is classified as falling in one of 5 clusters. To enable classification the input variables have been normalized to values between [0, 1].

SOM learning process is illustrated below as (Kohonen, 2001):

Step 1: Select output layer network topology.
Step 2: $t \leftarrow 0$: Initialize values for the weight vectors w_{ij} to random values confined [0, 1].
Step 3: Present an input vector I from the set of training data.
Step 4: Determine the output neuron j^* such that its weight vector wj^* is the closest to the input vector I:

$$\left\| w_{j*}(t) - I(t) \right\| = \min \left\| w_j(t) - I(t) \right\| \quad \text{for all } j.$$

Step 5: Update the weight vectors for all output neuron j within a specified neighborhood of j^*:

$$w_j(t+1) \leftarrow w_j(t) + \eta(t) \times h(t,r) \times \left[\mathrm{I}(t) - w_j(t) \right].$$

Step6: If t > T stop, else $t \leftarrow t + 1$ and go to Step 3.

In this algorithm, t is the current iteration number, T is a number of iterations fixed, and $\eta(t)$ is a learning rate which value is in between 0 and 1 as t increase $\eta(t)$ approach zero; $h\,(t, r)$ is a neighborhood function which can be constant or time-dependent. r means the topological distance between output neuron j and output neuron j^* $r = \|r_j - r_j *\|$ where r_j and r_{j*} are the coordinates of output neuron j and j^* respectively.

After grouping the clusters, it was essential to characterize (label) the different cluster, in order to show which lithological type of each cluster has the greatest possibility to represent. This labeling was done through the application of the results of logging responses obtained by Luo and Pan (2010). Table 1 shows the statistical results of logging responses from Luo and Pan (2010). Let define μ_{Vp}, μ_{DEN},μ_{PE},μ_{RD},μ_{CNL},μ_{GR},and μ_K the average values of Vp, DEN, PE, RD, CNL, GR, and K, respectively in each cluster obtained by SOM.

If $5.44 \le \mu_{Vp} \le 5.92$, $2.83 \le \mu_{DEN} \le 3.01$, $3.80 \le \mu_{PE} \le 4.78$, $0 \le \mu_{RD} \le 2134.56$, $31.90 \le \mu_{CNL} \le 40.58$, $2.08 \le \mu_{GR} \le 9.28$ and $0.21 \le \mu_K \le 0.41$ then they are called ultramafic rocks.

If $5.97 \le \mu_{Vp} \le 6.47$, $2.97 \le \mu_{DEN} \le 3.36$, $4.73 \le \mu_{PE} \le 6.11$, $5429.66 \le \mu_{RD} \le 13475.45$, $4.68 \le \mu_{CNL} \le 7.03$, $18.09 \le \mu_{GR} \le 41.49$ and $0.74 \le \mu_K \le 1.61$ then they are called eclogites rocks.

If $5.50 \le \mu_{Vp} \le 6.36$, $2.62 \le \mu_{DEN} \le 2.94$, $3.52 \le \mu_{PE} \le 5.68$, $0 \le \mu_{RD} \le 51744.39$, $3.80 \le \mu_{CNL} \le 11.36$, $30.29 \le \mu_{GR} \le 83.77$ and $1.38 \le \mu_K \le 3.26$ then they are called amphibolite rocks.

If $5.34 \le \mu_{Vp} \le 5.98$, $2.54 \le \mu_{DEN} \le 2.80$, $2.69 \le \mu_{PE} \le 4.05$, $0 \le \mu_{RD} \le 48365.08$, $-0.37 \le \mu_{CNL} \le 6.11$, $59.87 \le \mu_{GR} \le 107.73$ and $2.06 \le \mu_K \le 3.66$ then they are called paragneiss rocks.

If $5.48 \le \mu_{Vp} \le 6.14$, $2.45 \le \mu_{DEN} \le 2.65$, $2.05 \le \mu_{PE} \le 5.79$, $0 \le \mu_{RD} \le 64364.38$, $-2.22 \le \mu_{CNL} \le 1.64$, $109.19 \le \mu_{GR} \le 172.43$ and $3.85 \le \mu_K \le 5.59$ then they are called orthogneiss rocks.

Note that, limit superior and limit inferior with the above constraints are related to the results of logging responses obtained by Luo and Pan (2010).

4.2. k-Fold cross-validation analysis

ANN is purely an empirical model, therefore evaluating performance is essential to operational success, since the aim of ANN is to generalize effectively (Twomey and Smith, 1997). Generalization is a central issue for the development of mathematical and statistical models. It refers to the ability of a model to accurately represent the underlying data generation process, rather than the noise features of the training data (May et al., 2010). There is no systematic formulated method for ANN model evaluation performance; the common practice is to base model evaluation on a testing set (Twomey and Smith, 1997).

ANN users in geophysics are not necessarily familiar with statistical resampling theory. The consequence is that, in many applications, the most common method of evaluating performance which is found in most applications in logging problems, is the statistical test called train-test (hold-out cross-validation). It mostly consists of randomly dividing an entire available data point into two or three independent parts: training set and testing set or training set, validation set and testing set. Training set is used to construct the model and testing set is used to evaluate how the model behaves with independent data (new data). Validation set is used during training to decide when training should stop (Twomey and Smith, 1997).

Despite the popularity of this method, its main problem, however, is appropriate data splitting. In light of this, authors such as Van der Baan and Jutten (2000) and Reitermanova (2010) stated that, in their opinion, any bias resulting to the inappropriate split of entire available data could have a negative effect on the performance of ANN. Furthermore, most importantly, for both schema of train-test, the hold-out sampling is generally performed only once during ANN development (May et al., 2010). In this sense, the model generalization ability is not guaranteed.

To avoid the problem of splitting data set, the use of k-fold cross-validation (Hereafter *KCV*) technique is suggested (Van der Baan and Jutten, 2000, and references therein). For this reason, one of the major aims of this study is to use *KCV* technique to evaluate SOM performance in terms of classification accuracy in the context of lithology clustering tasks by sampling rotation on the CCSD-MH data and in addition, to compare its performance with those classified by FFNN. In other words, mathematically, we would like to distinguish which method offers better generalization performance and hence leads to better classification results in terms of sampling rotation.

In *KCV* technique the dataset D is divided into k partition nearly equal-sized independent set D_i such that $D = \bigcup_{i=1}^{k} D_i$. Then k iterations are conducted to train and test the model in such a way that within each trial, the model is trained on the union of any $k-1$ of the k disjoint set while the remaining set is used for testing (Stone, 1974). In the present context, any $k-1$ of the k disjoint sets is being used to train SOM and FFNN scheme with the objective to classify

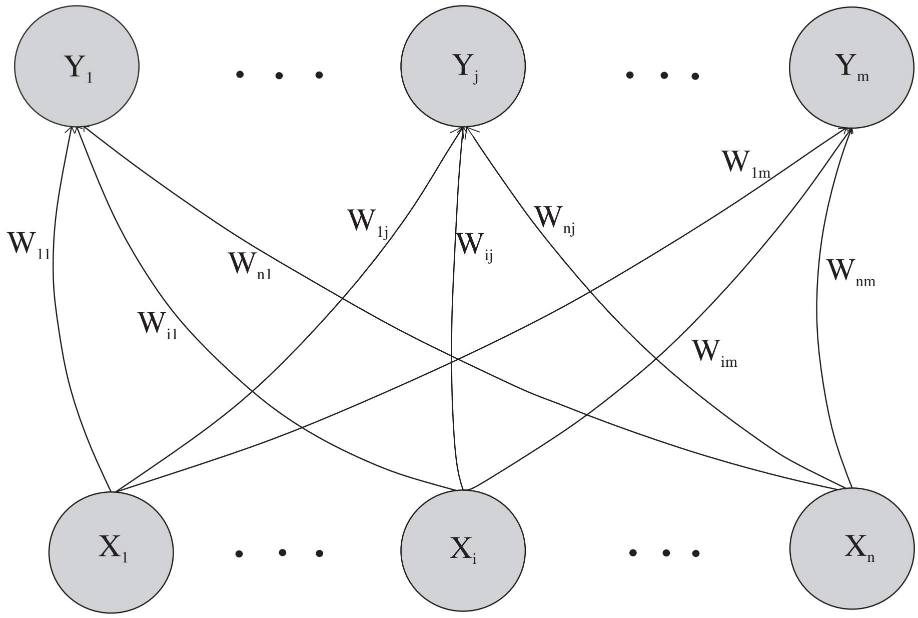

Fig. 3. Architecture of SOM model utilized.
Adopted from Chang et al., 2002.

the CCSD-MH log data into orthogneiss, paragneiss, eclogite, amphibolite, and ultramafic rocks. The remaining set is being used to test the generalization ability of the two methods with unseen log data.

Using k-fold cross-validation an important question is that how one should choose a proper value of k parameter? With a large number of k, the estimator will be very accurate. In this context, scientists can observe and understand the variability of the accuracy which is an advantageous feature when facing method comparison in this study; however, computational time increases with the number of k. On the other hand, the smaller the number of k, the larger the bias of the estimator. As recommended by Rodriguez et al. (2010), the use of $k = 10$ is appropriate for comparing classifiers using *KCV t*echnique. Following this reasoning, the CCSD-MH was divided into 10 disjoint interval depths (Table 2). In each trial it leaves out one interval depth for testing and the remaining for training. This process was continued until all interval depths had been passed once for testing. In this way, SOM (as well as FFNN) model is trained, and tested in 10 trials. Therefore, the average classification accuracy over all 10 disjoint subsets was a good indicator for the rotation of sample performance of the classifiers. Moreover, to have a better idea of the generalization ability of the classifiers with large unseen log data, we also tested each classifier using the whole data set.

4.3. Feed Forward Neural Network

It is pertinent to mention that, in the case of FFNN, the core description was definitely needed to show to a neural network what lithology the geophysicist desires to predict from the well log data. In this experimentation, the core description were classified and encoded in binary code into five categories. Then this information was utilized in the calibration of CCSD-MH geophysical log data. Therefore, the geophysical log values should produce (1 0 0 0 0) at the output neuron if the lithology is orthogneiss; (0 1 0 0 0) for paragneiss; (0 0 1 0 0) if eclogite is present; (0 0 0 1 0) to indicate amphibolite and (0 0 0 0 1) for the presence of ultramafic rocks.

A FFNN scheme with a three-layer (one input–one hidden layer–one output model) was considered in this research. The scaled conjugate gradient (SCG) algorithm was used for training. It uses a step size scaling mechanism which avoids time-consuming line search per training iteration, thus, making the algorithm faster than other second order algorithms (Moller, 1993). A sigmoid transfer function was used for the hidden nodes and the output nodes. For more detailed description of FFNN and the underlying concepts, reference may be made to Bishop (1995) and Haykin (1999).

5. Application

The 10-fold cross validation technique results and average correct classification percentages for SOM and FFNN in classifying crystalline metamorphic rocks from CCSD-MH logs are given in Tables 3 and 4. From the visual inspection of Table 3, across the 10 subsets, the correct classification rates for individual training subset fluctuation goes from 77.8% to 83.12% for SOM and range from 70.21 to 92.32% for FFNN. Hence, in terms of average correct classification accuracy model (training) FFNN show comparable results to SOM (Table 4). Again from Table 3, the correct classification rates for individual testing subset run from 77.44% to 82.75% for SOM and between 45.18% and 86.53% for FFNN. Therefore, in terms of average correct classification accuracy (testing), SOM shows comparable results to FFNN with slightly better performance for SOM; because, SOM gave an average of 79.86% over the 10 subsets, slightly higher than 72.84% achieved by FFNN

Table 1
The statistical results of logging responses from Luo and Pan (2010).

Lithology	Vp (km/s)	DEN (g/cm^3)	PE (b/e)	RD (Ohm.m)	CNL (p.u)	GR (API)	K (%)
Ultramafic rocks	5.68 ± 0.24	2.92 ± 0.09	4.29 ± 0.49	629.43 ± 1505.13	36.24 ± 4.34	5.68 ± 3.60	0.31 ± 0.10
Eclogites	6.22 ± 0.25	3.165 ± 0.195	5.42 ± 0.69	9452.555 ± 4022.895	5.855 ± 1.175	29.79 ± 11.7	1.175 ± 0.435
Amphibolite	5.93 ± 0.43	2.78 ± 0.16	4.60 ± 1.08	20,603.77 ± 31,140.62	7.58 ± 3.78	57.03 ± 26.74	2.32 ± 0.94
Paragneiss	5.66 ± 0.32	2.67 ± 0.13	3.37 ± 0.68	20,254.76 ± 28,110.32	2.87 ± 3.24	83.80 ± 23.93	2.86 ± 0.80
Orthogneiss	5.81 ± 0.33	2.55 ± 0.10	3.92 ± 1.87	30,813.56 ± 33,550.82	−0.29 ± 1.93	140.81 ± 31.62	4.72 ± 0.87

Table 2
10-Fold cross-validation technique structure.

Sequence of matching data along the CCSD-MH										
Subset	First trial Subset 1	Second trial Subset 2	Third trial Subset 3	Fourth trial Subset 4	Fifth trial Subset 5	Sixth trial Subset 6	Seventh trial Subset 7	Eighth trial Subset 8	Ninth trial Subset 9	Tenth trial Subset 10
Interval depth										
100–521 m (3332)	Testing	Training	Training	Training	Training	Training	Training	Training	Training	Training
521–942 m (3332)	Training	Testing	Training	Training	Training	Training	Training	Training	Training	Training
942–1365 m (3332)	Training	Training	Testing	Training	Training	Training	Training	Training	Training	Training
1365–1831 m (3332)	Training	Training	Training	Testing	Training	Training	Training	Training	Training	Training
1831–2291 m (3333)	Training	Training	Training	Training	Testing	Training	Training	Training	Training	Training
2291–2729 m (3333)	Training	Training	Training	Training	Training	Testing	Training	Training	Training	Training
2729–3206 m (3333)	Training	Training	Training	Training	Training	Training	Testing	Training	Training	Training
3206–3866 m (3333)	Training	Training	Training	Training	Training	Training	Training	Testing	Training	Training
3866–4519 m (3333)	Training	Training	Training	Training	Training	Training	Training	Training	Testing	Training
4519–5025 m (3333)	Training	Training	Training	Training	Training	Training	Training	Training	Training	Testing

(Table 4). As mentioned earlier, the aim of ANN is to generalize effectively. Generalization gives us a more convincing estimate of the validity of the ANN. Therefore, generalization is taken as the analytical accuracy using the neural classifiers in this study. Based on this, we can say that, SOM performed slightly better than FFNN in classifying crystalline metamorphic rocks whether RD, CNL, V*p*, DEN, PE, GR, and K logs are used as input pattern.

To test the sensitivity (statistical significance) of the classifiers at 5% threshold in classification decision, the statistical data analysis Student's paired t-test (Hsu and Lachenbruch, 2008, 2014) was used. The most typical use of this statistical tool is to compare means as being used in this study, thus, to test whether there are real differences between the two neural classifiers in crystalline rock classification, based on the average testing correct classification accuracy.

Let's suppose: μ_1 = average (mean) is the correct SOM classification accuracy and μ_2 = average is the correct FFNN classification accuracy across all testing subsets. The hypotheses for paired t-test could be written as follows:

Null hypothesis (H_0) H_0: $\mu_1 - \mu_2 = 0$ (there is no real difference between the two means)
Alternative hypothesis (H_a) H_a: $\mu_1 - \mu_2 \neq 0$ (there is real difference between the two means)

The results obtained from Kolmogorov–Smirnov (Kolmogorov, 1933; Smirnov, 1939) [KS-statistic = 0. 176; *p-value* = 0.20; *p-value* > 0.05] and Shapiro–Wilk (Shapiro and Wilk, 1965) [SW-statistic = 0.894; *p-value* = 0.189; p-*value* > 0.05] nonparametric tests indicated and confirmed that the assumption of normality is met in data for Student paired sample

Table 3
10-Fold cross-validation results.

Methods		Subset 1		Subset 2		Subset 3		Subset 4		Subset 5	
		Training (29,994)	Testing (3332)	Training (29,994)	Testing (3332)	Training (29,994)	Testing (3332)	Training (29,994)	Testing (3332)	Training (29,993)	Testing (3333)
SOM	#Correct	24,481	2677	23,437	2580	24,490	2677	23,656	2603	24,363	2688
	#Incorrect	5513	655	6557	752	5504	655	6338	729	5630	645
	%Correct	81.62	80.34	78.14	77.44	81.65	80.34	78.87	78.12	81.23	80.64
FFNN	#Correct	26,823	2723	24,391	2503	24,061	2357	23,731	2017	21,058	1506
	#Incorrect	3171	609	5603	829	5933	975	6263	1315	8935	1827
	%Correct	89.43	81.72	81.32	75.12	80.22	70.74	79.12	60.53	70.21	45.18
	NHN	40		44		45		40		44	

Methods		Subset 6		Subset 7		Subset 8		Subset 9		Subset 10	
		Training (29,993)	Testing (3333)	Training (29,993)	Testing (3333)	Training (29,993)	Testing (3333)	Training (29,993)	Testing (3333)	Training (29,993)	Testing (3333)
SOM	#Correct	23,361	2593	24,930	2758	24,855	2681	24,024	2660	24,882	2699
	#Incorrect	6632	740	5063	575	5138	652	5969	673	5111	634
	%Correct	77.89	77.79	83.12	82.75	82.87	80.44	80.10	79.80	82.96	80.97
FFNN[a]	#Correct	23,622	2179	24,024	2604	25,527	2667	27,689	2884	27,356	2838
	#Incorrect	6371	1154	5969	729	4466	666	2304	449	2637	495
	%Correct	78.76	65.37	80.10	78.13	85.11	80.02	92.32	86.53	91.21	85.15
	NHN	46		44		40		46		44	

[a] Several FFNN models were trained and their performances were tested. Only the optimal models are included in this study. An optimal model was identified based on the highest correct classification rate in terms of the testing subset result. Note that #Correct illustrates the number of correct classification; it is defined by the number of data points classified by the computational neural network which matches those of core data. #Incorrect indicates the number of incorrect classifications; it is represented by the number of data points classified by the computational neural network which mismatches those of core data. %Correct shows the percentage of correct classification. The values between parentheses ( ) indicate the number of available data points. NHN shows the number of hidden nodes. SOM mean Self-Organizing Map algorithm while FFNN indicates Feed Forward Neural Network with scaled conjugate gradient algorithm.

Table 4
Average (mean) correct classification rates for SOM and FFNN across the 10 subsets using CCSD-MH logging data.

Statistic	Training		Testing	
	SOM	FFNN	SOM	FFNN
Mean (%)	80.84	82.78	79.86	72.84

t-test. For more detailed description of nonparametric tests and the underlying concepts, reference may be made to Conover (1999), and Razali, and Wah (2011).

The application of paired *t*-test shows that, the *p*-value for the difference between the two mean *t*-test is $p = 0.110$ (t-statistic $= 1.770$). Since this *p-value* is greater than 0.05, the decision would be that there is no statistically significant difference between the two methods (H_0 should not be rejected). This finding does not necessarily mean that there is no meaningful difference between the two methods in the actual situation, but there is not enough evidence to conclude that the difference about 7% between the two methods in terms of average classification accuracy in the context of lithology prediction is statistically significant at 5% level. In other words there is less confidence to conclude that SOM superiority is statistically significant over FFNN in crystalline metamorphic rocks classification tasks. The conclusion to be drawn is that Student paired sample *t*-test failed to bring out a statistically reliable difference between the two methods on CCSDMH data at 5% level thereby suggesting that the null hypothesis (H_0) may be true. That is, there is no real difference between the two computational neural networks in crystalline rocks classification task using CCSD-MH log data.

Making reference to testing on the whole data, the resulting lithology obtained by the two computational neural classifiers and corresponding core description are shown in Fig. 4. It was observed that the predicted lithologies match satisfactorily with the actual lithologies from core data. All lithologies detected in the core data are present in predicted lithologies by the two methods. The correct classification archived by SOM was 83, 25% (27,744 of 33,326 points), comparable to 81.02% (27,001 of 33,326 points; 46 hidden neurons) archived by FFNN.

A detailed performance of the two neural classifiers into each category of metamorphic rocks using the whole data is reported in Fig. 5. It follows from Fig. 5, the highest correct classification rates is obtained for ultramafic rocks (92–94%). In contrary, the lowest correct classification rate is achieved for amphibolite (47–58%). This may be due to complex geologic features of these rocks, showing that the variation of mineral composition in amphibolite is not relatively stable. This fact may have been more difficult for both SOM and FFNN to recognize amphibolite in an optimal way. The correct classification rates in orthogneiss, eclogite and paragneiss rocks run from 92 to 93%, from 84 to 87% and from 57 to 69% respectively.

Again from Fig. 5, the correct classification rate bars depict that both SOM and FFNN provide satisfactory results in predicting crystalline rocks from CCSD-MH data. Practically the same pattern of correct classification rate bars for SOM is comparable to FFNN as well with SOM offering slightly better classification in orthogneiss, eclogite and paragneiss rocks. One the other hand, FFNN shows a slightly better classification in ultramafic and amphibolite rocks. These results are indicating that the use of logging data in classifying crystalline rocks is very significant feature. SOM and FFNN confirmed that even if only geophysical log data are available an adequate lithology interpretation is still feasible.

It was of interest to compare our results with those of Pan et. al (2010). In their study, FFNN-BP was used and it was found to be more accurate in training sample. The FFNN-BP reached 95.27% correct classification rates in predicting crystalline rocks at CCSD-MH. However, the robustness of the model was not tested. In other words, the authors have given more attention to the model selection i.e. training but have not taken the performance of the model on independent dataset into account in their analysis. The basic question which must be answered is that: how robust was their model performance in predicting crystalline rocks?

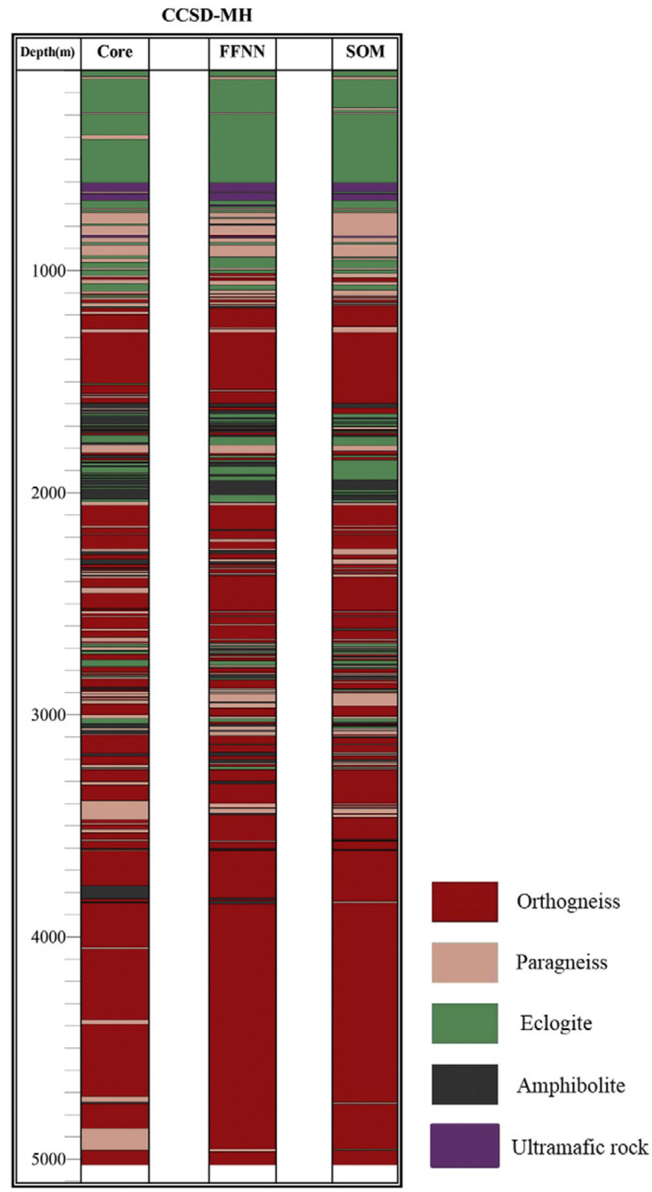

Fig. 4. Lithology obtained by core descriptions FFNN and SOM, respectively.

Guidance to the answers to the question shows that their work has limitations in terms of generalization ability. Since it cannot be guaranteed that the model is generalized well. Poor generalization can be characterized by the phenomenon of over-fitting: the model describes the trained data well, but fails to generalize into new data. The test on new data aims at evaluating the accuracy of the results predicted by the neural network.

As mentioned before, to use ANN as a prediction tool, a model must be able to produce sensible output on new data. Model robustness is important when the model is used for classification tasks as in this study. On the issue of ANN development, Poulton (2002) points out that: “I often advise people that training a feed-forward neural network is about 10% of the effort involved in an application; deciding on the input and output data coding and creating good training and testing sets is 90%”. Based on the aforementioned, this current study has employed 10-fold cross validation to investigate the robustness of the two neural classifiers in lithology classification with respect to sampling

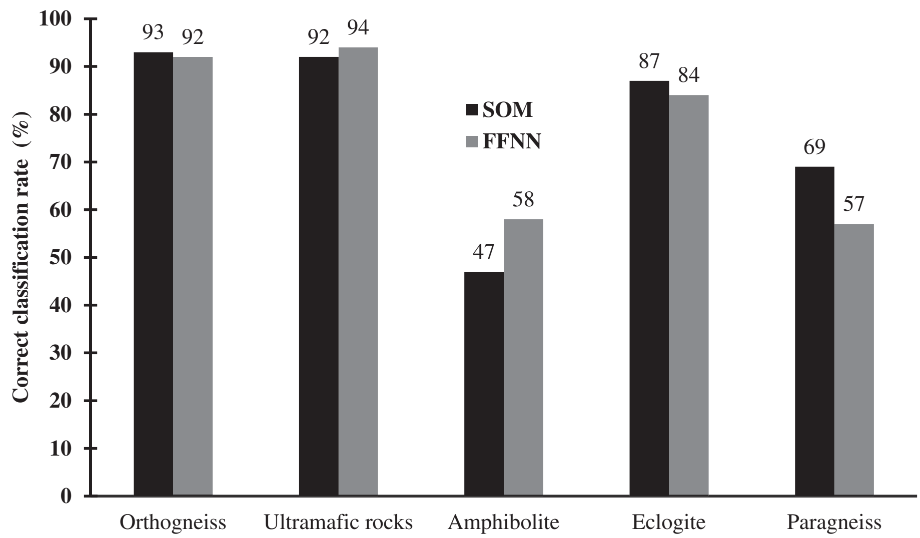

Fig. 5. Performance of SOM and FFNN into each category of metamorphic rocks from CCSD-MH using whole data.

rotation in crystalline rocks. Furthermore, the two computational methods were also tested using the whole dataset. Moreover, we feel that in this study, sufficient data points were used (33,326 data points) to assure adequate ANN training.

6. Summary and conclusion

The identification of lithology is fundamental for understanding the subsurface. It plays an important role on the description of the rocks and their classification.

In this study, an investigation into the applicability of the existing classification methods in literature on crystalline rocks using geophysical well log data from CCSD-MH data were carried out. In light of this, we investigated SOM performance and compared to FFNN performance in crystalline rock classification. The statistical resampling technique and hypothesis testing paired test sample were used to assess the classification accuracy.

This study has demonstrated that complex nonlinear geophysical log data from crystalline rocks can be processed by artificial intelligence model for lithology identification tasks; and they have the potential of improving lithology classification accuracy from geophysical well logs. Several authors such as Maiti et al. (2007), Maiti and Tiwari (2009), Maiti and Tiwari (2010) and Bosch et al. (2013) have also proven in their studies that artificial intelligence can be integrated as a powerful tool for analyzing complex nonlinear geophysical log data from crystalline rocks.

From the overall results in this study, the consensus of both methods is helpful for the lithology classification using geophysical well logs (RD, CNL, V*p*, DEN, PE, GR, and K) from crystalline rocks. The performance of SOM and FFNN are reasonably comparable. In the sense the magnitude of difference between SOM and FFNN is not statistically significant. It is interesting to note that SOM was able to achieve impressive accuracy rate for classifying complex changes in the lithology successions of CCSD-MH. This was done without any prior information (core description) in its learning mechanism but instead it has tried to discover and classify structure having similar characteristics in large, complex and nonlinear well log data; which is visibly a much more difficult task. This advantage of self-organizing may be a very important factor in the cases when no core data is available, providing on-line real time responses to geophysicists during the exploration phase. The SOM classifier has the latent benefit that geophysicists may understand the explicit explanation between geophysical logs and lithology behind it better than they do for the FFNN, which is habitually seen as a “Black box”. Additionally, SOM can avoid the unresolved problem of number of hidden layer and hidden neuron in building an FFNN (Poulton, 2002). SOM may therefore serve as a good practical alternative technology in crystalline rocks predictions task in real-time-based drilling research. It is reasonable to say that no work ever covers all of the research that the researchers envisaged covering when they commenced the project. The same is true for this study. The main theme of this study has been to consider the ANNs as a method of recognizing lithology in the context of crystalline rocks using geophysical logs. However, there is evidence from previous works that other methodologies can be applied for lithology identification tasks. For this reason, we see for example, fractal analyses, wavelet transform in which promising new research can be done. They are very appropriate for the analysis of geophysical data as they inherently identify incremental information across scales. Fractal analyses and wavelet transform capabilities have been proven in several studies (Ouadfeul 2012;Perez-Munoz et al. 2013 and references therein; Karacan et al. 2014). However, these methods may reach their limits when being applied to relatively short and noisy geophysical data. The present study was done using a single deep well (CCSD-MH). Lacking at least two wells, we were motived to use cross validation analysis. This technique is well known to give valuable understandings on the reliability of the models with regard to sampling rotation. However, it gets more interesting in applying our models to other geologically complex zone of interest. We see additionally this scenario in future research works.

Acknowledgment

The work was supported by the National Natural Science Foundation of China Project Fund No. 41074086.Research on model experiments, forward and inversion of induced polarization in ground - well mode. 2011/01-2013/12.

References

Baldwin, J.L., Bateman, A.R.M., Wheatley, C.L., 1990. Application of neural networks to the problem of mineral identification from well-logs. Log Anal. 3, 279–293.

Bartetzko, A., Delius, H., Pechnig, R., 2005. Effect of compositional and structural variations on log responses of igneous and metamorphic rocks. I: mafic rocks. In: Harvey, P.K., Brewer, T.S., Pezard, P.A., Petrov, V.A. (Eds.), Petrophysical Properties of Crystalline Rocks. Geological Society, London, Special Publications, pp. 255–278.

Benaouda, D., Ferrier, G., Whitmarsh, R.B., Rothwell, R.G., MacLeod, C., 1999. Inferring lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the Ocean Drilling Program. Geophys. J. Int. 136, 477–491.

Bhatt, A., Helle, H.B., 2002. Determination of facies from well logs using modular neural networks. Pet. Geosci. 8, 217–228.

Bishop, C.M., 1995. Neural networks for pattern recognition. Oxford Press.

Bosch, D., Ledo, J., Queralt, P., 2013. Fuzzy logic determination of lithologies from well log data: application to the KTB project data set (Germany). Surv. Geophys. 34 (4), 413–439.
Briqueu, L., Gottib-Zeh, S., Ramadan, M., Brulhet, J., 2002. Traitement des diagraphies à l'aide d'un réseau de neurones du type «carte auto-organisatrice»: application à l'étude lithologique de la couche silteuse de Marcoule (Gard, France). C.R. Geoscience 334, 331–337.
Busch, J.M., Fortney, W.G., Berry, L.N., 1987. Determination of lithology from well logs by statistical analysis. Soc. Petrol. Eng. Form. Eval. 2, 412–418.
Chang, H.C., Merkel Kopaska, D.C., Chen, H.C., Durrans, S.R., 2000. Lithofacies identification using multiple adaptive resonance theory neural networks and group decision expert system. Comput. Geosci. 26, 591–601.
Chang, H.C., Kopaska-Merkelb, D.C., Chen, H.C., 2002. Identification of lithofacies using Kohonen self-organizing maps. Comput. Geosci. 28, 223–229.
Conover, W.J., 1999. Practical nonparametric statistics. Third Edition. John Wiley & Sons, New York.
Delfiner, P., Peyret, O., Serra, O., 1987. Automatic determination of lithology from well logs. Soc. Petrol. Eng. Form. Eval. 2, 303–310.
Harvey, P.K., Brewer, T.S., Pezard, P.A., Petrov, V.A. (Eds.), 2005. Petrophysical Properties of Crystalline Rocks. Geological Society, London, Special Publications, p. vii.
Haykin, S., 1999. Neural networks: a comprehensive foundation. Prentice-Hall, New Jersey.
He, L.J., Hu, S.B., Yang, W.C., Wang, J.Y., Yang, S.C., Yuan, Y.S., Cheng, Z., 2006. Temperature measurement in the main hole of the Chinese Continental Scientific Drilling. Chin. J. Geophys. 49 (3), 671–678.
Hsu, H., Lachenbruch, P.A., 2008. Paired t test. Wiley Encyclopedia of Clinical Trialspp. 1–3 http://dx.doi.org/10.1002/9780471462422.eoct969.
Hsu, H., Lachenbruch, P.A., 2014. Paired t test. Wiley StatsRef: Statistics Reference Online http://dx.doi.org/10.1002/9781118445112.stat05929.
Jing, J., Wei, W., Jin, S., Ye, G., Deng, M., 2007. A study on the classification and well-logging identification of eclogite in the main hole of Chinese continental scientific drilling project. J. China Univ. Geosci. 18 (4), 357–365.
Karacan, C.O., Olea, R.A., 2014. Inference of strata separation and gas emission paths in longwall overburden using continuous wavelet transform of well logs and geostatistical simulation. J. Appl. Geophys. 105, 147–158.
Kohonen, T., 2001. Self-organizing maps. 3rd Edition. Springer, Berlin.
Kolmogorov, A., 1933. Sulla determinazione empirica di una legge di distribuzionc, 1st. Ital. Attuari. G. 4 pp. 1–11.
Luo, M., Pan, H.P., 2010. Well logging responses of UHP metamorphic rocks from CCSD main hole IN Sulu Terrane, Eastern Central China. J. Earth Sci. 21 (3), 347–357. http://dx.doi.org/10.1007/s12583-010-0098-9.
Maiti, S., Tiwari, R.K., 2009. A hybrid Monte Carlo method based artificial neural networks approach for rock boundaries identification: a case study from KTB Borehole. Pure Appl. Geophys. 166, 2059–2090. http://dx.doi.org/10.1007/s00024-009-0533-y.
Maiti, S., Tiwari, R.K., 2010. Neural network modelling and an uncertainty analysis in Bayesian framework: a case study from the KTB borehole site. J. Geophys. Res. Solid Earth http://dx.doi.org/10.1029/2010JB000864.
Maiti, S., Tiwari, R.K., Kumpel, H.J., 2007. Neural network modelling and classification of litho-facies using well log data: a case study from KTB borehole site. Geophys. J. Int. 169, 733–746. http://dx.doi.org/10.1111/j.1365-246X.2007.03342.x.
May, R.J., Maier, H.R., Dandy, G.C., 2010. Data splitting for artificial neural networks using SOM-based stratified sampling. Neural Netw. 23 (2), 283–294.
Moller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. 6, 525–533.
Niu, X.Y., Pan, H.P., Wang, W.X., Zhu, L.F., Xu, D.H., 2004. Geophysical well logging in main hole (0–2 000 m) of Chinese Continental Scientific Drilling. Acta Petrol. Sin. 20 (1), 109–118 (in Chinese with English abstract).
Oja, E., 2002. Unsupervised learning in neural computation. Theor. Comput. Sci. 287, 187–207.
Ouadfeul, S.A (Ed.), 2012. Fractal Analysis and Chaos in Geosciences. InTech. ISBN: 978-953-51-0729-3. http://dx.doi.org/10.5772/3309.
Ouadfeul, S.A., Aliouane, L., 2012. Lithofacies classification using the multilayer perceptron and the self-organizing neural networks. Neural Information Processing: 19th international conference. Lecture Notes in Computer Science 7667. ICONIP, pp. 737–744.
Pan, H.P., Niu, Y.X., Wang, W.X., 2002. CCSD well logging engineering program. J. China Univ. Geosci. 13 (1), 91–94.
Pan, H.P., Niu, Y.X., Wang, W.X., 2005. Radioactive logging application in CCSD main hole. Earth Sci. J. China Univ. Geosci. 30 (Suppl.), 49–56.
Pan, H.P., Luo, M., Zhao, Y., 2010. Identification of metamorphic rocks in the CCSD Main Hole. In: Yue, S., Wei, H.L., Wang, L., Song, Y. (Eds.), Sixth International Conference on Natural Computation. IEEE, Eds, pp. 4049–4051.
Pechnig, R., Haverkamp, S., Wohlenberg, J., Zimmermann, G., Burkhardt, H., 1997. Integrated log interpretation in the German Continental Deep Drilling Program: lithology, porosity, and fracture zones. J. Geophys. Res. 102 (B8), 18363–18390.
Pechnig, R., Bartetzko, A., Delius, H., 2001. Effects of compositional and structural variations on log responses in igneous and metamorphic rocks. AGU Fall Meeting (San Francisco. Abstract V32C-0988).
Pechnig, R., Delius, H., Bartetzko, A., 2005. Effect of compositional variations on log responses of igneous and metamorphic rocks. II: acid and intermediate rocks. In: Harvey, P.K., Brewer, T.S., Pezard, P.A., Petrov, V.A. (Eds.), Petrophysical Properties of Crystalline Rocks. Geological Society, London , Special Publications, pp. 279–300.
Perez-Munoz, T., Velasco-Hernandez, J., Hernandez-Martinez, E., 2013. Wavelet transform analysis for lithological characteristics identification in siliciclastic oil fields. J. Appl. Geophys. 98, 298–308.
Poulton, M.M., 2002. Neural networks as an intelligence amplification tool: a review of applications. Geophysics 67, 979–993.
Pratsone, E.L., Anderson, R.N., Dove, R.E., Lyle, M., Silver, L.T., James, E.J., Chappell, B.W., 1992. Geochemical logging in the cajon pass drill hole and its application to a new, oxide, igneous rock classification scheme. J. Geophys. Res. 97 (B4), 5167–5180.
Razali, N.M., Wah, Y.B., 2011. Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lillefors and Anderson Darling tests. J. Stat. Model. Anal. 2 (1), 21–33.
Reitermanova, Z., 2010. Data splitting. In: Safrankova, J., Pavlu, J. (Eds.), WDS'10 Proceedings of contributed papers, part I–Mathematics and Computer Sciences. Matfyzpress, Prague, pp. 31–36 (ISBN 978-80-7378-139-2).
Rodriguez, J.D., Perez, A., Lozano, J.A., 2010. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans. Pattern Anal. Mach. Intell. 32 (3), 569–575. http://dx.doi.org/10.1109/TPAMI.2009.187.
Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52 (3/4), 591–611.
Smirnov, N., 1939. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. Bull. Math. Univ. Mosc. 2 (2), 3–14.
Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. J. R. Stat. Soc. Ser. B Methodol. 36 (2), 111–147.
Twomey, J.M., Smith, A.E., 1997. Validation and verification. In: Kartam, N., Flood, I., Garrett, J. (Eds.), Artificial Neural Networks for Civil Engineers: Fundamentals and Applications. ASCE Press, New York, pp. 44–64.
Van der Baan, M., Jutten, C., 2000. Neural networks in geophysical applications. Geophysics 65, 1032–1047.
Wang, W.X., Niu, Y.Z., Pan, H.P., 2005. Well logging analysis of ultrabasic rock section in main borehole of Chinese Continental Scientific Drilling. Earth Sci. J. China Univ. Geosci. 30 (Suppl.), 83–87.
Wolff, M., Pelissier-Combescure, J., 1982. FACIOLOG: automatic electrofacies determination. SPWLA annual logging symposium, pp. 6–9 (paper FF).
Xu, H.J., Jin, Z.M., Ou, X.G., 2006. Lithology determination of rocks from CCSD 100–2 000 m main hole by magnetic susceptibility and density using discriminant function analysis. Earth Sci. J. China Univ. Geosci. 31 (4), 513–519 (in Chinese with English Abstract).
Xu, Z., Yang, W., Ji, S., Zhang, Z., Yang, J., Wang, Q., Tang, Z., 2009. Deep root of a continent–continent collision belt: evidence from the Chinese Continental Scientific Drilling (CCSD) deep borehole in the Sulu ultrahigh-pressure (HP–UHP) metamorphic terrane, China. Tectonophysics 475, 204–219.
Yang, W.C., 2009. The crust and upper mantle of the Sulu UHPM belt. Tectonophysics 475, 226–234.
Zhang, Y., Salisch, H.A., McPherson, J.G., 1999. Application of neural networks to identify lithofacies from well logs. Explor. Geophys. 30, 45–49.