

Prediction of rock strength properties by data mining approach

G. S. Terra¹, M. C. A. Costa², N. F. F. Ebecken²

¹*CEFET-Campos/UNED*

²*COPPE/ UFRJ*

Abstract

Oil reservoir rocks can be characterized by both dynamic and static parameters. Dynamic parameters are derived from in situ tests, whilst static parameters are derived from laboratorials tests (generally destructive), which are carried on samples from the wells.

Despite the fact that dynamic and static tests generate distinct parameters, it is possible to establish relationships between these different parameters, so that one can obtain the static parameters without carrying static tests and therefore derive elastic properties of the material, such as elastic and shear moduli, etc.

In the last years the development of computational intelligence techniques has been growing very fast. They have been used successfully to evidence the relations of attributes of different sources. Some of these relations could not be easily obtained by mathematical formulation. The use of some techniques leads to new concepts and insights about data relationship. Neural networks(NN) is one of the most powerful techniques for describing relationship among data. Neural networks can predict values in a very precise way, even in the presence of few or missing attributes. It also has the ability of perceiving the influence the variation of one attribute has on another.

In the present work, data minig techniques are used to extract these relationships and predict elastic properties from dynamic parameters.

1 Introduction

Static tests are performed on samples extracted during the dynamics tests and determine rupture stress, shear modulus, Poisson and Young modulus. To extract

the samples we need to stop the dynamic test and this is very inconvenient, making limitations to the sampling process.

The dynamic properties are usually evaluated many times each meter, and the static properties are rarely determined. This is the main reason to try to establish precise tools to correlate them. Evolving neural network architectures it is possible to obtain very good estimators of static parameters derived from a dynamic parameter database. In this work the parameters database includes two types of litologies: sandstone (68 samples, 16 wells) and calcareous (28 samples, 7 wells). Table I summarizes the considered static and dynamic properties

Table 1 – Data Description

Name	Type	I/O	Description	Unit
Well	string		Well	
Depth	number		Depth	M
P st	number	I	static confined pressure	MPa
P dy	number	I	dynamic confined pressure	MPa
VP	number	I	shear wave velocity P	m/s
VS	number	I	shear wave velocity S	m/s
Rho	number	I	dynamic density	g/cc
PHI	number	I	dynamic porosity	%
Lit	number	I	Litology (sandstone and calcareous)	
Srup	number	O	rupture stress	MPa
Poisson st	number	O	Static Poisson modulus	GPa
Young st	number	O	Static Young modulus	Gpa
Shear st	number	O	Static Shear modulus	Gpa

The paper is organized as follows. In Section 2, we give a brief overview of Artificial Neural Networks model. Section 3 discusses characteristics of the data and the performance of the proposed ANN model. Conclusions are drawn in Section 4.

2 Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) (as compared to real ones) are mathematical systems that are comprised of a number of “processing units” that are linked via weighted interconnections. A processing unit is essentially an equation, which is often referred to as a “transfer function”. A processing unit takes weighted signals from other neurons, possibly combines them, transforms them and outputs a numeric result. Processing units are often considered crudely analogous to real neurons and since they are linked together in a mesh or network, the name Neural Networks was coined.

Many neural networks have their neurons structured in “layers” that have similar characteristics and execute their transfer functions in synchronization (at

the same time, virtually speaking). Nearly all neural networks have neurons that accept data and neurons that produce outputs.

The behavior of neural networks, how they map input data to output data, is influenced primarily by the transfer functions of neurons, how they are interconnected and the weights of those interconnections.

Typically, an architecture or structure of a neural network is established and one of a variety of mathematical algorithms are used to determine what the weights of the interconnections should be to maximize the accuracy of the outputs produced. Neural networks are “trained”, meaning they use previous examples to establish (learn) the relationships between the input variables and the predicted variables by setting these weights. Once these relationships are established (the neural network is trained), the neural can be presented with new input variables and it will generate predictions.

The applications uses Genetic Algorithms to evolve neural network structures while simultaneously searching for significant input variables to maximize the predictive accuracy of the resulting neural network models.

The effectiveness of the genetic algorithms’ capabilities cannot be overstated. For example, a problem consisting of finding the best combination (subset) of 20 inputs and up to 15 hidden nodes in a back propagation neural network is a combinatorial problem with over 16 million permutations. To train a network in a full search of all permutations would be a good project for a super computer. But with genetic algorithms, an excellent solution often appears in less than 1500 evaluations, which is 0.009% of the total possible configurations.

Using some statistical data analysis to assist, highly fit networks are often found in the first 30 to 50 neural networks evaluated. This is clearly an efficient means for discovering effective network structure/input combinations.

Several tests were performed and some results are discussed in the following.

3 Model prediction

In table I only a subset of the dynamic parameters obtained in *in situ* test are considered. Missing values and the small number of samples can justify the choice of this subset. It was selected in order to keep the biggest number of samples.

In this subset, the shear wave velocity P and S are very correlated (0.96) and this can be very evident during the training activities of the neural network models to predict rupture stress. Figure I indicates VP and VS for the 80 samples that were taken into account.

The correlation function was chosen as a performance measure of the networks related to the training and testing data during the training task. Some other indicators were considered and this is commented in the text that follows.

The rupture stress presents an asymptotic behavior when we plot this value against the static confined pressure. The friction angle and cohesion coefficient are obtained from this plot. This characteristic was used to eliminate some NN models, that means, high correlated and small error are not sufficient. We need to

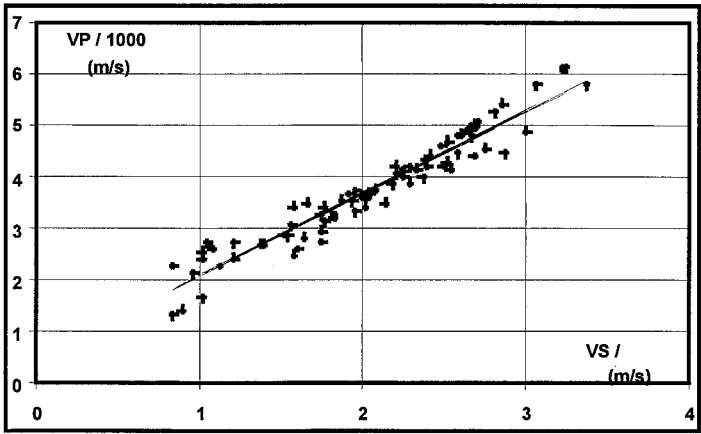


Figure 1 - Comparison of VP and VS

approximate the asymptotic behavior. Figure 2 shows this behavior. The two experimental values correspond to samples of the same well. The curves were generated by neural network predictions for different confinant pressures. The obtained values in the same laboratories pressure conditions are also indicated. This allows that the NN models can be use to predict values that cannot be monitored in the sample because the destructive aspects.

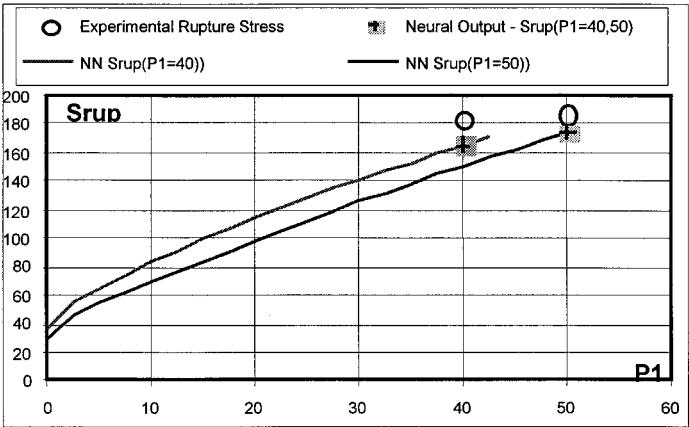


Figure 2 - Comparison of Srup and P1

- Two other criteria were adopted to select the best NN:
- motivated by the small number of samples and irregular distribution the NN models that predict extreme values are preferred
 - the NN results were compared with existing formulas from the literature, as indicated in Figure 3.

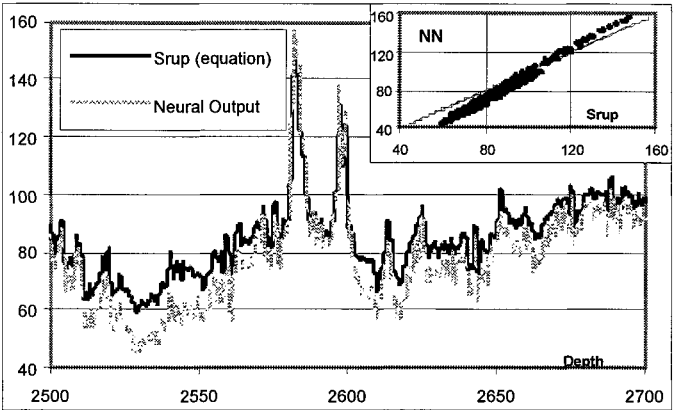


Figure 3 - Comparison of neural output and equation

The different NN architectures were shown in table 2. The neuron number in the input layer may be different from the number of variables because they were presented to the NN models after linear and nonlinear data transformations.

Table 2 – Architecture Obtained by the Genetic Algorithm

Output	Input							Architecture Input-hidden-output
	P1	P2	VP	VS	Rho	Phi	Lit	
Rupture Stress 1	x	x	x		x	x	x	5-9-1
Rupture Stress 2		x	x	x	x	x	x	6-9-1
Young st	x	x	x	x	x	x		5-8-1
Shear st		x	x		x	x		11-17-1
Poisson st	x	x	x		x	x		10-16-1

The greater neuron numbers in the input and hidden layers of the last NN models may be related to the complexity level intended to be approximate.

Figures 4 and 5 shows the NN models 1 and 2 obtained results for rupture stress. The used training, test and validation data are also indicated. Sixteen samples were removed from the database to generate the second NN models because the shear wave velocity S values were missing.

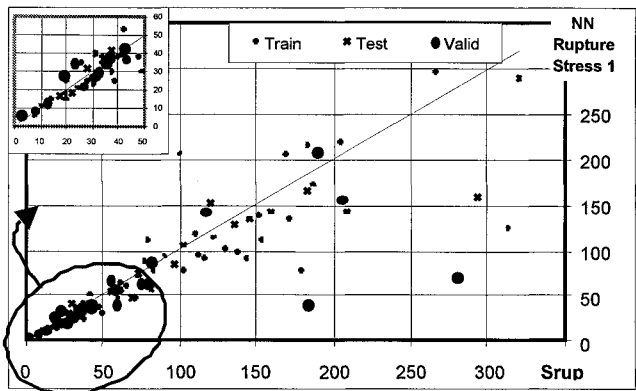


Figure 4 - Comparison of NN Srup 1 and Srup

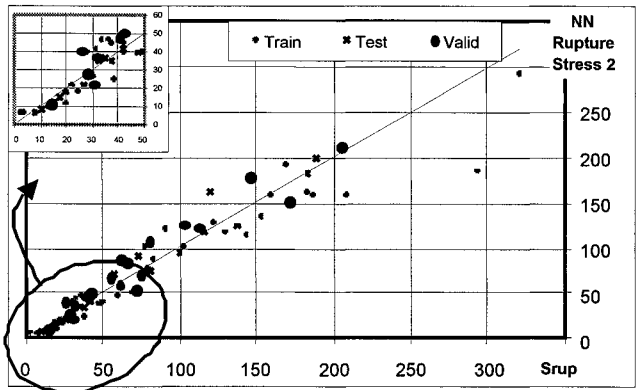


Figure 5 - Comparison of NN Srup 2 and Srup

Figures 6 shows the obtained values for the Young modulus.

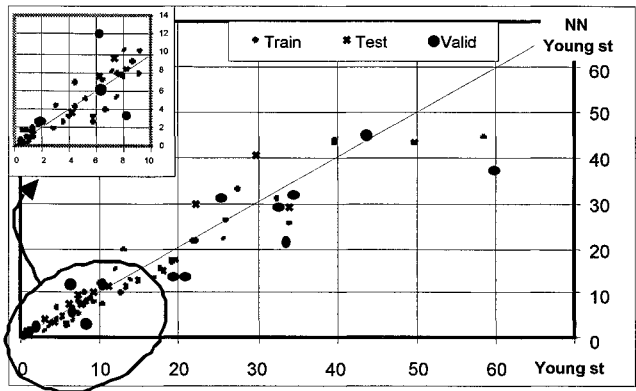


Figure 6 - Comparison of NN Young st and Young st

Figures 7 shows the obtained shear values.

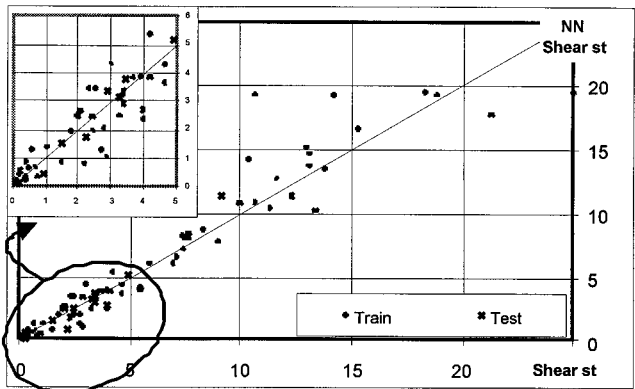


Figure 7 - Comparison of NN Shear st and Shear st

Figures 8 indicates the Poisson coefficients.

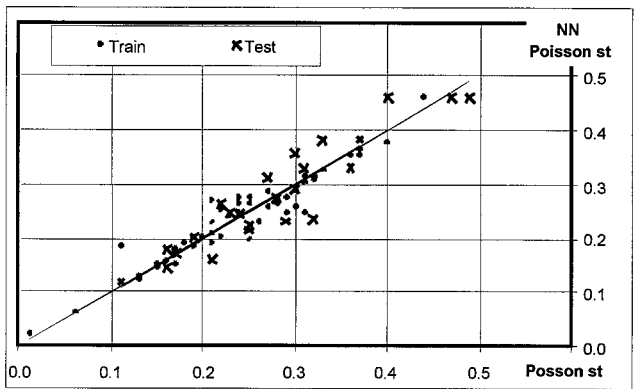


Figure 8 - Comparison of NN Poisson st and Poisson st

The comparison of the results are summarized in Table 3.

Table 3 – Linear Correlation and RMS

NN Output	Linear Correlation			RMS		
	Training	Test	Valid	Training	Test	Valid
Rupture Stress 1	0.839	0.938	0.680	38.4	33.7	54.6
Rupture Stress 2	0.963	0.980	0.978	22.5	12.1	19.5
Young st	0.967	0.958	0.900	3.3	3.6	7.9
Shear st	0.939	0.962	0.944	2.1	1.3	1.9
Poisson st	0.964	0.930	0.958	0.024	0.036	0.027

Despite the better performance of NN model 2, some data were removed because they were associated to wells with very peculiar characteristics.

Based on the described approach a software system called Neurock was developed and implemented on the Windows platform using Visual Basic.

Neurock selects the required well properties directly from the geological database and can visualize the prediction along the depth. The Mohr-Coulomb criteria was also included.

4 Conclusion

The search space of all the possible NN configurations to predict mechanics properties of rock is very large. We cannot guarantee that the obtained NN are the best choice. But some aspects induce that we have generate relevant results.

The NN models to predict rupture stress have shown a superior correlation when compared with statistical equations. Because they take into account the initial confinant pressure, they can reproduce with fidelity the laboratories tests.

The rupture stress x confinant pressure plots are generated and the friction angle and cohesion may be calculated.

The reduced number of samples considered are justified by the need of very expensive experiments. The number os samples with high rupture stress values was reduced.

As NN are an example learning methodology the selection of the most adequate architecture considered the correlation results, the evaluation with the statistical formulas, the rupture stress x confinant pressure diagram and the best produced accuracy.

To the elastic modulus, shear modulus and Poisson coefficient the selection criteria of NN models have corresponded to high correlation and dynamic equation comparison.

More than 150 NN were generated and we can conclude that high correlated and small RMS were generated and they satisfied the different criteria that were proposed.

References

- [1] "Advances in Knowledge Discovery and Data Mining", Fayyad, U.M., Piatetsky-Shapiro, G., Smith, P. and Uthurusamy, R., MIT Press, 1996
- [2] "C4.5: Programs for Machine Learning", Quinlan, J.R., Morgan Kaufmann Publishers, 1993
- [3] "Intelligent Miner for AIX; User's Guide", IBM Corp., 1996
- [4] "NeuralWorks Predict", Neural Ware Inc., 1997.
- [5] "Neuro Genetic Optimizer", Bio Comp Systems Inc., 1998
- [6] "NeuroShell 2", Ward Systems Group Inc., 1993
- [7] "Solving Data Mining Problems through Pattern Recognition", Kennedy, R.L., Lee, Y., Van Roy, B., Reed, C.D. and Lippman, R.P., Prentice Hall Ptr, 1997

- [8] "Data Mining", edited by N.F.F.Ebecken, WIT Press, Southampton, UK, 1998
- [9] "Data Mining II", edited by N.F.F.Ebecken and C.A.Brebbia, WIT Press, Southampton, UK, 2000
- [10] "Evaluation of the Cohesion and friction angle of oil reservoir rocks by dynamic tests", Soares, J.A., Technical Report, CENPES/DIGER/SEGRES, 1999.
- [11] "Neural Networks - A comprehensive Foundation", Haykin, S., 2 ed., Prentice Hall, 1999
- [12] "Neural Computing: Theory and Practice", Wasserman, P. D., New York: Van Nostrand Reinhold, 1989
- [13] "Soil Mechanics", Lambe, T.W., Whitman, R.V., SI version, John Wiley & Sons, 1979