Permeability Determination From Well Log Data

Shahab Mohaghegh, SPE, **Bogdan Balan,** SPE, and **Samuel Ameri,** SPE, West Virginia U.

Summary

We discuss and compare three different approaches for permeability determination from logs from a practical point of view. The three methods, empirical, statistical, and the recently introduced "virtual measurement," make use of empirically determined models, multiple variable regression, and artificial neural networks, respectively. We apply all three methods to well log data from a heterogeneous formation and compare the results with core permeability, which is considered to be the standard. Our comparison focuses on the predictive power of each method.

Introduction

Reservoir management strategies are as realistic as the "image" of spatial distribution of rock properties. Permeability is the most difficult property to determine and predict. Many investigators[1-10] have attempted to capture the complexity of permeability function in models with general applicability. While these studies contribute to a better understanding of the factors controlling permeability, they demonstrate that it is an illusion that a "universal" relation between permeability and variables from wireline logs can be found.

The regression approach, which uses statistics instead of "stiff," deterministic formalism, tries to predict a conditional average, or expectation of permeability.[11-15] The newest method, called "virtual measurement,"[16,20] makes use of the artificial neural networks that are model-free function estimators. Neural networks are flexible tools that can learn the patterns of permeability distribution in a particular field and then predict permeability from new data by generalization.

To compare the capabilities of these approaches, we apply all three methods to wireline log data from a heterogeneous oil-bearing formation and compare the results with core-determined permeability, which is considered to be the standard. We test these methods for their model development as well as their predictive capabilities.

Empirical Models

Empirical models are based on the correlation between permeability, porosity, and irreducible water saturation. **Table 1** presents the four empirical models used the most: Tixier, Timur, Coates & Dumanoir, and Coates. All these methods, except Coates & Dumanoir, assume certain values for cementation factor and/or saturation exponent and are applicable to clean sand formations where conditions of residual water saturation exist. Coates & Dumanoir have proposed an improved empirical permeability technique. With the support of core and log studies, they adopted a common exponent, *w*, for both the saturation exponent, *n*, and cementation exponent, *m*. Coates & Dumanoir also presented a method for testing whether the formation is at irreducible water saturation. However, they noted that if the reservoir is heterogeneous, it may fail that test and still be at irreducible water saturation. Their method is the first that satisfies the condition of zero permeability at zero porosity and when $S_{wirr} = 100\%$. Because of the corrections provided, this method can be applied to formations that are not at irreducible water saturation and to shaley formations. Values for the exponents *m* and *n* are not needed because they are found as a result of the computation.

The cementation factor, *m*, and the saturation exponent, *n*, are the biggest sources of uncertainty in permeability determination by means of empirical models.[3] They can be obtained by laboratory measurements, which is seldom the case, or approximated according to some general guidelines and/or experience. Methods for deducting the cementation factor have had a long history. In this study, we use the method based on the establishment of a "water line" in a zone that is 100% water saturated.[15]

Multiple Variable Regression

Multiple regression is an extension of the regression analysis that incorporates additional independent variables in the predictive equation. In this study, the dependent variable is the logarithm of permeability because permeability seems to be log-normal and the independent variables are well log variables. Wendt and Sakurai[12] established a general procedure for permeability prediction by multiple variable regression. They also pointed out the shortcomings of using this technique. When the regression method is used for prediction, the distribution of predicted values is more narrow than that of the original data set. Kendall and Stuart[15] explained this, stating that the regression model "...either exhibits a property of a bivariate distribution or, when the regressor variables are not subject to error, gives the relation between the mean of the dependent variable and the value of the regressor variables." That is, the regression provides the best estimate of the average. The assumption that the error is related only to the dependent variable (permeability measurements) and not to the independent variables (log variables) can be verified by comparing repeat runs of properly calibrated instruments with the main runs of the logs, provided that there is no bias in the measurement. Logs of acceptable quality have errors with a relatively small unbiased scatter that is a function of the physics of the tool, its response characteristics, and the borehole environment. If the deviations are indeed random, then they would be expected to be normally distributed with a mean value of zero.

The correlation matrix of all independent and dependent variables should be analyzed to establish whether there is a dominant independent variable or whether the independent variables are essentially uncorrelated with each other. This gives the analyst some guidelines for selecting the variables and the order in which they should enter the model. However, sensible judgment is still required in the initial selection of variables and also in the critical examination of the model through analysis of residuals.

Virtual Measurement

Neural networks, unlike conventional programs, are general-purpose systems that attempt to achieve good performance by dense interconnection of simple computational elements. For this reason, they are also called connectionist models. The solution to a problem is not explicitly encoded in the program, but is "learned" by supplying examples of previously solved problems to the network. After the network has learned how to solve the example problems, it is said to be "trained." This is called "supervised training." New data from the same knowledge domain can then be entered to the trained neural network that then outputs a solution. There are also neural networks that learn unsupervised, like Kohonen's self-organizing map network. There are also methods that allow scientists to extract fuzzy rules from a developed neural model. These rules can relate inputs (log responses) to output (permeability) by use of a series of fuzzy rules by dividing the domain of each variable into fuzzy subsets. Implementation of this technique in reservoir characterization is currently under investigation.

Virtual measurement technique was applied successfully in determination of permeability from well log data.[17-20] The major advantage of neural network solution is that it does not require all the pa-

Original SPE manuscript received for review 18 September 1995. Revised manuscript received 3 April 1997. Paper peer approved 10 July 1997. Paper (SPE 30978) first presented at the 1995 SPE Eastern Regional Conference and Exhibition held in Morgantown, West Virginia, 17–21 September.

TABLE 1—MOST USED EMPIRICAL MODELS		
Investigator	Year	Equation
Tixier[2]	1949	$\left(\frac{K}{20}\right)^{1/2} = \frac{2.3}{R_o(d_w - d_o)}\frac{\Delta R}{\Delta D}$
Timur[5]	1968	$K = 0.136\frac{\phi^{4.4}}{S_{wi}^2}$
Coates & Dumanoir[6]	1974	$K^{1/2} = \frac{C}{w^4}\frac{\phi^{2w}}{R_w/R_{ti}}$
Coates[7]	1981	$K^{1/2} = 100\frac{\phi^2\left(1 - S_{wirr}\right)}{S_{wirr}}$

rameters and the relationship between them to be specified explicitly. Because the neural networks learn to solve problems through examples, they are especially suited for subjective and interpretative processes that humans easily can perform intuitively, but which we cannot describe in terms of an algorithm or set of equations.

Ground Rules for This Study

The purpose of this study is to test the capability of each method to match the presented data (model development) (Part A) and test the applicability of the most promising methods for permeability prediction from well logs (Part B). We designated a heterogeneous formation in West Virginia for this study. Granny Creek field produces from Big Injun sand, which is a highly heterogeneous formation. Located approximately 25 miles northeast of Charleston, West Virginia, Granny Creek field is structurally situated on the northwest flank of a syncline that strikes north 15 to 20° east to south 15 to 10° degrees west. Upper Pocono Big Injun sandstone is the oil-producing formation in the Granny Creek field. Development of this field started in 1916 and continued for 30 years. Production throughout the field has been continued until the present day. The crude produced in this field is a paraffin-based Pennsylvania grade oil. It has been estimated that this field has a total production of 6.5 to 6.75 million barrels of oil. A moderately successful water flooding operation was initiated during the 1970's and early 1980's. A tertiary recovery CO_2 pilot project was conducted beginning in 1976. The Pocono Big Injun sandstone is a well documented heterogeneous formation.[21-22]

We chose eight wells (1107, 1108, 1109, 1110, 1126, 1128, 1130, and 1134) that had both geophysical log data and core analysis for this study. **Fig. 1** shows relative location of these wells. Only gamma ray, deep induction, and density logs were available for all the eight cored wells. We only present the data and the results for one well, 1107. Data and results for the other seven wells are quite comparable to Well 1107 and are available upon request. All the well logs are compatible in terms of depth and resolution and are corrected for different effects. We feel that the test of Part B of the study is the more important of the two tests because it reflects actual and practical field use of each method. We performed Part A to determine the best candidates for Part B. The procedure of Part B is as follows:

1. Seven of the eight wells are chosen to develop the models.

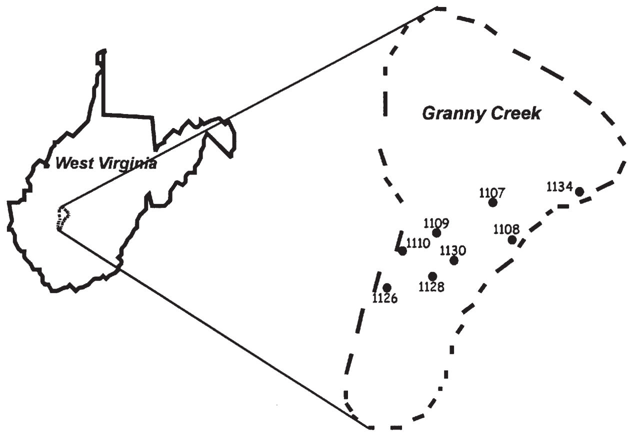

Fig. 1—Granny Creek field in West Virginia.

2. The developed models will be applied to the eighth well. Using the eighth well's log data, a permeability profile for the well will be predicted.
3. The predicted permeability profile will be compared with actual laboratory measurements of the permeability for this well. The technique that performs better under these circumstances should be the superior method.
4. Steps 1 through 3 will be repeated by substituting the eighth well (verification well) with one of the seven (model development) wells. This is to ensure the robustness of the methods.

Results and Discussions

A. Model Development. During the model-development portion of this study, we used core and corresponding well log data from the eight wells. Four empirical methods (Table 1) as well as multiple variable regression (MVR) and virtual measurement [artificial neural networks (ANN)]) techniques were applied to the log data to match the available core permeability. This was performed on all the eight wells and similar results were obtained. **Table 2** shows the typical results that we achieved from this exercise. Complete data and detail analysis for this part of the study are available[23] and can be provided upon request.

From Table 2 (and rest of the analysis), it was clear that MVR and virtual measurement techniques are the most promising methods to be used as permeability prediction techniques. Therefore, during the second part of this study (developing predictive models), we investigated only these two techniques.

B. Permeability Prediction. In this portion of the study, our goal was to develop a robust model that could predict the permeability with only well log data for wells from which core data is not available. Fig. 1 shows the relative location of the eight wells that we used. The approximate distance between Wells 1110 and 1134 is about 2 miles. In the first trial, we used all wells except 1110 to develop the multiple regression and virtual measurement models. Variables used for this development were gamma ray, bulk density, and deep induction log responses. Once the models were developed,

TABLE 2—TYPICAL RESULTS (WELL 1107) FOR MODEL DEVELOPMENT							
	Permeability						
Statistics	Core	MVR	ANN	Tixier	Timur	C&D	Coates
Arithmetic Mean	5.75	4.44	6.39	16.17	31.35	9.44	36.07
Geometric Mean	1.77	1.79	2.76	1.75	6.37	1.06	6.07
Sample Variance	38.6	30.9	47.4	439	1338	187	1874
Standard Deviation	6.21	5.57	6.89	20.9	36.6	13.67	43.30
Standard Error	1.12	1.00	1.24	3.76	6.57	2.46	7.78
Median	2.90	2.37	4.20	3.79	11.90	3.06	15.36
Correlation Coefficient	1.00	0.61	0.91	0.43	0.51	0.61	0.59

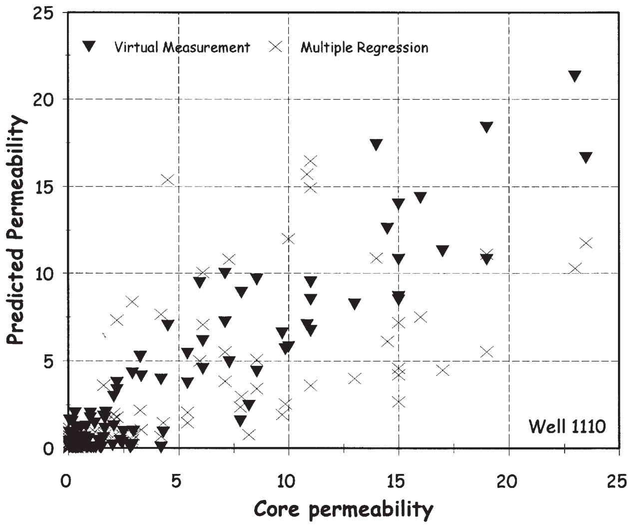

Fig. 2—Crossplot of both models performance against core measurements for Well 1110.

we applied them to Well 1110. The multiple regression model had the following form:

$$k = 126.4669 + 0.0011 * \gamma - 50.2924 * \rho_D + 0.0625 * I_D. \quad (1)$$

The neural model developed by virtual measurement technique cannot be represented with mathematical equations. We developed the virtual measurement model with a back propagation neural network with 18 hidden neurons in the midlayer and logistic activation function in all hidden and output neurons. **Fig. 2** is a crossplot of both models' performance against core measurements. Data in this figure are those used to develop the model (data used for training). From Fig. 2, it is clear that multiple regression tends to overestimate the low-permeability value while underestimating the higher-permeability value. The tendency to average the entire data set to achieve reasonable values for statistical indicators is usually one of the weak points of all regression methods. This behavior has been one of the more consistent characteristics of this method. On the other hand, virtual measurement seems to be less vulnerable to these types of shortcomings. Virtual measurements' relative inaccuracies seem to be consistent throughout the entire domain, which will give scientists a better handle on the problem. It is also noticeable that multiple regression models predict negative permeability values in many occasions, while the virtual measurement model never makes such a mistake. Because of the complex structure that enables a virtual measurement model to learn from experience, it has learned that permeability never takes a negative value. Adaptation of neural networks to the knowledge that has been presented to them in the form of input-output pairs is one of their strong points. This characteristic sets virtual measurement techniques apart from stiff and rigid statistical approaches.

We apply the developed model to Well 1110, keeping in mind that data from this well were not used during the model development. **Fig. 3** shows the prediction of both models with continuous lines, while core measurements are shown with circles. Again, multiple regression clearly shows a tendency to average out the entire data set, while virtual measurement shows better consistency in following the actual trend in permeability variation. In low-permeability range (from 1,913 to 1,938 ft and from 1,870 to 1,876 ft), virtual measurement technique performs similar to multiple regression. High-permeability values occur in bottom of the formation in a thin section at 1,940 ft and then also between 1,877 and 1,908 ft. In both cases, virtual measurement's predictions are closer to core measurements than multiple regression. It is interesting to note the sharp changes in permeability at 1,940 and 1,906 ft and how closely virtual measurement's prediction detects the trend and follows it, while the tendency of averaging is clear in the multiple regression method.

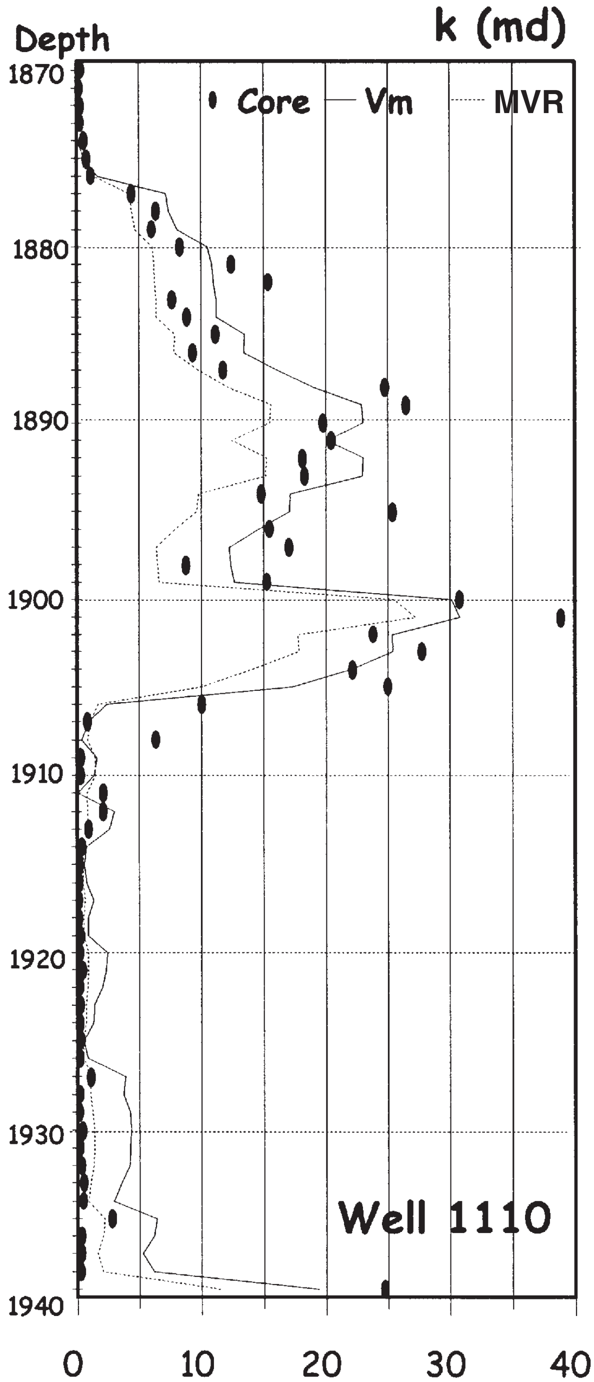

Fig. 3—Prediction models permeability vs. core measurements for Well 1110.

To ensure that it is not an isolated incident when virtual measurement method outperforms multiple regression, we repeated the previous exercise. This time, data from Well 1110 are put back into the data set that is used to develop the model and data from Well 1126 is removed from that data set and put aside for testing the models.

Fig. 4 shows the crossplot that shows the behavior of both models with respect to the development data set, and **Fig. 5** is the permeability prediction for Well 1126 using the developed models. Almost all the previous discussion holds true in this case. Again, during the high-permeability intervals (from 1786 to 1806 ft and from 1828 to 1832 ft), virtual measurement technique outperforms the multiple regression method. Virtual measurement consistently performs better than any other technique that is currently available for predicting permeability from well log responses.

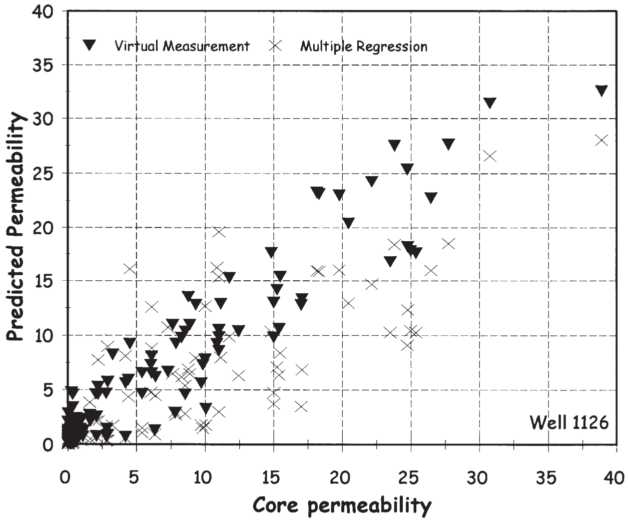

Fig. 4—Crossplot of both models performance against core measurements for Well 1126.

Table 3 provides some detail statistics on the predictive models generated (training) with virtual measurements as well as multiple variable regression techniques. This table also shows the same statistics on use of the developed models to perform actual prediction (testing) on data that were not used during the model development. The main reason for virtual measurement's superiority is the use of artificial neural networks. Neural networks, because of their ability to process data in a parallel and distributed fashion, can discover highly complex relationships between input and output. Neural network's superior ability in pattern recognition is a known fact, and many disciplines in science and engineering take advantage of these abilities.

Conclusions

Verifiable and accurate permeability prediction from well logs in a well with no core measurement data is the bottom line for any technique that claims permeability-prediction capabilities. Many methods use certain core data, such as effective porosity and water saturations, to predict permeability (empirical models). Other methods solely use log data for this purpose, but do not perform adequately once new data are used (multiple variable regression). Virtual measurement technique uses neural networks to predict permeability from well log responses. As we show, virtual measurement can predict permeability values for entire wells without prior exposure to their log or core data and with accuracies that are unmatched by any other technique. The ability of neural networks to learn from experience and then generalize this learning to solve new problems sets it apart from all conventional methods.

We showed that virtual measurement performs better than multiple regression method in predicting permeability from well logs in new wells. We also showed that this characteristic of virtual measurement technique is not accidental and works for any combination of wells in model development and testing.

Nomenclature

γ = gamma ray
ρ_D = bulk density
I_D = deep induction

References

1. Archie, G.E.:"The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics," *Trans.*, AIME (1942) **146,** No. 1, 54.
2. Tixier, M.P.: "Evaluation of Permeability From Electric-Log Resistivity Gradients," *Oil & Gas J.* (June 1949) 113.
3. Wyllie, M.R.J. and Rose, W.D.: "Some Theoretical Considerations Related to the Quantitative Evaluation of the Physical Characteristics of Reservoir Rock From Electric Log Data," *Trans.*, AIME (1950) **189,** 105.

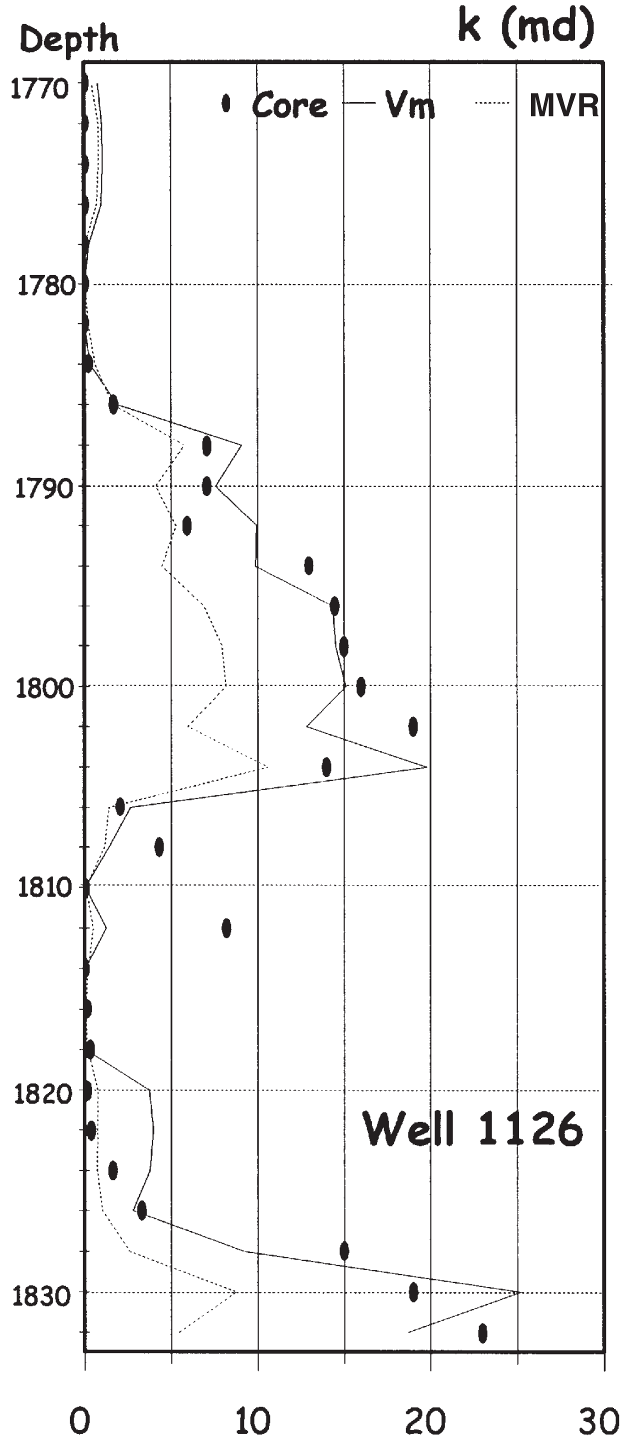

Fig. 5—Prediction models permeability vs. core measurements for Well 1126.

4. Pirson, S.J.: "Handbook of Well Log Analysis," Prentice-Hall Inc., Englewood Cliffs (1963).
5. Timur, A.: "An Investigation of Permeability, Porosity, and Residual Water Saturation Relationship for Sandstone Reservoirs," *The Log Analyst* (July–August 1968) **9,** No. 4, 8.
6. Coats, G.R. and Dumanoir, J.L.: "A New Approach to Improved Log-Derived Permeability," *The Log Analyst* (January–February 1974) 17.
7. Schlumberger Ltd.: "Log Interpretation Charts," Houston, Texas (1987).
8. Kapadia, S.P. and Menzie, U.: "Determination of Permeability Variation Factor V From Log Analysis," paper SPE 14402 presented at the 1985 SPE Annual Technical Conference and Exhibition, Las Vegas, Nevada, 22-25 September.

TABLE 3—STATISTICS ON RESULTS OBTAINED BY VIRTUAL MEASUREMENT AND MULTIPLE VARIABLE REGRESSION METHODS

Well	Data Set	Method	Mean	Std. Dev.	Std. Error	R Squared	Corel. Coef.
1110		Core	3.33	5.21			
1110	training	VM*	2.60	4.33	1.88	0.87	0.93
1110	training	MVR**	2.19	3.59	3.68	0.51	0.71
1110		Core	7.94	9.86			
1110	testing	VM*	8.48	8.56	3.25	0.89	0.94
1110	testing	MVR**	5.21	6.21	3.22	0.89	0.94
1126		Core	4.65	7.43			
1126	training	VM*	4.89	6.88	2.25	0.91	0.95
1126	training	MVR**	3.29	5.28	3.88	0.73	0.85
1126		Core	5.99	7.13			
1126	testing	VM*	6.02	6.90	3.05	0.83	0.90
1126	testing	MVR**	2.76	3.09	3.96	0.71	0.87

*Virtual Measurements using Artificial Neural networks
**Multiple Variable Regression

9. Bloch, S.: "Empirical Prediction of Porosity and Permeability in Sandstones," *AAPG Bulletin* (July 1991) **75,** No. 7, 1145.
10. Ahmed, U., Crary, S.F., and Coates, G.R.: "Permeability Estimation: The Various Sources and Their Interrelationships," *JPT* (May 1991) 578.
11. Yao, C.Y. and Holditch, S.A.: "Estimating Permeability Profiles Using Core and Log Data," paper SPE 26921 presented at the 1993 SPE Eastern Regional Meeting, Pittsburgh, Pennsylvania, 2-4 November .
12. Wendt, W.A., Sakurai, S., and Nelson, P.H.: "Permeability Prediction From Well Logs Using Multiple Regression," Lake, L.W. and Caroll, H.B., Jr. (eds.), *Reservoir Characterization*, Academic Press, New York City (1986).
13. Draper, N.R. and Smith, H.: "Applied Regression Analysis," Wiley & Sons, New York City (1981).
14. Dubrule, O. and Haldorsen, H.H.: "Geostatistics for Permeability Estimation," Lake, L.W. and Caroll, H.B., Jr. (eds.), *Reservoir Characterization*, Academic Press, New York City (1986).
15. Doveton, J.H.: "Geological Log Analysis Using Computer Methods," *AAPG Computer Applications in Geology* (1994) No. 2, 169.
16. McCormak, M.P.: "Neural Networks in the Petroleum Industry," RE2.1, expanded *Abstracts* with biographies, technical program, Soc. of Exploration Geophysicists (1991) **1,** 285–89.
17. Wiener, J.M., *et al.*: "Predicting Carbonate Permeabilities from Wireline Logs Using a Back-Propagation Neural Network," CM2.1, expanded *Abstracts* with biographies, technical program, Soc. of Exploration Geophysicists (1991) **1.**
18. Osborne, D.A.: "Permeability Estimation Using a Neural Network: A Case Study from The Roberts Unit, Wasson Field, Yoakum County, Texas," *Trans.*, AAPG South West Section (1992) 125.
19. Mohaghegh, S. *et al.*: "Design and Development of An Artificial Neural Network for Estimation of Formation Permeability," paper SPE 28237 presented at the 1994 SPE Petroleum Computer Conference, Dallas, 31 July–3 August.
20. Mohaghegh, S. *et al.*: "A Methodological Approach for Reservoir Heterogeneity Characterization Using Artificial Neural Networks," paper SPE 28394 presented at the 1994 SPE Annual Technical Conference and Exhibition, New Orleans, 25–28 September.
21. Donaldson, A. *et al.*: *The Fluvial-Deltaic Big Injun Sandstone in West Virginia*, final reports, DOE/BC/14657-15, U.S. DOE, Bartlesville Project Office (1992).
22. Ameri, S. *et al.*: "Permeability Evaluation in Heterogeneous Formations Using Geophysical Well Logs and Geological Interpretations," paper SPE 26060 presented at the 1993 SPE Western Regional Meeting, Anchorage, 26–28 May.
23. Balan, B., Mohaghegh, S., and Ameri, S.: "State-of-the-Art in Permeability Determination From Well Log Data, Part 1—A Comparative Study, Model Development," paper SPE 30978 presented at the 1995 SPE Eastern Regional Meeting, Morgantown, West Virginia, 17–21 September.

SI Metric Conversion Factors

$$\text{ft} \times 3.048^{*} \qquad \text{E}-01 = \text{m}$$

*Conversion factor is exact.

SPEFE

Shahab Mohaghegh is an associate professor of petroleum and natural gas engineering at West Virginia U., Morgantown. His research interests include application of virtual intelligence in formation evaluation, reservoir, and production engineering. He holds BS and MS degrees in natural gas engineering from Texas A&I U. and a PhD degree in petroleum and natural gas engineering from Pennsylvania State U. Mohaghegh is a member of the *SPE Formation Evaluation* Editorial Review Committee. **Bogdan Balan** is a data modeling engineer at Schlumberger's Austin Product Center in Austin, Texas. Previously, he worked as a reservoir engineer for the Romanian Oil Co. (Petrom R.A.). He holds a BS degree from the Romanian Oil & Natural Gas Inst. and an MS degree from West Virginia U., both in petroleum engineering. **Samuel Ameri** is chairman and professor in the Dept. of Petroleum & Natural Gas Engineering at West Virginia U. Previously, he worked at Schlumberger Well Services. Ameri is West Virginia U. SPE Student Chapter faculty sponsor. He received the 1989 SPE Regional Service Award. He served on the 1986–89 Continuing Education Committee and was chairman of the program committee for the 1989 Eastern Regional Meeting.

Mohaghegh

Balan

Ameri