A Statistical Analysis of the Accuracy and Reproducibility of Standard Core Analysis

David C. Thomas: Amoco Production Company, Tulsa, Research Center
Virgil J. Pugh: Amoco Production Company, Tulsa Research Center

Abstract: Porosity, grain density, and air permeability are routinely measured on most cores taken by Amoco Production Company. These basic data are used to calibrate logs, estimate reserves, and to evaluate reservoir quality and description. Because these data are used in all phases of reservoir development and management, assurance that the data are accurate and knowledge of the limitations of the methods are extremely important. Amoco has prepared core sample sets that have been tested repetitively in house and by commercial laboratories around the world.

This paper discusses the data accumulated and shows the expected accuracy for measurements in a single laboratory by multiple operators over a long period of time. The results are compared to data on the same samples measured by other laboratories. Acceptable ranges of variation within a laboratory and between laboratories are defined and compared to those limits developed by experience. The type of equipment and the procedures used to acquire the data are included in the paper.

INTRODUCTION

Amoco's Rock Property Group has evaluated commercial core analysis laboratories for their accuracy and reliability for many years. Sets of field samples, including a variety of rock types, have been prepared for quality assurance tests of commercial laboratories. This paper describes the quality assurance program in place at Amoco and discusses the statistical analysis of our accumulated data.

Porosity, air permeability, and grain density are measured on most cores taken by Amoco Production Company. These basic data are used to calibrate logs, identify potentially productive intervals, and to select samples for additional study. Usually, the analyses are done on a foot-by-foot basis for the entire core so the volume of data is large. It has proven impractical for Amoco to have an internal laboratory group to make these measurements; consequently, commercial laboratories analyze most of our cores with Amoco's laboratory efforts concentrating on quality assurance evaluations of those laboratories.

Quality assurance and quality control are frequently confused with one another or used interchangeably. Quality control is the internal evaluation that a laboratory makes of any test to ensure that their data are correct. In core analysis, this might be measuring the permeability, porosity, and grain density of calibrated rock samples and checking the bulk volume of known samples or reference billets at regular intervals. Other quality control techniques include a series of samples run repetitively to determine reproducibility of the method within the laboratory and between operators in that laboratory. The analyst should keep records on the results and calibrations that are made and carefully document any problems that are found. Quality control procedures should be ingrained into the daily operations of a laboratory and are the responsibility of the laboratory supervisor and analysts. Most laboratories that Amoco has evaluated have some internal quality control programs.

Quality assurance is the critical examination of the basic assumptions made in the methods and continual review of procedures used by all the laboratories within the company. Quality assurance includes critical evaluation of the accumulated quality control data to verify the accuracy of the methods and pertinence of the procedures. Quality assurance requires a broader view of data quality and imposes a self-critique of one's efforts.

Amoco's role has been to provide quality assurance on core analysis services purchased from commercial laboratories. The program aids the commercial laboratories because they are shown how they have performed relative to Amoco's analyses and are given suggestions on how to correct problems that have been detected.

EXPERIMENTAL METHODS

Porosity and Grain Density: Core plug porosities are determined by direct measurement of grain volume and bulk volume. Grain volume (GV) is measured using the double-cell Boyle's law method as described in API RP-40, Section 3.5.11, Modification A (1). The equipment shown schematically in Figure 1 has been modified to use a precision pressure transducer (Setra Model 204) and a thermal ballast chamber. This 3 liter stainless steel pressure vessel, partially filled with clean copper beads, holds helium at the working pressure (about 100 psig) to minimize adiabatic cooling of the system. Rapid repetitive operation cooled the apparatus enough to change the calibration between tests, introducing a small systematic bias.

Bulk volume (BV) is measured by the Archimedes mercury immersion method with the apparatus shown in Figure 2. A cup of mercury is placed on a single-pan electronic balance and a pronged fork is immersed to a reference mark. The balance is tared and the plug to be measured is forced into the mercury with the fork to the

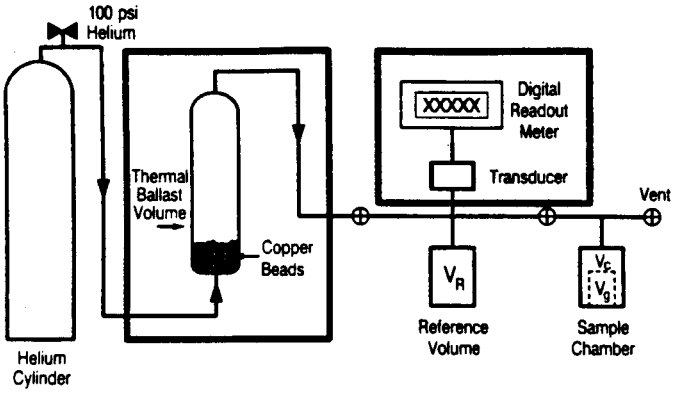

Figure 1: Double-cell Boyle's Law porosimeter: V_c = volume of chamber, V_g = grain volume of sample, and V_R = reference volume.

reference mark (3–7 mm), depending on the plug size. The technologist returns the fork to the same position by eye. Because the forks are made of 18 gauge stainless steel wire, the errors induced in positioning the fork are insignificant (less than 0.001 cc). More complicated electronic methods have been used, but were discarded because of the nuisance involved in keeping them running. A heavy weighted base is necessary to hold the sample firmly in place to ensure accurate weights. This method cannot be used with unconsolidated samples in lead sleeves or samples that have very large surface vugs. Such samples must be measured by direct measurement of the pore volume in a Hassler core holder or, as last resort, by caliper measurement of bulk volume. Such results are less accurate and must be noted in any report. The samples chosen for our quality assurance program have been selected to allow use of the methods described in this paper. The quality assurance of unconsolidated or vuggy core analysis is beyond the scope of this paper.

The bulk volume is calculated using the following equation:

$$BV = \frac{\text{Mass of mercury displaced}}{\text{Density of mercury at measurement temperature}}.$$

It is important to use the correct mercury density for the laboratory temperature. A temperature variation of 5°C will induce a systematic error of 0.02% in the bulk volume.

Porosity and grain density are calculated using the following equations:

$$\text{Porosity (\%)} = [(BV - GV)/BV] \cdot 100$$

and

$$\text{Grain density (g/cc)} = \text{Plug dry weight (g)}/GV \text{ (cc)}.$$

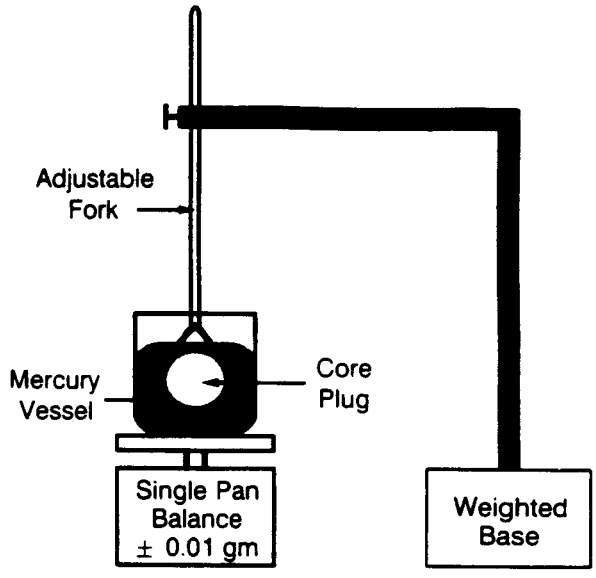

Figure 2: Plug bulk volume measurement apparatus.

Many laboratories use a calibrated screw pump (the mercury pump) to measure bulk volume. This technique is very good if care is taken to minimize the immersion depth of the sample in mercury. The most common implementation of the method immerses the sample about 2 inches. This mercury height induces a capillary pressure of about 1 psi to the top of the plug and makes the bulk volumes measured with the system systematically low because of mercury conforming to microscopic roughness and entering large pores.

The error induced by excessive immersion was tested using 1 inch and ¾ inch diameter plugs cut from coarse and fine-grained sandstones. The results (Figure 3) show that size is more important than texture, because of the changing surface to volume ratio. Steel billet measurements are shown to verify that the effects seen were caused

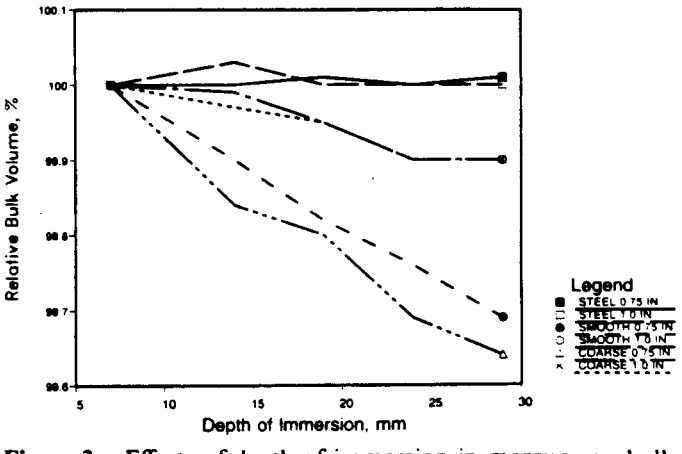

Figure 3: Effects of depth of immersion in mercury on bulk volume of ¾ and 1 inch diameter samples.

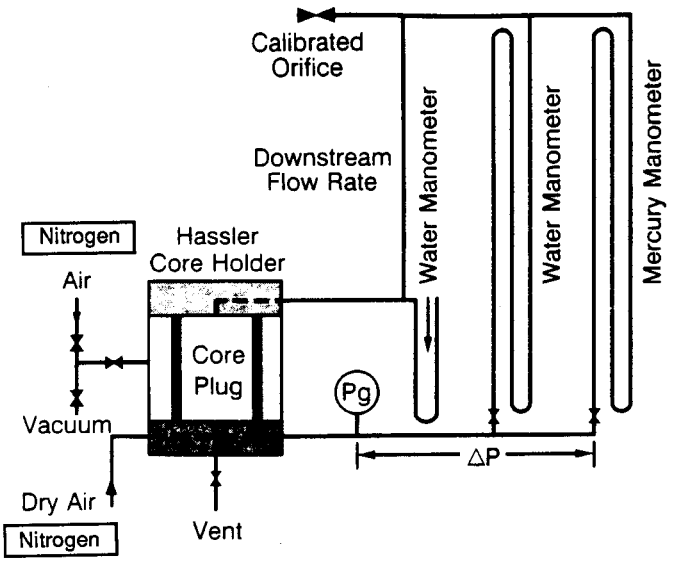

Figure 4: Manual core plug permeameter.

by immersion depth rather than some experimental artifact. Immersion of the samples 21 mm farther than normal induced an error of about 0.1% in bulk volume in 1 inch plugs. The smaller plugs showed a larger relative error in bulk volume because of their increased surface area to volume ratio. This error is systematic and can be seen in the porosity as a negative bias of about 0.2 porosity % over a wide porosity range. Amoco has seen systematic errors of −0.4 porosity % that are directly attributable to depth of immersion errors.

Bulk volume measurement errors are the most common reason for a laboratory's failure of Amoco's quality assurance tests. A bulk volume method in which no capillary pressure effects occur would be ideal; however, such a system is practically impossible. Immersion of samples in mercury to minimal depths has proven to be the simplest, most accurate, and most reproducible method available.

Air Permeability: Core plug permeability is measured using the procedures described in API RP-40, Sections 3.5.15.1 and 3.5.15.2 (1). No modifications have been made to the apparatus, shown schematically in Figure 4, except the use of dry nitrogen for both the annulus pressure and for the flowing gas. This permeameter has remained unchanged since the early 1950's. Whereas the accuracy of the results could be improved with electronic transducers, the simple manometer system is still used in a large number of laboratories worldwide, and therefore is tested in Amoco's quality assurance program. The accuracies produced by this system have proven adequate for the usual applications of air permeability. Modernized manual permeameters and newly developed automated permeameters are being tested in Amoco's laboratories. Those evaluations are beyond the scope of this paper.

The results are calculated using Darcy's law for compressible fluids:

$$K_g = \frac{2{,}000 P_0 Q_0 \mu_g L}{(P_1^2 - P_0^2)A},$$

where

K_g = gas permeability (millidarcies),
P_0 = barametric pressure (atm),
P_1 = upstream pressure (atm),
Q_0 = gas flow rate (cc/sec.),
μ_g = gas viscosity (cp),
L = core length (cm), and
A = core area (cm^2).

The most common problems found in air permeability measurement are system leaks that affect either the differential pressure measurement or the gas flow rate. Improper calibration or damage to the orifice used for gas flow rate calculation is the next in frequency. Care must be taken in handling the orifice to avoid mechanical damage. The orifices should be stored in a desiccator to minimize contact with potentially damaging atmosphere.

The air permeability being measured is run at low confining pressure to estimate an ambient, unstressed value. The pressure applied by the core holder is kept at low levels (less than 500 psig) to avoid rock compressibility effects. This low confining pressure can induce errors in the measured permeabilities if the pressure used is so low as to allow bypassing of gas around the outside of the rock.

DISCUSSION

Since 1978, Amoco Research has performed 126 quality assurance studies on 65 laboratories representing 35 service companies. The fraction of laboratories meeting Amoco's quality assurance criteria increased from about 50% in 1978 to about 65% in 1987. The studies use either a selection of plugs from a recent job performed for Amoco by the laboratory or quality assurance test plugs prepared by Amoco Research.

Specific guidelines are used in comparing commercial laboratory data to Amoco's values to evaluate data quality. The criteria stated in Table 1 were developed from more than 30 years of experience in core analysis. They represent deviations large enough that random variation in test results cannot explain a laboratory exceeding them. The differences must result from calculation errors, procedures, operator efficiency, or equipment. While these criteria are good estimates of the errors associated with the methods, a statistical test was needed to ensure that the criteria are fair and represent field practices. Amoco has accumulated enough information about known samples to produce statistically-based criteria for core plug

Table 1: Maximum acceptable deviation in standard core plug analysis.

Measurement	Experience-based	Statistically derived[1]	Mean coefficient of variance, %
Porosity	±0.5 Por %	±0.23%	0.67
Grain density	±0.01 g/cc	±0.0093 g/cc	0.13
Air permeability:			
0.01–0.1md	±30%	±21%	8.0
0.1–1.0 md	±25%	±21%	8.0
1.0–50 md	±15%	±13%	5.0
50 md–1 darcy	±15%	±8%	3.0

[1] A single measurement made on the same sample that falls outside the specified limit is 99% likely to be in error, assuming that both tests are performed with the same standard deviation.

data that can withstand critical evaluation. A major objective of the work reported here was to verify that the traditional confidence limits are valid statistically.

Statistical Analysis of Core Plug Data: The samples were selected to cover the range of air permeability (0.01–1,000 millidarcies), porosity (3–40%), and grain density (2.6–2.9 g/cc) usually found in oil and gas reservoirs. Friable samples or those with active clays were avoided because the plugs must be able to survive repeated handling, shipment, and oven drying. Approximately 1,900 tests on about 70 plugs by Amoco analysts and 600 analyses of the same plugs have been performed by service company laboratories. The results obtained on tests of these plugs have been collated and summarized in a database.

The data are shown in graphical form in Figures 5–7. These figures show the average of all Amoco measured results compared to the average for all commercial laboratory tests on the same samples. The reference line shown on each graph is a one-to-one correspondence line that should be fit by the data if no overall bias is present. The correspondence between samples is good when viewed on this scale. Certain samples show significant deviation from the main trend line. These deviations reflect averages of relatively few measurements on a sample where one or more measurements were significantly in error. For example, in Figure 6 three points lie significantly above the trend line and form a second line. These data are influenced by one laboratory that had serious procedural errors. A similar problem is represented in Figure 7, where three grain density averages are well off the trend line for the same reasons.

Another estimate of the variability is shown in Figures 8–10. These figures show the scatter of data obtained by service companies compared to the mean value for all Amoco measurements on the same samples. The data

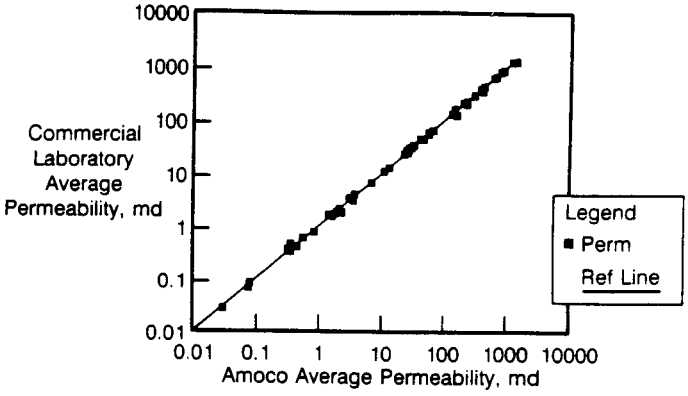

Figure 5: Comparison of average air permeability measurements on all quality assurance samples.

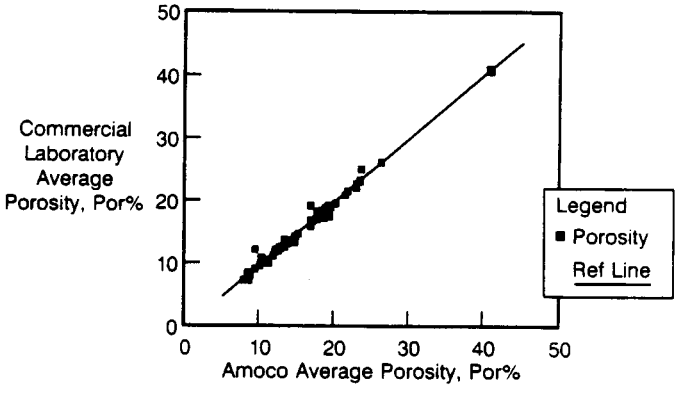

Figure 6: Comparison of average porosity measurements on all quality assurance samples.

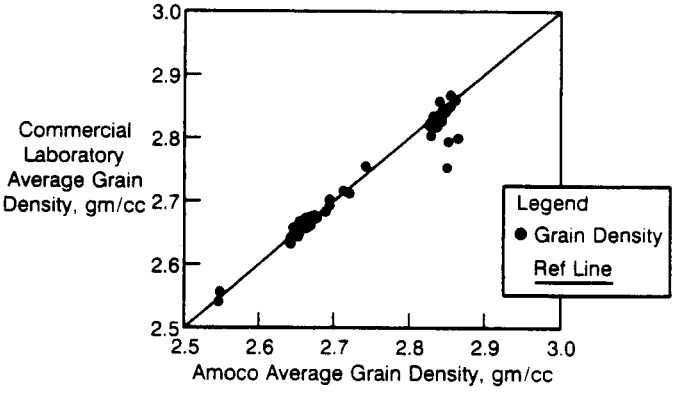

Figure 7: Comparison of average grain density measurements on all quality assurance samples.

have been normalized to allow a meaningful comparison to be made. If a service company result was the same as the average of all Amoco measurements, it was assigned a value of 1.00. All measurements made by service companies are represented on the figures. The comparison shows that the service company results cluster around the mean values in a normal distribution as expected when

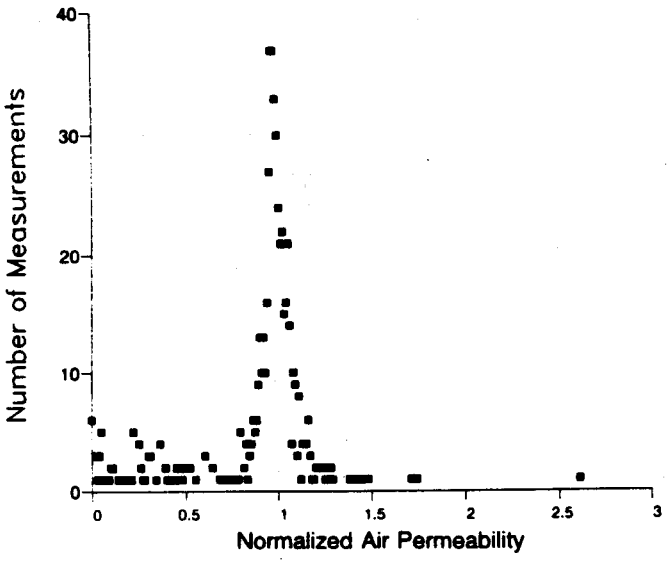

Figure 8: Distribution of air permeability results measured by service companies on Amoco quality assurance samples.

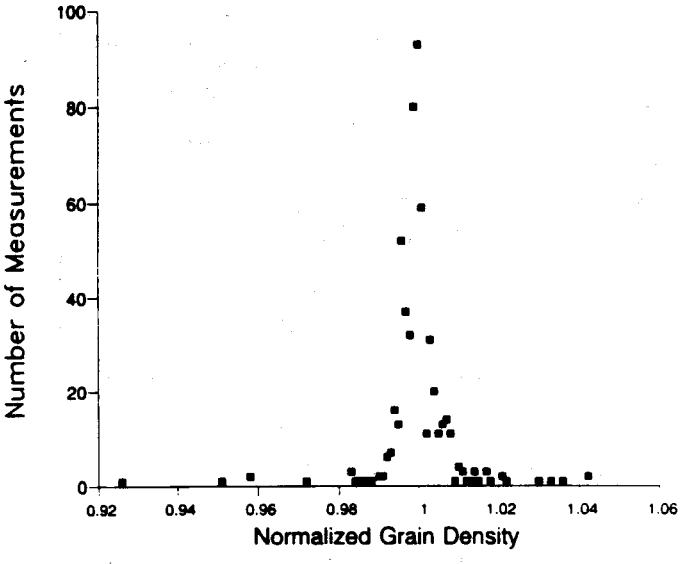

Figure 10: Distribution of grain density results measured by service companies on Amoco quality assurance samples.

the main component of the error is random variation. Systematic error would be represented by cluster of data that does not fit the normal distribution. The permeability data shown in Figure 8 show such a systematic error in the method. A large number of measurements are low relative to the mean. It is speculated that the difference is caused by insufficient stabilization time being allowed in some air permeameters. This result surprised the authors, who expected any bias to be high because of bypassing of gas around the samples.

A more detailed analysis of Amoco data was used to develop the statistically-derived confidence limits shown in Table 1. All the data were analyzed to measure the mean, standard deviation, and coefficient of variance (see the Appendix for definitions). The coefficient of variance is used to provide a basis to compare the results relative to one another over a broad range of values, such as 0.01–1,000 millidarcies in air permeability. The larger the coefficient of variance, the more random error is associated with the measurement. This variation reflects the accuracy with which the experimental measurements can be made. Analysis of the air permeability measurement shows that low permeabilities are less accurately measured than high permeabilities; consequently, the acceptable variation in the measurement is greater.

The variability in grain density and porosity over the range tested is much smaller and shows no particular trend in coefficient of variance. The average coefficient

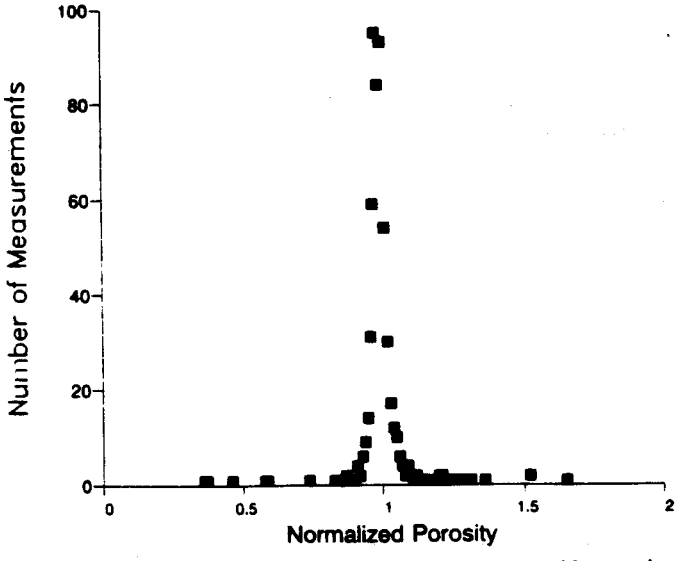

Figure 9: Distribution on porosity results measured by service companies on Amoco quality assurance samples.

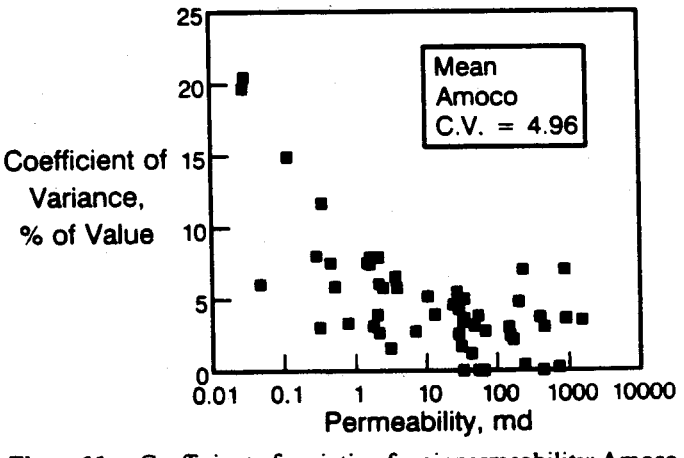

Figure 11: Coefficient of variation for air permeability: Amoco results.

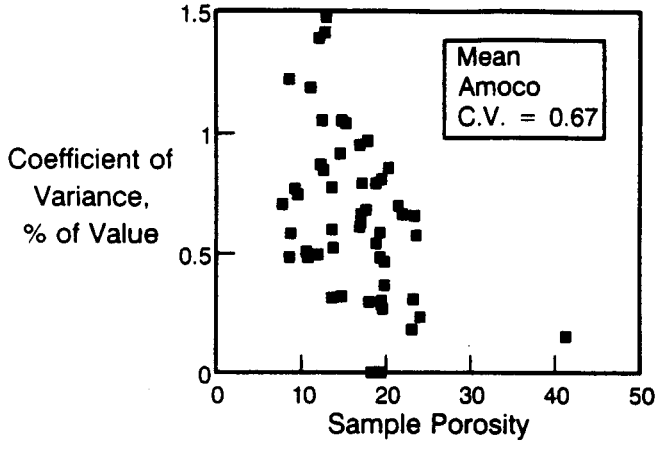

Figure 12: Coefficient of variation for porosity: Amoco results.

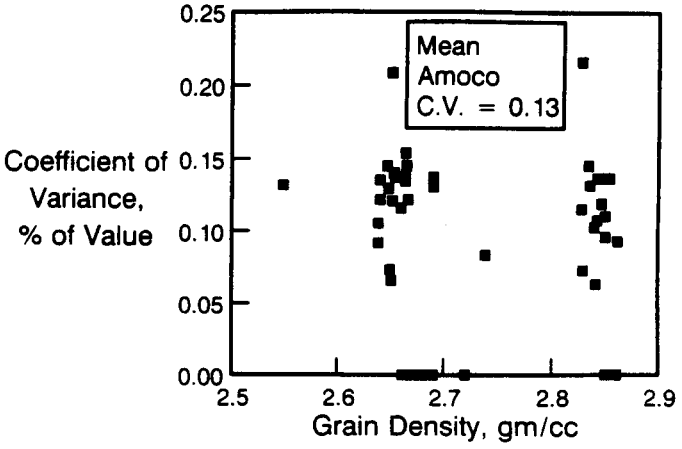

Figure 13: Coefficient of variation for grain density: Amoco results.

of variance gives a good estimate of the expected random error in the measurements.

The samples are assumed to not change between tests and that random variation in the analytical result is the dominant source of variability. Variation around the mean is usually represented by the Gaussian distribution. The area under the curve between $-s$ and $+s$ is 68.26% of the total area. If the only errors in the measurement are random variation around the mean, 68.26% of all tests on that sample will produce a result that falls within $\pm s$. Similarly, 95.44% of all measurements are expected to be between $\pm 2s$. This range is a confidence interval.

The confidence limits shown in Table 1 represent the 99% confidence level that compares a single measurement with the mean. A single measurement made on the same sample that falls outside the specified limit is 99% likely to be in error. The experience-based allowable variation used by Amoco to qualify laboratories is larger than the 99% confidence limit in each case. Those limits are quite conservative and allow for considerable differences in laboratories that will still pass Amoco's criteria.

CONCLUSIONS

1. The experience-based limits on deviation in core analysis have been shown to be valid statistically. Rejection by these criteria are at or above the 99% confidence limits for the expected variation in the methods.
2. Statistical analysis of core data shows that air permeability measurements as practiced in the industry have a systematic negative bias that must be addressed to improve overall core analysis results.

SUMMARY

Amoco's Rock Property Group has evaluated commercial core analysis laboratories for their accuracy and reliability for many years. This paper analyzes the data accumulated and shows that the confidence limits used as judgement criteria are statistically valid. Acceptable ranges of variation within a laboratory and between laboratories are defined. The type of equipment and the procedures used to acquire the data are described. An improved method for bulk volume measurement on plug samples is given. A systematic error in air permeability is demonstrated from the data analysis.

REFERENCES

1. API (1960), "Recommended Practice for Core Analysis Procedure," American Petroleum Institute, RP-40.
2. Blaedel, W. J. and Meloche, V. W. (1963), "The Theory of Error and the Treatment of Quantitative Data" in *Elementary Quantitative Analysis*, 2nd Ed., New York: Harper and Row, pp. 37–58 and 613–644.
3. Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978), *Statistics for Experimenters*, New York: Wiley.
4. Worthham, A. W. and Smith, T. E. (1959), *Practical Statistics in Experimental Design*, Dallas, TX: Dallas Publishing House.

APPENDIX

Statistics Used in Quality Assurance Evaluations

Mean: The mean is the average test result for all tests on the individual samples (2–4). Mathematically, it is

$$\bar{x} = \left(\frac{\sum_{i}^{n} x_i}{n} \right)$$

DEVIATION ($x_i - \bar{x}$): The deviation is the difference between the mean of all measurements and any single measurement on a sample. This statistic is used heavily in our program because most data are provided as single measurements on a sample (2–4).

Sample Standard Deviation, s: The sample standard deviation is a measure of the spread, or deviation, of the sample values around the mean. If the sample being tested does not change,

the standard deviation gives information about the random errors in the measurement method. It is the positive square root of the sample variance and has the same units as the mean (2–4). Mathematically, it is

$$s = \left[\frac{\sum_{1}^{n} (x_i - \bar{x})^2}{n - 1} \right]^{1/2}$$

Coefficient of Variation, CV: The CV is the standard deviation of a sample divided by the mean of the sample expressed as percent (2–4):

$$CV = (s/\bar{x}) \cdot 100.$$

The coefficient of variation normalizes standard deviations so that comparisons between samples can be made. It also allows calculation of a better estimate of the standard deviation to be expected by use of a relatively small number of measurements on a larger number of samples. The CV also shows the interaction of the mean and the standard deviation. Such effects are caused by the sensitivity of the method to measurements made during the test. For example, accuracy of a given test may be limited by the accuracy with which the sample can be weighed or with which the pressure drop can be measured. Such analysis is helpful in error analyses on analytical methods.

Confidence Intervals and Limits: A confidence interval is an estimate of the interval in which the true value lies. The more commonly used term, confidence limit, has essentially the same meaning as confidence interval. A confidence limit is "A range of values within which a sample observation can be expected to fall at a specified probability level." Usually, variation around the mean is discussed as "falling within certain [e.g., 90%] confidence limits" (2).

These values help the experimenter estimate the confidence interval for a given test. The confidence limits (μ) around the mean were calculated, after normalization, using the following equation where $z = 2.58$ (ref. 2):

$$\mu = \bar{x} \pm z\left(\frac{s}{\sqrt{n}}\right).$$