Estimating the mean permeability: how many measurements do you need?

Patrick W.M. Corbett and Jerry L. Jensen

Introduction

The question of the number of permeability measurements that are required to estimate the mean permeability of a reservoir interval is not generally considered during reservoir evaluation. Permeability sampling is traditionally carried out with core plugs (1.0–1.5 inch or 2.5–3.8 cm cylindrical samples) at a spacing of about 1 foot or 0.3 m. This spacing has evolved from considerations of practicality, cost and the need for core conservation, and takes little account of the underlying rock variability. The development of low-cost, automated probe permeameters (e.g. Halvorsen and Hurst 1990) removes these constraints and permits a reconsideration of the question of sample spacing.

During reservoir characterization, two parameters defined by the distribution of permeability measurements in the unit of interest (lamina, bed or formation) are particularly important:

• the value in the centre of the range around which observations tend to cluster, which is estimated by the average, harmonic average or median, and
• the spread about the centre, as given by the standard deviation, for example.

The appropriate choices for these two parameters, which can have a major influence on predicted fluid flow characteristics, will depend on the variability of permeability and the intended application. In low-variability sediments the mean permeability can be estimated by the arithmetic average of the sampled values and a few samples will suffice to characterize the rock. Highly variable sediments, however, will have regions of very high and very low permeability. The upper extremes will govern the horizontal permeability of the reservoir unit because the bulk of the horizontal fluid flow takes place along the most permeable layers. The lower extremes will govern the vertical permeability because the rate of vertical fluid flow will predominantly be limited by the least permeable layers (Archer and Wall 1986). Rocks that are highly variable require many samples to ensure that the extremes are identified.

In recent work from probe permeameter studies (Hurst and Rosvoll 1991; Corbett and Jensen 1992a) a method has been suggested for estimating the minimum number of samples required to characterize a rock unit. The method uses the coefficient of variation, Cv (defined below), in a convenient rule of thumb. Strictly speaking, the rule of thumb only applies to sample numbers for the arithmetic average but, as we show later, it is also useful for other statistics. We also show evidence linking Cv and geological variability, so that permeability variation, geological variation, and the required number of samples can all be related.

Cv and the prediction of number of samples for the arithmetic average

Cv is defined as

\[ Cv = \frac{\text{Standard deviation}}{\text{Arithmetic average}} \]

The value of Cv can be roughly estimated from experience with previous studies or a geological description. A better estimate can be obtained from a pilot survey of the interval under study. Corbett and Jensen (1992b) outline an approach to obtaining enough data for accurate Cv estimates. The procedure requires an initial estimate of Cv based on about 25 or more measurements. The required number of data, \( N_0 \), derived from a method described by Hurst and Rosvoll (1991), is given by

\[ N_0 = 100 \, \text{Cv}^2 \]  

(1)

to give an estimate of the mean within ±20% of the true value (Corbett and Jensen 1992a). For example, if 30 samples give \( \text{Cv} = 0.5 \), then no more samples need be taken since \( N_0 = 25 \). If those samples give \( \text{Cv} = 0.7 \), however, then \( N_0 = 49 \) and at least another 19 measurements should be made. In practical terms, one would halve the spacing with infill measurements to take 60 samples, bearing in mind that the estimate of Cv is likely to change with more data. It is important that the samples be distributed uniformly (to eliminate bias) over the sample interval or grid. One should also avoid selecting an initial number of samples that results in a

---

2Department of Petroleum Engineering, Heriot-Watt University, Edinburgh EH14 4AS, UK.
3The words 'mean' and 'average' are synonyms in common parlance. However, in statistical analysis of sampled datasets, the term 'mean' is reserved for the true (mean or average) value for the whole population under consideration; and this can only be estimated by calculating the 'arithmetic average' of the measured sample values.
sample spacing equal to any obvious periodicity in the rock unit.

If an error margin other than 20% is desired, (1) can be adjusted accordingly (Hurst and Rosvoll 1991). The number of samples required, however, increases dramatically as the tolerance is reduced. For example, a sample with \( \text{Cv} = 0.9 \) gives \( N_0 = 81 \) samples for a 20% margin. For a 10 or 5% margin, \( N_0 \) increases to 325 or 1300, respectively. At this time, given the level of uncertainty for many other reservoir parameters (e.g. bulk rock volume) and considering the cost associated with sampling, permeability to within \( \pm 20\% \) is considered to be an appropriate tolerance.

**Heterogeneity classes**

In this paper we use \( \text{Cv} \) to define the following heterogeneity classes:

- \( 0.0 < \text{Cv} < 0.5 \) Homogeneous
- \( 0.5 < \text{Cv} < 1.0 \) Heterogeneous
- \( \text{Cv} < 1.0 \) Very Heterogeneous

These classes have been defined on the basis of a review of published variability measures for a range of sedimentary types (Fig. 1). In extracting these data, we have used (1) to eliminate those studies that are based on too few samples.

Figure 1 shows that low \( \text{Cv} \) values are associated with sediments of higher energy, better sorted facies, while high \( \text{Cv} \) values are found where mixed lithologies, facies or pore types are present. Analysis of the relationship between sedimentology and \( \text{Cv} \) variability is the subject of continued research. However, Fig. 1 suggests that (1) and the geological description can give sample sizes for reservoir characterization.

With \( \text{Cv} < 0.5 \), the permeability distributions are normally or approximately normally distributed (Goggin et al. 1988; Corbett and Jensen 1992a), and are considered to be effectively homogeneous (i.e. the level of heterogeneity will not have a significant impact on flow performance). When \( \text{Cv} > 0.5 \), permeability distributions tend to be progressively more skewed and the rock unit more heterogeneous.

**Application of \( N_0 \) in three North Sea formations**

Equation (1) was originally developed (Hurst and Rosvoll 1991) assuming normal distributions and was, therefore, thought to be of limited practical use because permeability in sediments has often been observed to have non-normal (i.e. skewed) distributions (e.g. Goggin et al. 1988). Permeability data can also be statistically correlated (i.e. nearby samples have similar properties), which might also limit the effectiveness of the method. We have found, however, that (1) is robust and useful (Corbett and Jensen 1992a).

Three formations have been used to examine the variation of the \( N_0 \) value. The formations were selected on the basis of their having comprehensive data sets and varying degrees of heterogeneity. In each case, probe permeameter measurements have been used. Intervals from the Rannoch, Etive and Rotliegender formations were chosen with \( \text{Cv} \) values of 0.48 (homogeneous), 0.95 (heterogeneous) and 3.1 (very heterogeneous), respectively. The Rannoch and Etive formations are successive marine, shoreface sequences of the lower Brent Group.

![Fig. 1. Coefficient of variation (Cv) and optimum number of samples (\( N_0 \)) for a range of sedimentary types. (Expanded from Corbett and Jensen 1992b.)](image-url)
(Middle Jurassic) and the Rotliegendes formation is an aeolian sequence (Early Permian). For the investigated shoreface intervals, the Etive is more heterogeneous than the Rannoch Formation. Using the above rule of thumb, the appropriate numbers of samples ($N_0$) are 23, 91 and 935, respectively. This means that the core plugs taken (numbering less than 15 in the Brent intervals and approximately 500 in the Rotliegendes interval) are insufficient in number to estimate the mean permeability of the intervals investigated (4 m thickness for the Brent
By varying the sample spacing (but maintaining a regular spacing throughout the interval of interest) we can see how variable the estimates provided by the arithmetic average are with respect to the rule-of-thumb predictions. We generated subsets by resampling the data at different spacings following the procedure described in Corbett and Jensen (1992a). To compare the results, given the difference in permeability levels of selected formations, we used normalized quantities.

The normalized arithmetic average is the estimated
arithmetic average of the subset sample divided by the arithmetic average of the total sample. The normalized sample number is the sample number, \( N \), divided by \( N_0 \). The results are shown in Fig. 2. Samples with numbers of data points greater than \( N_0 \) (i.e. where the normalized sample number > 1) will clearly estimate the mean within ±20% for all three formations.

The dotted line envelopes in Fig. 2 are defined by the functions \( 1.0 \pm 0.2 \sqrt{\frac{N_0}{N}} \). These envelope functions are based on the assumption that errors are normally distributed, and imply that the arithmetic average variability decreases as \( N^{-1/2} \) for uncorrelated measurements. Since permeability data are often statistically correlated (e.g. Corbett and Jensen 1992a), the fact that the averages stay within the envelopes is an unexpected result.

Similar plots in Fig. 3 for the harmonic average (defined as the inverse of the average of the inverses of the measured permeability values) of the same data, show that more (up to twice) the number of samples may be required than for the arithmetic average, if ±20% tolerance is required. If we only take \( N_0 \) samples, then the harmonic mean is only estimated to within ±40%. The additional samples reflect the sensitivity of the harmonic average to low values. As expected, note that the relative differences between arithmetic (Fig. 2) and harmonic (Fig. 3) averages increase as the Cv increases. In the facies sampled here, the low-permeability zones tend to be thin relative to high permeability zones. As a result, the harmonic average of a sample tends to underestimate the harmonic mean of the population. An additional problem is that low permeability laminae can be very thin and below the resolution of even the probe perimeter (approximately 5 mm for a 3.6 mm injection diameter, Corbett and Jensen, 1992a) when sampled along vertical profiles.

Appropriate sampling from a knowledge of \( \text{Cv} \)

From a knowledge of variability in a wide range of sediments, we can begin to determine \( N_0 \) for a range of reservoir rock types (Fig. 1). Using this plot as a guideline, some general sampling rules can be suggested. Unless the variability of the reservoir is known to be low, we recommend an initial 25 samples (i.e. sufficient to satisfy the assumptions in the derivation of (1) and also sufficient to characterize homogeneous media) be collected at a regular spacing (to avoid bias). \( \text{Cv} \) should then be determined. If the \( \text{Cv} \) value is < 0.5, we can be confident in our estimate of \( \text{Cv} \) and sufficient samples have been collected for estimation of the arithmetic mean to within ±20%.

For heterogeneous rocks with \( \text{Cv} \) up to 0.7, 50 samples would be appropriate. For \( \text{Cv} \) up to 1, 100 samples would suffice. It should be remembered that if \( N < N_0 \), the estimates of \( \text{Cv} \) are also based on insufficient samples and are likely to change with additional sampling.

Anisotropy, as measured by differences between the arithmetic and harmonic averages, tends to be significant for rocks with \( \text{Cv} > 0.5 \). These averages (the latter with careful appreciation of the problem of resolving thin laminae) can be used to estimate horizontal and vertical permeabilities (Archer and Wall 1986, p. 83), provided that thin, unresolved laminae are not a problem for the interval of study. For very heterogeneous materials, adequate estimation of the arithmetic or harmonic means will require many samples. A consideration of available sample interval or area will determine whether core plugs are practical or if a probe perimeter is required.

The accuracy of other characterization measures is also being examined by us using the \( N_0 \) criterion. Preliminary results based on empirical studies (e.g. Corbett and Jensen 1992a) suggest that the harmonic average is one of the most demanding statistics in terms of the number of data required for accurate estimation. Limited computer modelling of measurements which we have done has supported these findings, but further work is needed. For example, it is important to note that the determination of spatial statistics (e.g. correlation lengths or variogram ranges) can be more sample intensive than descriptive statistics. If spatial statistics are required, \( N_0 \) samples may not be appropriate. In sedimentary rocks, nested correlation lengths (e.g. at lamina and bed scales) are common and separate sampling programmes (at a spacing of about one-tenth of the expected correlation length) need to be defined for each scale (Corbett and Jensen 1992a).

Comments and conclusions

An easy to use rule-of-thumb based on a measure of variability, termed \( \text{Cv} \), can be used to determine the appropriate number, \( N_0 \), of samples for estimation of the arithmetic mean permeability of rock samples to a tolerance of ±20%. If the harmonic mean is to be estimated to the same tolerance, this number should be doubled.

An initial 25 samples, evenly spaced over the stratigraphic interval of interest, can be used to estimate the variability. Depending on the \( \text{Cv} \) from this initial survey, this sample spacing can be subdivided if more samples are found to be needed.

The \( N_0 \) test can be used to distinguish reasonable estimates of the arithmetic means from those that are potentially unreliable.

Sampling programmes at arbitrarily fixed spacings (i.e. the 0.3 m spacing of core plugs) run the risk of undersampling heterogeneous materials and oversampling homogeneous materials. This practice leads to biased estimates of ‘average’ properties with unknown variability.
A knowledge of variability in reservoir rock types will ensure the design of an appropriate and cost-effective sample programme. This knowledge will improve as a wide range of media is adequately sampled in cores and at outcrop.

Acknowledgements

The authors would like to thank Christian Halvorsen for data acquisition. Statoil and BP are also thanked for technical and financial support, respectively, and their partners in the Statfjord Field for allowing publication of the Brent data. The ideas in this paper have greatly benefited from discussions with other members of the Edinburgh Reservoir Description Group, particularly Jon Lewis, Kjell Rosvoll and Philip Ringrose.

Received and accepted 20 December 1991.

References