# Petroleum Exploration[1]

There is a unique true distribution of rock properties in each petroleum reservoir. This unique distribution is the result of a complex sequence of physical, chemical, and biological processes. Although some of these depositional and diagenetic processes are well understood, it is impossible to define the initial and boundary conditions in sufficient detail to provide a deterministic picture of the reservoir. There is unavoidable uncertainty in the distribution of rock properties. *Geostatistical techniques* are being increasingly used to quantify this uncertainty and to create numerical models that mimic the physically significant features of rock property variation.

The uncertainty is assessed by creating alternative realizations of the reservoir. The *response variables* of interest, such as peak production rate and recoverable reserves, are computed with each realization. A distribution or histogram or uncertainty can be assembled by pooling together the results from a number of realizations. This assessment of uncertainty leads to more informed reservoir management decisions.

To be reliable, each geostatistical realization must be constrained to all available information about the reservoir. This information comes from a variety of sources, representing different volume scales, and with different precision. The challenge in geostatistics is to simultaneously honor the variety of available data without assuming information that we do not have.

Hand contouring is one method that has been used to model reservoirs. Although this approach allows the main geological trends in rock properties to be accounted for, there are some problems (1) it is not possible to construct a model with the short scale variability or *heterogeneity* encountered in real geological formations, (2) complex data from seismic and historical production can not be accounted for, and (3) there is no assessment of uncertainty in the reservoir response. There are a number of computer aided mapping techniques based on inverse distance, splines, kriging, or local smoothing. These computer-based contouring

[1]Submitted to McGraw-Hill Inc. by Clayton V. Deutsch on September 16, 1996

methods suffer the same problems as hand contouring. They are appropriate for visualizing large scale trends but inappropriate for representing detailed heterogeneities and assessing uncertainty.

**Data Integration.** Information about the reservoir comes from a wide variety of sources, in a wide variety of formats, covering a huge range of scales from the pore to the basin scale. Moreover, many data sources measure rock properties indirectly related to the rock properties of interest. For example, seismic data measures sonic properties that are imprecisely related to the properties being mapped such as fluid saturations and porosity. Accounting for information from different scales with different precision is a fundamental challenge faced by reservoir geologists and engineers, and more generally earth scientists.

A second challenge faced by earth scientists is that some data are global and some are local. A correlation between porosity and permeability of 0.6 or knowledge of coarsening-upward trends in porosity are *global* data. A porosity measurement of 22.5% or a density of 2.14 are *local* data.

Core data and well log-derived properties provide the only *hard* data on the reservoir rock properties (lithofacies, porosity, permeability, and fluid saturations) that must be mapped. Historically, only vertical wells were drilled providing detailed information on the vertical succession of rock properties. It is becoming increasingly common, however, to drill horizontally through the reservoir interval. Horizontal well data provides valuable additional information on lateral changes in rock properties.

In addition to *hard* well data there are many indirect sources of information. These indirect data, such as seismic, are called *soft* data.

One source of soft data is a conceptual geological understanding of the reservoir depositional system. For example, knowing the direction of sediment transport, at the time of deposition, provides information on large scale trends, stacking patterns, and the spatial continuity of the rock properties. Analogue data from outcrops or densely drilled similar fields also provides conceptual geological data.

Seismic provides a wealth of information about a reservoir. Often, structural surfaces such as the top of the reservoir, intermediate stratigraphic surfaces, and important faults can be interpreted from the seismic data. In addition to large scale geometric information, seismic is being increasingly used to provide information on internal reservoir heterogeneities. Seimic attributes, from an inversion or statistical analysis of the seismic traces, are calibrated to the vertically averaged rock type proportions and porosity.

Only recently have geostatisticians been faced with repeated 3-D seismic surveys or 4-D seismic. 4-D seismic provides additional information that further reduces uncertainty and improves the predictions from flow simulation. The information may include (1) fault transmissibility, e.g., the presence of sealing faults or fluid flow across non-sealing faults, (2) presence of high-permeability regions, that is, regions of preferential fluid movement, and (3) more reliable measurements of lithofacies proportions and porosity. The value of 4-D seismic is that it allows us to subtract away much of the geologic uncertainty in the reservoir and overburden, while highlighting temporal changes in reservoir conditions related to production.

Another important source of *soft* data is from well test and historical production data. This data can be interpreted to yield information on reservoir boundaries, connectivity between wells, and volume averaged properties around wells. In general, honoring dynamic pressure or historical production data is quite difficult. In practice, trial-and-error history matching is still the most common approach at the final stage of modeling. The most promising approach to production data integration is to interpret the dynamic data to yield spatial data at some coarse resolution. Detailed geostatistical realizations are then constrained to that intermediate-scale *soft* spatial data.

The challenge addressed by geostatistics is to create detailed 3-D realizations of the reservoir rock properties that are consistent with all of the relevant *hard* and *soft* data.

**Geostatistical Modeling.** The major roadblock in data integration is the difference in the format (not only measurement units) under which each data type is presented. Some information types are interpretive in nature yet could be critical in the understanding of the reservoir, some data are categorical (e.g., facies types), some are continuous (e.g., porosity and permeability), some appear as constraints (e.g., intervals derived from seismic data), and some are prior probability distributions. A standardization of formats and units that does not tamper with the information content is an important first step towards data integration. The approach taken in geostatistics is to code as much data as possible as probabilities.

The probabilistic language and methodology are unique and universal in that they are not linked to any particular field of application or data type. Probability values are unit-free; in their cumulative form they are standardized in the ultimate standard interval $[0, 1]$. The concept of prior probability distributions allows a common coding of diverse data related to the same goal. The concept of Bayesian updating provides a methodology for merging prior distributions into a single posterior probability distribution. All original *hard* and *soft* data and the final posterior probability distribution are coded as cumulative probability values in the interval $[0, 1]$. From the posterior cumulative distribution function probability intervals can be derived, simulated values can be drawn for the unsampled value, or a single "best" estimated value can be retained for any given optimality criterion.

**Staged Modeling.** As illustrated on **Fig. 1(a)**, there is a vast interwell region to be filled in by geostatistical techniques. Most often, the *hard* well data are hundreds to thousands of feet apart. More extensive *soft* data provide imprecise information at a relatively large scale. For illustration, the black intervals at the wells in **Fig. 1(a)** are shale and the white intervals are sandstone. The task is to establish the distribution of shale/sandstone in the gray-shaded interwell region. Seismic or production-related *soft* data may provide information on the proportion of shale/sandstone between the wells; it can not identify the precise location of the shales.

A staged modeling approach will be considered whereby the main structural control (gridding) will be established first, then the sandstone/shale (lithofacies), and finally continuous rock properties like porosity and permeability are assigned within each lithofacies.

Most reservoirs are hosted in sedimentary rock formations. An essential character of these formations is a stratigraphic layering that can often be correlated between wells. These surfaces correspond to some specific time, where the sedimentary process changed, in the depositional history of the reservoir. For example, the sea level could have dropped causing changes in the locations of erosion and deposition. An important first step in petroleum reservoir modeling is to establish this gridding. **Fig. 1(b)** shows five stratigraphic surfaces defining four layers. Each layer has its own gridding style depending on whether or not erosion occurred at the bounding surface. After establishing a geologically-realistic gridding style, the lithofacies, porosity, and permeability are modeled sequentially.

**Stochastic Modeling.** The distribution of rock properties in a reservoir is not random nor is it smoothly varying between the available well data. Geostatistical techniques allow multiple realizations to be created that are constrained to different types of data. One important piece of information is a global measure of spatial correlation. Geostatistical modeling techniques are based on the theory of random variables and the concept of Bayesian updating that provides a methodology for merging prior probability distributions into a single posterior probability distribution. Geostatistics may also be seen as a branch of applied statistics that places emphasis on multiple data types and the spatial context of that data.

**Fig. 1(c)** and **Fig. 1(d)** show two different geostatistical realizations. Both of these realizations honor the well data and have the same appearance, that is, the shales have about the same size size and aspect ratio. The differences between multiple realizations quantifies our uncertainty in the sandstone/shale distribution.

In practice, geostatistical realizations of continuous rock properties such as porosity and permeability are constructed after the lithofacies (shale/sandstone) are modeled. The resulting complete realizations are then used to address reservoir management questions. For example, the hydrocarbon recovery could be established for a particular realization by fluid flow simulation. Repeating the flow simulation with multiple geostatistical realizations allows us to build a histogram of uncertainty. The 5% low and 95% high values on this histogram tell us the 90% probability interval for the recoverable reserves.

*Clayton V. Deutsch*

**Bibliography.** H. H. Haldorsen and E. Damsleth, Stochastic Modeling, Journal of Petroleum Technology, pp 404-412, April 1990; E. Issacs and R. M. Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, 1989; C. V. Deutsch and A. G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, 1992.

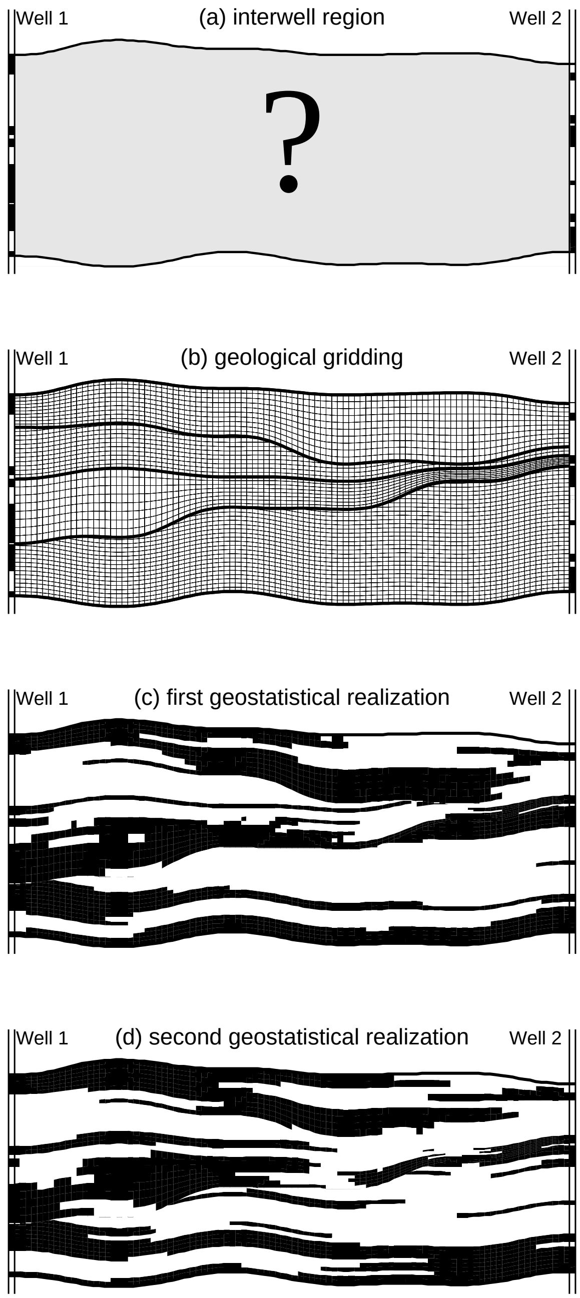


Figure 1: Illustration of geostatistical modeling. The white intervals at the well are sandstone and the black intervals are shale. The problem addressed by geostatistics is to map the vast interwell region as shown in (a). The first step is to establish a geological grid system between interpreted stratigraphic surfaces (b). Then, geostatistical realizations can be generated (c) and (d).