Determining Facies from Wireline Logs using Discriminant Analysis

Michael J. Ranger
Petroleum Consultant

Reservoir simulation requires a matrix of cells to model the geotechnical properties of lithologic intervals (facies) in the subsurface. Typically, a control set of reservoir facies determined from well core are assigned or "pigeonholed" into one of a small superset of lithofacies, representing distinct reservoir engineering properties. These are then used to populate the matrix using some gridding procedure, almost always geostatistical Kriging. Core being a physical sample of the reservoir, geologists can construct a realistic facies model from direct observation. But in areas of poor core control, one may need to interpret facies from wireline logs alone. Some lithologies such as clean sand, coal or 100% clay intervals are easy to differentiate from wireline logs, but mixed and interbedded lithologies are more problematic. The statistical technique known as Discriminant Analysis offers a solution.

Discriminant analysis is a multivariate statistical technique allowing probabilistic discrimination of a data set into discrete populations in multidimensional space. It is rooted in the traditional cross-plot analysis used to discriminate clusters in two dimensions, but extended to n-dimensional space (Figure 1). It is also somewhat analogous to multiple regression, but where the dependent variable is categorical rather than a continuous real number. Applied to wireline logs, each log represents a single linear dimension of multidimensional space, whereas a "facies" corresponds to a multidimensional cluster or ellipsoid in this space. Several examples from Athabasca Oil Sands reservoirs provide an overview of the technique. The analyses are easily carried out on various statistical software packages, including JMP from SAS Institute Inc. (used by the author).

Discriminant analysis utilises a control or "training" set of data. In this application the control set consists of all wells with a common, standardised, set of modern wireline logs and overlapping core recovery. The cores are manually logged so that intervals are assigned to an established reservoir facies scheme. Each facies is considered to be characterised by a set of probability distributions of the multivariate parameters (i.e. wireline log readings). From the control set of data a linear "Discriminant Function" is calculated, which can be used to identify from which population (i.e. facies) non-cored intervals most probably originate (Gupta and Johnson, 2002). A quality check is easily run by applying the Discriminant Function to the wireline parameters from the cored wells, and comparing the predicted facies to the actual core facies (Figure 2). Most statistical software also determines a goodness-of-fit probability score for each facies assignment.

Data preparation is a vital step before the Discriminant Analysis procedure can be implemented. Cores from the control wells must be carefully depth corrected to their corresponding wireline logs. Any uncertainty in this step for a particular well may be sufficient to exclude that well from the control group. Furthermore, one must be satisfied that all wireline logs from a well are on depth relative to one another. Discriminant Analysis proves to be highly sensitive to errors in core-log and log-log correlations. Three digital log curves are normally used to derive the Discriminant Function in non-consolidated formations such as the McMurray Formation in Athabasca: the Gamma Ray log and the Neutron and Density logs (both in consistent porosity units). The Deep Resistivity log can also be used, but the possibility of non-lithology effects such as variable saturation and changes in R_w often make the resistivity logs difficult to work with. If the resistivity logs are used, all resistivity readings should first be converted to log base10, because resistivity readings have a logarithmic response to the formation and formation fluids. This helps produce a normal probability distribution of resistivity values for each facies. The Gamma Ray logs also require standardization to bring the sand and clay limits to a consistent response. Because the neutron log is sensitive to formation fluids, control data over gas zones can be omitted before deriving the Discriminant Function, and then manually assigned after application of the Discriminant Function to non-cored wells. Alternatively, gas-saturated sands may be considered as a distinct lithofacies and allowed to be classified as one member of the set of predicted facies. Zones of other unusual features such as tight cemented horizons and radioactive sands can be treated in a similar manner.

Additional parameters besides wireline logs can also be used in the Discriminant Function if they add information that helps differentiate lithofacies. For example, in some Athabasca cores, clay breccias are observed to occur almost exclusively towards the bottom of the middle McMurray. In this case, using stratigraphic depth as a discriminant parameter has proved to be highly effective in differentiating breccias from interbedded facies, which tend to be more common in the upper part of the reservoir. On the other hand it has been suggested that the difference between the density porosity and neutron porosity curve response be added as an additional parameter (Gupta and Johnson, 2002). Porosity curve separation is often used as an indicator of clay content in a reservoir. However, including a simple function of two other existing parameters does not add an additional degree of freedom to the Discriminant Function, and can be shown to have absolutely no effect.

Comparing the predicted facies derived from well logs (pseudocore) to the original core description generally shows a close correlation (Figure 2) except over gas zones or other unusual features as explained above. When such zones are manually adjusted, the results are highly satisfactory. In the final step of populating a data matrix, it may be judged preferable to utilise the actual core descriptions if available. Then, over intervals of lost or missing core, as well as for all wells that have wireline logs only, the derived facies can be used. Alternatively, there is a case to be made for using only the facies as derived from the Discriminant Function since these can be considered more objective and unbiased, whereas the original core descriptions are to some extent subjective and may be somewhat biased by the skills and degree of experience of the core logger.

References

GUPTA, R. and H.D. JOHNSON, 2002. High resolution facies architecture of heterolithic tidal deposits: An integrated outcrop and electrofacies analysis of a complex Reservoir. In M. Lovell and N. Parkinson (eds.), Geological applications of well logs: AAPG Methods in Exploration No. 13, pp. 161-184.

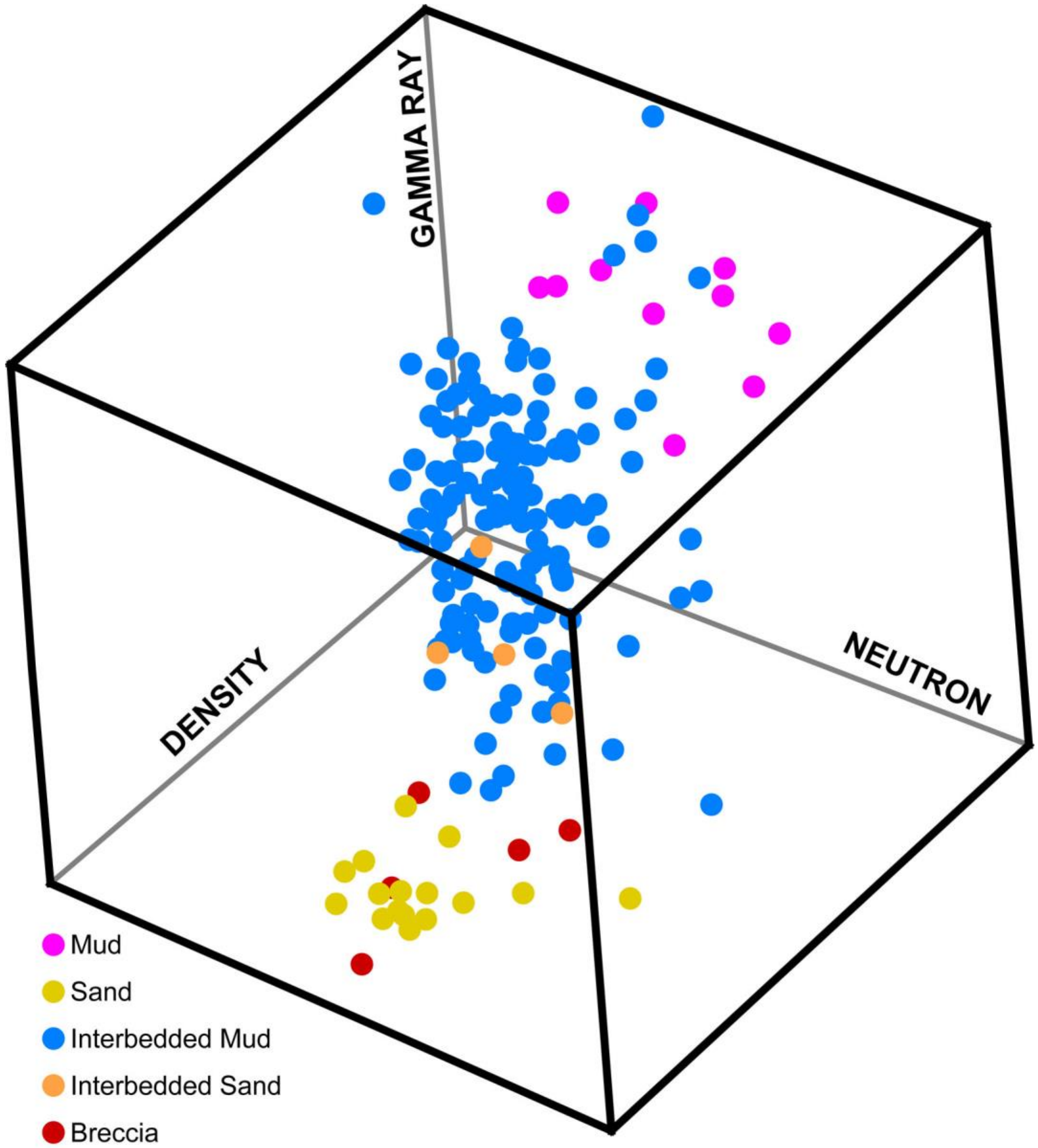

Figure 1. Facies discriminated as clusters in multi-dimensional space.

In this example from the Athabasca McMurray Formation the Gamma Ray, Density and Neutron porosity logs represent 3 axes in space. Data points are successive log readings using colour to represent facies observed in core.

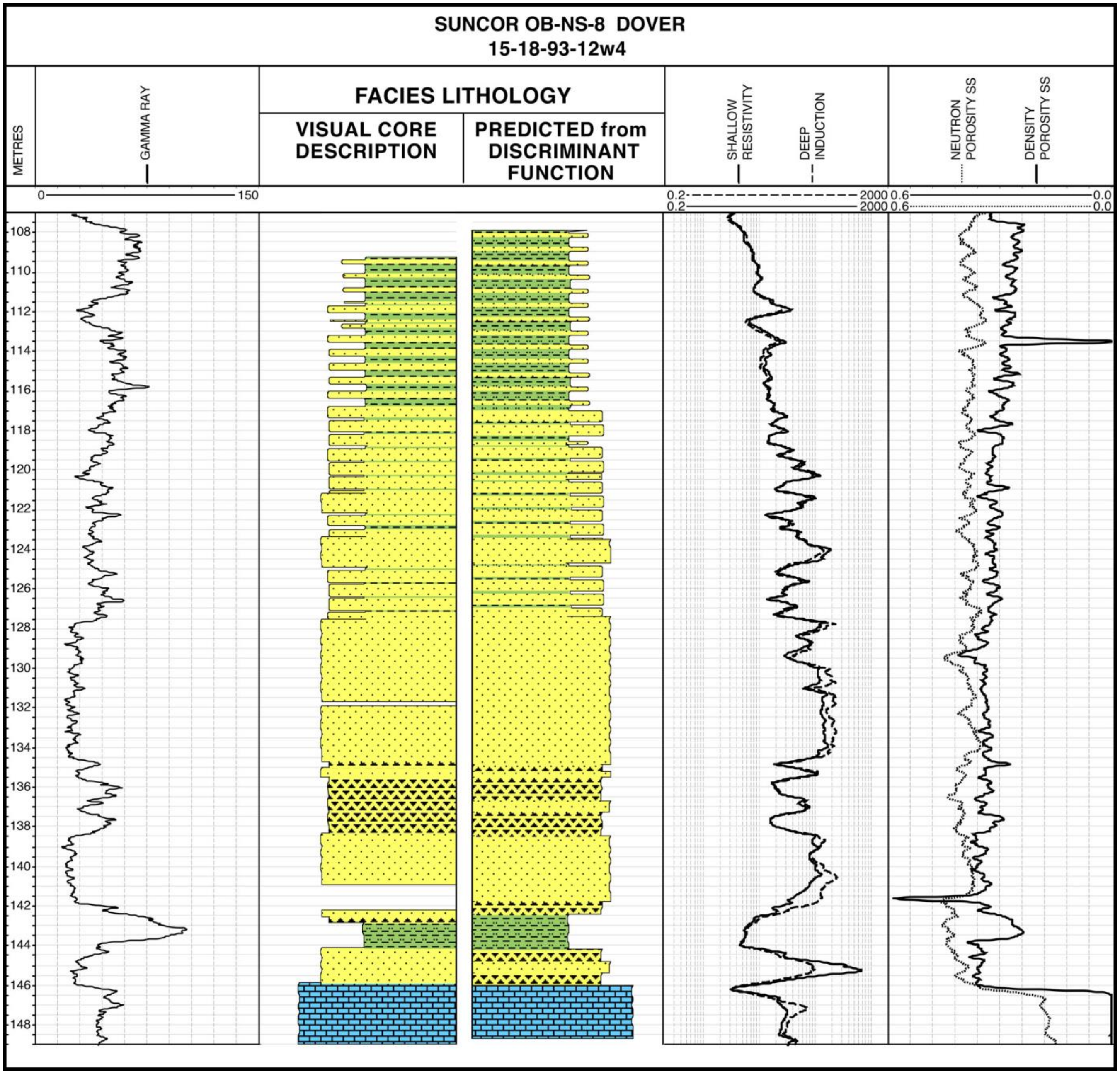

Figure 2. Predicted vs. observed facies from an Athabasca well.

Predicted facies succession on the right is the result of the Discriminant Analysis run on the Gamma Ray, Neutron and Density wireline logs. On the left is the lithofacies succession originally observed and logged from core in the same well. Lithofacies from the Discriminant Analysis are a simple superset of the observed facies (sand, mud, interbedded mud, interbedded sand, clay breccia and limestone). Note the coal response on the porosity logs at approximately 141.5 m. The coal lies within a lost core interval, and was the only occurrence of coal in the project. Therefore the Discriminant Function could not be "trained" for coal, and in this case would require manual assignment.