Identification of log-units in argillite geological formations based on local and spatial statistics of well-log properties

A. Rabaute[1], M. H. Garcia[2], J. Espitalier[2], J. Becker[3]

1. GEOSUBSIGHT, 119, Avenue de Stalingrad, 91120 Palaiseau, France (alain.rabaute@geosubsight.com)
2. KIDOVA, 155 avenue Roger Salengro, 92370 Chaville, France (michel.garcia@kidova.com)
3. NAGRA, Hardstrasse 73 Postfach 280 5430 Wettingen, Switzerland (Jens.Becker@nagra.ch)

In previous work (Garcia et al., 2011), a statistical approach was devised to identify argillaceous log-units from well-logging data and to study the spatial (lateral and vertical) variability of well-log properties in identified units. The approach combines two statistical methods: 1) Principal Component Analysis (PCA) to decorrelate the well-log properties and reduce the problem dimension to a few (one or two) components that explain most of the overall variability of the well-logging data, 2) fuzzy k-means clustering to classify the data by using the previously calculated principal components and local statistics computed from them. The advantage of the fuzzy or generalized k-means method is that it provides the probability that each well-logging data point belongs to each class, instead of just assigning each data point to one class. It also allows to determine an optimal number of classes.

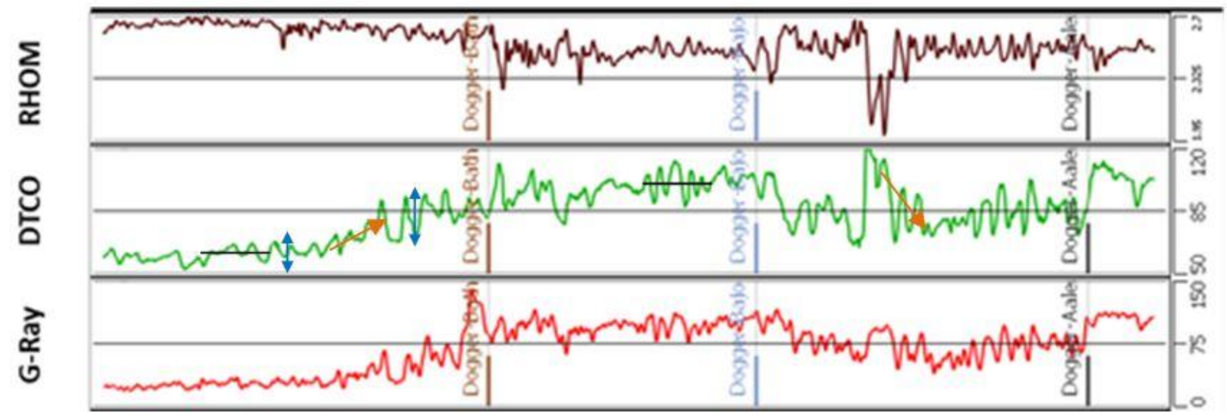

Figure 1. Example of property logs from the Benken dataset. Local statistics signatures are indicated: blue arrows for local and cyclic variabilities, black segments for stationary zones associated with plateaus, and orange arrows for transition zones showing trends.

Improvements in the statistical approach was sought to obtain classes that can be more directly related to log units by avoiding capturing unrelevant small scale heterogeneities. Two aspects of the approach have been corrected, yet preserving the principle of using PCA and k-means clustering: 1) the classification is now based only on local and spatial statistics derived from the principal components, without including their values, 2) a kernel-based method (Silverman, 1986) is used to compute local and spatial statistics. The kernel-based method allows to compute consistent statistics that are representative of an observation scale: the higher the kernel size, the higher the observation scale.

The performances of the improved approach is demonstrated on a argillite formation located in Switzerland (Benken dataset). Figure 1 shows well-logging data measured through the studied argillite formation. Three types of statistical signatures can be observed from these data: local (small scale) variability of well-log properties (blue arrows), stationary zones corresponding to constant average values (black plateaus), and transition zones corresponding to progressively increasing or decreasing property values (orange arrows). The following local and spatial statistics are proposed to capture these statistical signatures: local mean or trend (measure of location related to plateaus), local variance (measure of spread related to local variability), and trend derivative (measure of spatial variability related to transition zones).

To perform the classification, the local and spatial statistics are calculated on (decorrelated) principal components that result from relevant well-logging properties. The calculations are repeated for different kernel sizes of 1, 3 and 5m that are chosen consistently with the scales of heterogeneity to recognize (Figure 2).

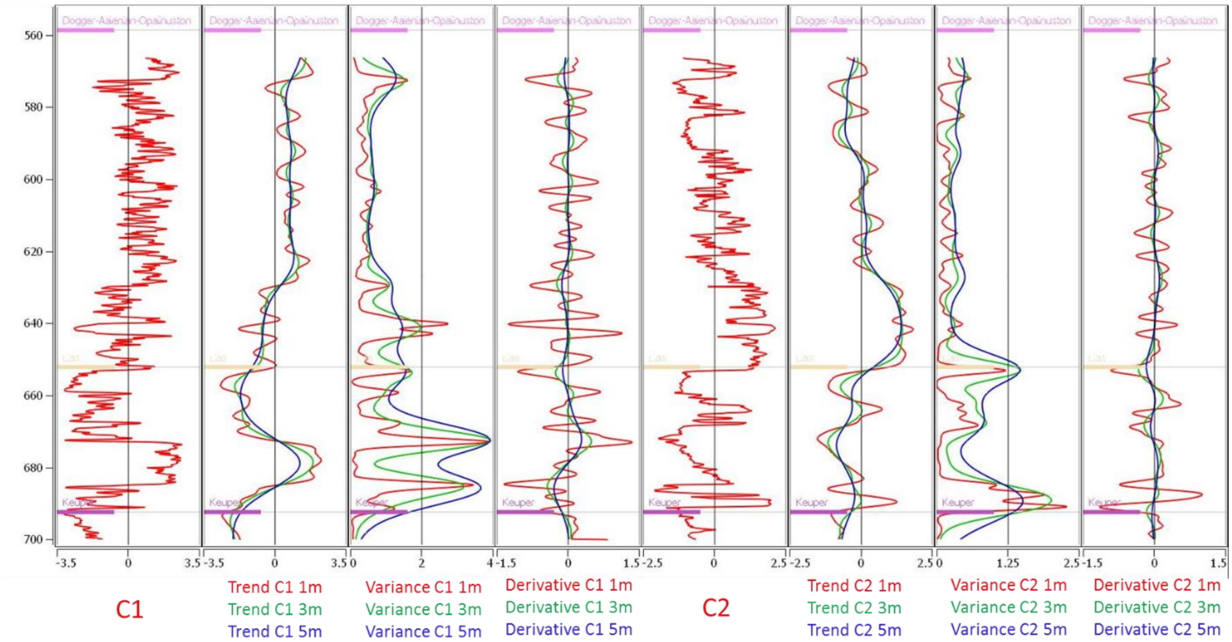

Figure 2. Logs of PC1 and PC2 over the studied argillite formation, and logs of local and spatial statistics calculated from them with kernel sizes of 1, 3 and 5 m (local mean or trend, local variance, trend derivative)

The classification based only on the variability proves to be robust in identifying log units, no matter the kernel size of 1, 3 or 5 m. The kernel-size of 1 m depicts high frequency variability, which reflects small scale lithological variations. Two main class types can be identified. The stationary classes are those for which the trend derivatives of both PC1 and PC2 stay close enough to zero to assume that PC1 and PC2 are stationary within the class interval. Inside those classes, the low or high local variance is to be interpreted as measuring the local variability of well-log properties due to small scale heterogeneities. The transitional classes are identified as associated with trend derivatives of PC1 or PC2 that is far enough from zero to reveal some continuously changing well-log properties. Figure 3 shows the classification results obtained with this dataset for the three kernel sizes. In order to interpret the classes in terms of log units, all available well-log properties can be used, not only those retained from PCA, to compute statistics per class or unit. The validity of the interpretation is checked by using drilling logs and core properties. The possibility of using the approach to identify particular lithological or hydrogeological units is also discussed.

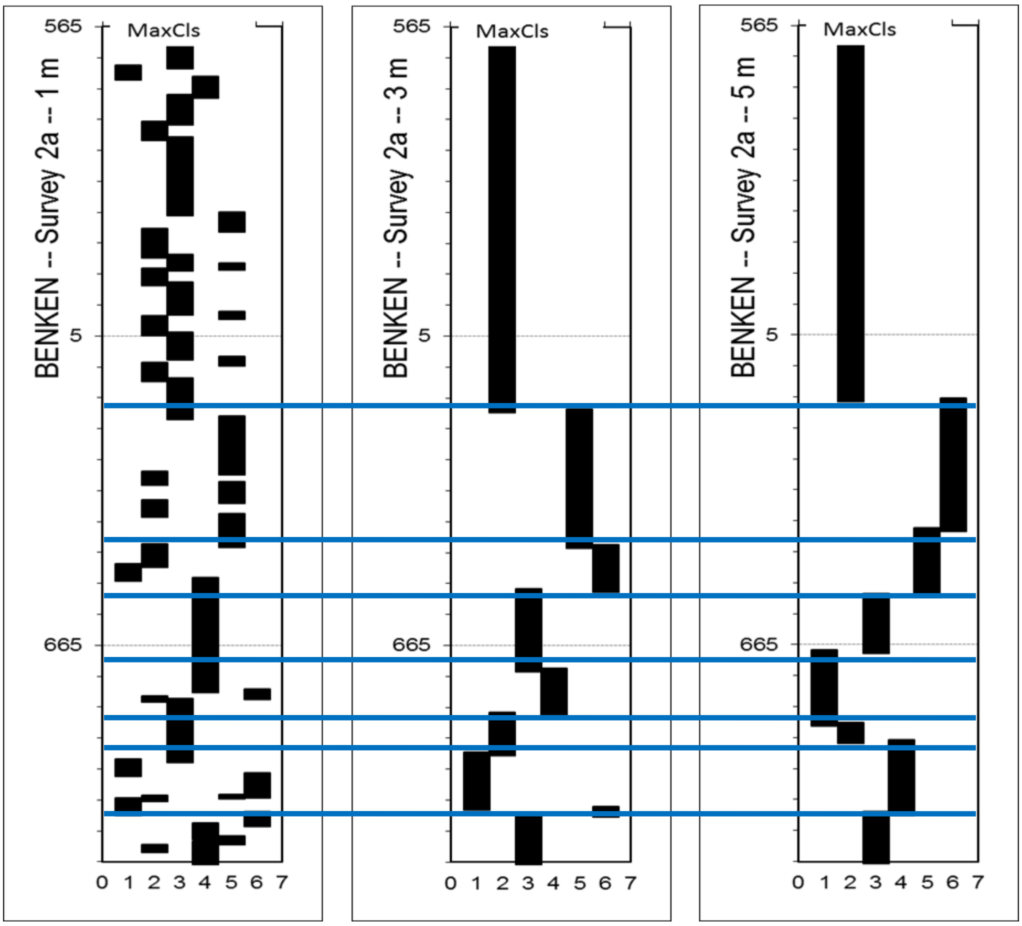

Figure 3. Classification results obtained from local statistics calculated with different kernel sizes. The blue lines show consistent log unit changes at the different observation scales. The heterogeneities captured at the scales of 3 and 5 m look similar, though their boundaries slightly differ: the higher the observation scale, the thicker the transition units and the thinner the stationary units. At the scale of 1 m, heterogeneities of smaller size are also identified, yet the higher scale heterogeneities are preserved.

References:

Garcia, M., A. Rabaute, B. Yven, D. Guillemot (2011), Multivariate and spatial statistical analysis of Callovo-Oxfordian physical properties from lab and borehole logs data: towards a characterization of lateral and vertical spatial trends in the Meuse/Haute-Marne Transposition Zone, *Phys. Chem. Earth*, Vol. 36, doi:10.1016/j.pce.2011.07.086.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.