Statistical correlation between geophysical logs and extracted core

David Price[1], Andrew Curtis[2], and Rachel Wood[2]

ABSTRACT

The precise correlation of extracted core to geophysical borehole logs is often problematic in intervals in which no sufficiently large or distinctive features span the gap between the outer core surface and the borehole wall. This impairs our ability to use techniques that require accurate correlation as a prerequisite (e.g., to develop field-based, upscaling and downscaling relationships of porosity distributions at or below log resolution). We propose a new method for such situations in which we correlate a cascade of statistics of core features with those of image logs taken from the borehole wall. Each statistic is used individually to calculate a likelihood function of possible correlation locations. These results are combined using a joint-likelihood function, and with other prior information using Bayesian techniques, to bring all available information to apply to the final correlation solution. The technique is demonstrated with computerized tomography for a core section and image-log data extracted from a typical Middle Eastern carbonate reservoir. Using lithologic criteria alone, the correlation was constrained to a 609.6 mm (2 ft) region, which constitutes prior information. Using the mean, variance, and geostatistical-range parameter, our method further constrains the correlation to 18.29 mm (0.72 in), only seven times larger than Fullbore Formation MicroImager resolution. Thus our method allows further interpretation to be based on correlation accuracies as small as 25.4 mm (1 in).

INTRODUCTION

Middle Eastern carbonate formations hold 60% of the earth's remaining hydrocarbon reserves (Ramakrishnan et al., 2001). Efficient recovery of these reserves requires an understanding of fluid flow, which in turn requires a detailed understanding of porosity and permeability within carbonate reservoirs. However, compared with siliciclastic formations, carbonate rocks have complex porosity distributions as a result of both biologic genesis and subsequent diagenetic overprinting (Choquette and Pray, 1970). The result of these processes is a highly heterogeneous porosity network with length scales of pore geometries that range over many orders of magnitude.

Most information used to quantify carbonate porosity comes from geophysical logs or extracted core. Wireline logging tools do not measure porosity directly, but measure another physical property (e.g., bulk density, neutron density, P-wave velocity, etc.) which can be converted to, or used as a proxy for, porosity. The tools typically provide continuous data coverage in each well, but with insufficient resolution to detect individual pores (Rider, 2002).

However, porosity can be imaged directly in extracted core sections, in 2D using a scanning electron microscope (SEM) or in 3D using computerized tomography (CT). To image the smallest pores in typical Middle Eastern reservoir rocks, which can be less than $1\ \mu m$ in diameter, currently require a maximum sample size on the order of 1 mm (Arns et al., 2005). Nevertheless, by using a range of different-sized core samples, a library of fabric images with known porosity distributions at multiple length scales can be constructed (Anselmetti et al., 1998). This information is limited to the regions of extracted core, which is in turn limited by the significant cost of subsurface extraction.

If low-resolution logs and high-resolution data from extracted core are available for the same interval of the well, statistical upscaling and downscaling relationships can be developed to predict relevant statistics of the detailed porosity distribution from the resolution-log data. These relationships can then be applied to regions of the well with logs, but without extracted core, improving our understanding of porosity and therefore fluid flow prediction within heterogeneous reservoirs (Delhomme et al., 1996; Tilke et al., 2006). However, there is a significant barrier to creating such relationships: relating statistics of high-resolution core information to lower-resolution log measurements requires that for each log measurement the corresponding section of core can be identified. The spatial accuracy of any correlation has to be equivalent to the resolution of log mea-

surements. In reality, other than well intervals with pronounced marker horizons, or other features that are visible both in core and logs, the accuracy with which the relative position of core and logs can be matched typically deteriorates to ~1 ft or greater. The task of matching logs and core intervals in depth is called the correlation problem.

We propose a solution to the correlation problem for depth intervals between marker features by matching a cascade of statistics of core and log measurements. Statistics are calculated for each data set individually and correlations between them are identified. The matching of individual statistics produce multiple possible correlation locations. By combining the results from a number of statistics using a joint-likelihood function, a powerful prediction tool is created.

We demonstrate the method using data from a borehole through the Lower Cretaceous carbonate Shu'aiba Formation, which holds many prolific hydrocarbon reservoirs throughout the Middle East (Buchem et al., 2002). The Shu'aiba was deposited in a shallow marine carbonate ramp setting and grades broadly from packstones and wackestones to intermixed rudist grainstones and algal boundstones (Alsharhan, 1987; Pittet et al., 2002). Much of the lime mud has recrystallized as microrhombic calcite crystals (1–10 μm), and a considerable proportion of the Shu'aiba's porosity lies between these crystals with typical pore diameters of less than 2 μm (Budd, 1989; Moshier, 1989).

We apply the developed method to correlate fullbore formation microimager (FMI) log data with a section of carbonate reservoir core scanned by CT. Although core and logs contain many more detailed features within their complex porosity distributions, these features do not span the interval between the outer core surface and borehole wall (an annulus of rock approximately 50.8 mm [2 in] thick is lost during the coring process), and hence cannot be used for direct visual correlation. As a result, it was only possible to locate the core section to within a 2-ft section of the log data using the standard method of visually matching lithologic features. We attempted higher-resolution correlations using the following porosity statistics: mean, variance, and spatial range. Each statistic was calculated over a vertical interval (rather than only a slice) of data to ensure that it was not influenced unduly by the presence or absence of vugs. As expected, these statistics produce relatively uncertain results when applied on their own. However, when combined in a likelihood function they enable the correlation to be constrained to within 18.29 mm (0.06 ft or 0.72 in).

DATA

We will demonstrate the method on the following data set: A section of extracted core was CT scanned to create a 3D volume representing variations in X-ray density. Log data were taken from a resistivity tool, the FMI. Because we use these data sets throughout this paper, we begin by describing them in more detail.

Core data

The extracted core was 120 mm long and came from a section of rudist rudstone. The only obvious large-scale features within the core are rudist shell fragments, typically elliptical, as much as 25 mm (1 in) long and 15 mm (0.6 in) wide. Outer shells of the rudist valves are composed of low-magnesium calcite and hence are generally preserved; The inner valves, originally composed of aragonite, are now represented by either recrystallized calcite or large

vugs. The core was CT scanned to produce a 3D volume made of volume elements known as voxels. The voxels are cubic with dimensions of 0.165 mm (0.006 in) which represents the resolution of the tomogram. CT assigns each voxel a value representing the average linear X-ray attenuation of material bounded within that voxel. Because no siliciclastic material has been observed in the formation, and the abundance of high-magnesium echinoderm debris is minimal, the core sample was assumed to be composed largely of low-magnesium calcium carbonate. Because the carbonate core is essentially monomineralic, changes in linear attenuation reflect changes in density (Duliu, 1999; Ketcham and Carlson, 2001). It follows that observed changes in CT value can be explained mainly by changes in porosity of material within each voxel, which is caused by variations in the distribution of micropores with length scales beneath the tomogram resolution.

Log data

The FMI has a series of electrode pads which press against borehole walls. Current is passed into the formation and measured at the electrodes. These measurements are converted to local resistivity estimates and plotted as an image. The resolution of FMI data is defined by the electrode spacing and is 2.54 mm (0.1 in) (Safinya et al., 1991). Note that this is more than one order of magnitude larger than the CT data resolution.

FMI data from the borehole were converted to a measure of porosity. Porosity estimates were calculated using the following relationship from Tilke et al. (2006):

$$\Phi = \Phi_{\text{ave}} \left[\frac{R_{\text{ave}}}{R} \right]^{1/m} \tag{1}$$

where Φ is porosity, R is the resistivity of the borehole wall, m is the cementation exponent, and the subscript *ave* indicates an average value for a given depth. The relationship is derived from the classic Archie saturation equation. The average porosity value for a given depth was taken from a conventional porosity logging tool. The cementation exponent was taken from laboratory core measurements (Ragland, 2002). In this borehole, m was 2.0 for the interval of investigation, within measurement uncertainty (Tilke et al., 2006).

FMI data were available for nearly 700 ft of the borehole and a 609.6 mm (2-ft) interval had been identified previously as the region from where the core was thought to have been extracted by using bedding planes and other distinctive lithologic marker features. Although the FMI interval had been constrained, it remained uncertain to within five times the length of the extracted core section.

DATA PROCESSING

Before any statistical methods could be used, a number of processing steps were applied to the CT scan of the extracted core to make it comparable to FMI-derived data. There were artifacts created during the scanning process which were removed. Data had to be converted to an estimate of porosity to match the FMI. Finally CT data were averaged spatially to a resolution matching that of the FMI.

CT scanners measure the attenuation of an X-ray beam passed through a sample. The sample is rotated and attenuation is measured at regular angular intervals. A back-projection algorithm is applied to the data to create a 3D-reconstruction tomogram of the sample's attenuation properties.

This technique works well for scanners using single-energy frequency X-ray beams (monochromatic signals). Most scanners, however, use X-ray beams that are composed of a range of energies, which creates artifacts in the tomogram because X-ray attenuation is energy dependent. High-energy X-rays are attenuated preferentially compared with low-energy X-rays. The incident X-ray beam therefore changes energy composition as it passes through the sample. Outer parts of the sample will be measured using the complete beam when facing the source and an attenuated beam when facing the detector. If the sample is circular, the center of the sample will be measured by a beam of similar (depleted) energy composition at all rotations. The resulting artifact, known as beam hardening, is manifested as a gradient from high CT values at the edge to low CT values at the center of the image. This gradient does not reflect actual heterogeneity within the material.

The artifact is reduced during the scanning process by using filters to preattenuate high energies in the X-ray beam (Curry et al., 1990). However, some residual beam hardening remains in the tomogram of the extracted core. It is most noticeable in a vertically averaged profile through the core (Figure 1).

To remove the artifact, the averaged vertical profile was fitted with a second-order polynomial (Figure 1). Horizontal slices were extracted from the CT volume and center of the core within each slice determined. Then each voxel within the slice was assigned a correction factor from the polynomial, determined by its radial distance from the center of the core.

Conversion of X-ray density to porosity estimates

To enable statistics to be compared with different data sources they must reflect a common property. The key to our analysis is the assumption that the core sample can be considered essentially monomineralic. Under this assumption, differences in X-ray attenuation can be attributed to averaging over varying porosity below the resolution of CT data. Therefore a method was devised to derive porosity estimates from CT data.

CT data were stored using 8 bits and hence lie in the range 0–255. The value 0 is assigned to the least attenuating voxel of the scan whereas the value 255 is assigned to the most attenuating voxel. Using the assumption above, low attenuation areas (low CT values) correspond to porous areas of the core. The lowest CT values, $C(x,y,z) = 0$, will be assigned to large pores and the air surrounding the core, and have porosity, $\Phi(x,y,z) = 1$. The high attenuation areas (high CT values) will have low porosity, closer to 0.

The following relationship was used to convert CT values to porosity estimates. It assumes a dense carbonate region exists with zero porosity at the length scale of CT data which has been assigned the value C_{max}:

$$\Phi(x,y,z) = \frac{C_{max} - C(x,y,z)}{C_{max}}, \qquad (2)$$

where $C(x,y,z)$ is a CT value at each location and $P(x,y,z)$ is the corresponding porosity estimate.

The maximum CT value, corresponding to a carbonate voxel, correlates with the outer calcite layers of rudist shells (large, oval features observed in Figure 1a. The low-magnesium calcite outer parts of rudist valves possess a prismatic microstructure with virtually zero porosity (Figure 2). This made them resistant to the recrystallization process that created microporosity in the Shu'aiba. Hence

these shell fragments have little or no porosity and correspond to high CT values compared with the surrounding matrix.

However, the maximum overall CT value within the core was 150 and occurred in slice 359 (Figure 3). This voxel does not correlate with a fragment of rudist shell. It is an isolated and sharply defined anomaly. This suggests that it was caused by an anomalous material, probably a series of diagenetic pyrite framboids.

The presence of pyrite within the core sample threatens the assumption of monomineralic composition. However, similar peaks in CT value are observed only in a limited number of slices. We therefore assume that pyrite is rare, and volumetrically insignificant, within this section of core.
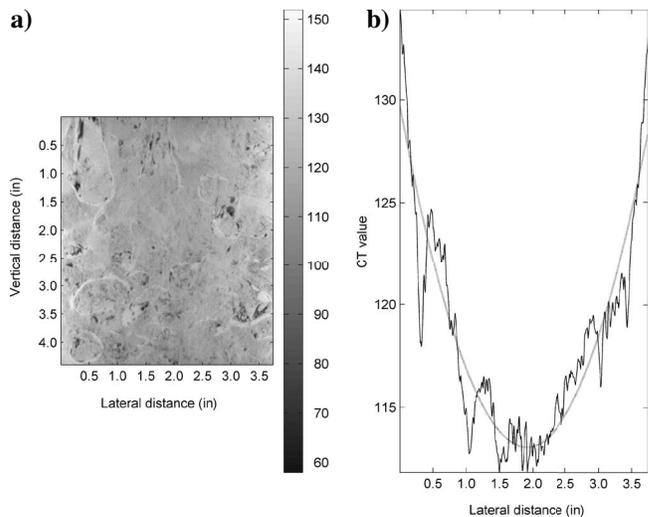

Figure 1. (a) Vertical slice through the center of the CT volume. Gray scale represents differences in CT value and hence X-ray attenuation. (b) Vertically averaged CT values from the slice (solid) with fitted polynomial (dotted) used to correct for beam hardening.
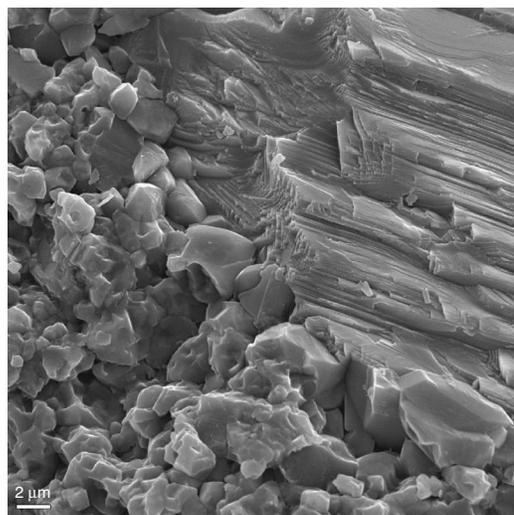

Figure 2. SEM image of carbonate matrix with microporosity (left), and the outer part of a rudist valve composed of virtually zero porosity, prismatic, low-magnesium calcite (right). The white bar (lower left) is a 2 μm length scale.

The background maximum was then taken as the maximum CT value for a carbonate voxel. This was observed to be 144 ± 2. This range of values correlates with fragments of rudist shell.

The tomogram and CT data have resolutions of 0.165 mm (0.006 in) and 2.54 mm (0.1 in), respectively. To enable statistical comparison between data sets, they were converted to the same resolution. This was achieved by averaging CT data to the same resolution as FMI data using a 3D kernel of weights:

$$Kernel(i,j,k)$$
$$= \frac{\left(1 + \sin\left(\left(\frac{i}{15}\right) \times \pi + \frac{\pi}{2}\right)\right) \times \left(1 + \sin\left(\left(\frac{j}{15}\right) \times \pi + \frac{\pi}{2}\right)\right) \times \left(1 + \sin\left(\left(\frac{k}{15}\right) \times \pi + \frac{\pi}{2}\right)\right)}{8}$$
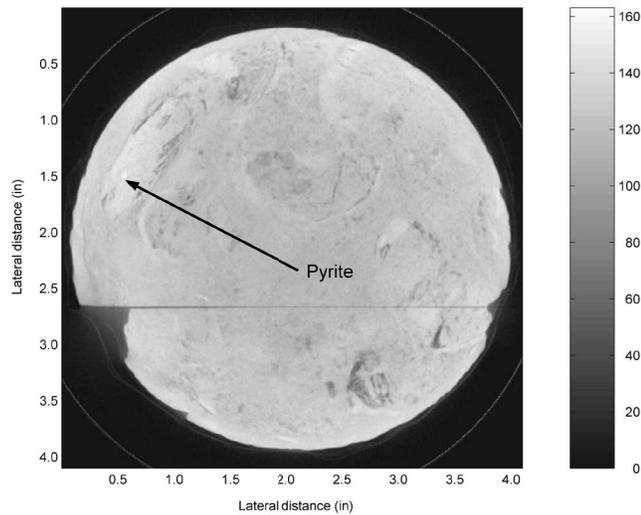
(3)

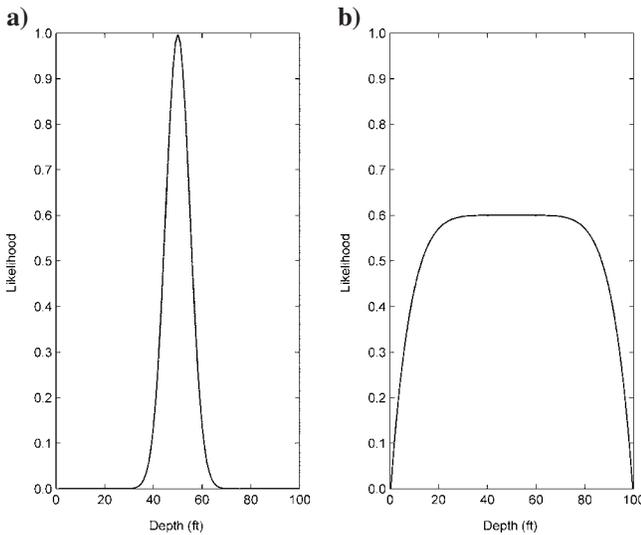Figure 3. Horizontal slice through the CT volume containing maximum CT value.

Figure 4. Two likelihood functions showing different entropy values. (a) Low entropy = large amount of information, (b) High entropy = small amount of information.

where i, j, and k are the discrete x-, y-, and z-coordinates of each voxel in the kernel discretised by the CT data at 0.165 mm. The kernel size was $31 \times 31 \times 31$ voxels. The choice of kernel size was determined to make the central section, with weights above 0.5, correspond to an FMI-sized pixel (15×15 CT voxels).

CORRELATION METHODOLOGY

To compare individual statistical values, between either FMI inside and outside the targeted depth range or CT and FMI data, we used a nondimensional misfit measure:

$$M(x) = \sqrt{\frac{(T(x) - T)^2}{\sigma^2}} \qquad (4)$$

where $M(x)$ is the misfit at depth x, $T(x)$ is the observed value of any particular statistic in FMI data at depth x, σ is the standard deviation of $T(x)$, and T is the reference value of the statistic in the area of interest. For statistics other than the mean, the value of σ^2 is taken to be the variance of each statistic $T(x)$ within the 2-ft region of the well from where the core was extracted.

The misfit for statistic i is normalized to lie between 0 and 1 by converting to a likelihood function $L_i(x)$ using (Tarantola, 1994)

$$L_i(x) = \exp\{-[M(x)]^2\}, \qquad (5)$$

which is valid for Gaussian uncertainty. We therefore assume that the uncertainty of the values of the statistics is Gaussian. The likelihood increases with improved fit (unlike the misfit), and also can be calculated for more than one statistic. A joint-likelihood function for n independent statistics is given by:

$$J(x) = \prod_{i=1}^{n} L_i(x). \qquad (6)$$

The proposed correlation method requires statistics to be identified that have predictive power when used to correlate core data to geophysical data. These statistics were identified by calculating the entropy of normalized likelihood functions, using:

$$H = -K\sum_{i}^{I} p_i \ln p_i \qquad (7)$$

where p_i is the normalized likelihood at location i and K is a constant relating to the measurement scale. The entropy value is the negative of the information represented by the normalized likelihood function. Low entropy values indicate that a likelihood function contains more information than one with high entropy values (Shannon, 1948). An example of two likelihood functions, each with different entropy values, is shown in Figure 4. In Figure 4a, the likelihood function has a low entropy value because the function has a single strong signal. In Figure 4b, the likelihood function has a very high entropy value because it is much broader and contains little or no information.

Finally, we always have prior information about the likely correlation of core and logs (e.g., from the closest marker horizons, or operationally measured core and log depths). If this information is described by a probability distribution $\rho(x)$, then our final uncertainty in correlation depth is represented by $P(x) = \rho(x)J(x)$ by Bayes's theorem. The prior is chosen to be uniform: constant within the

609.6 mm (2 ft) interval and 0 elsewhere. Hence, $P(x)$ is proportional to $J(x)$ within the interval and $P(x) = 0$ outside it.

STATISTICS

A number of statistics were applied to the data sets. The first four moments of data distributions were used, as well as the geostatistical-range parameter. Because the range parameter was calculated on a curved surface instead of a plane, a new technique for its automatic determination was developed.

The following equations 8–11 are the first four moments of the distribution of values $g_1, \ldots, g_N$:

$$\text{Mean:} \quad \bar{g} = \frac{1}{N} \sum_{j=1}^{N} g_j \tag{8}$$

$$\text{Variance:} \quad \sigma^2 = \frac{1}{N-1} \sum_{j=1}^{N} (g_j - \bar{g})^2 \tag{9}$$

$$\text{Skewness:} \quad \text{Skew} = \frac{1}{N} \sum_{j=1}^{N} \left[\frac{g_j - \bar{g}}{\sigma} \right]^3 \tag{10}$$

$$\text{Kurtosis:} \quad \text{Kurt} = \left[\frac{1}{N} \sum_{j=1}^{N} \left[\frac{g_j - \bar{g}}{\sigma} \right]^4 \right] - 3 \tag{11}$$

The mean and variance are well known. The skewness and kurtosis are nondimensional. Positive skewness values indicate a distribution with a longer asymmetric tail toward larger values and vice versa. *Kurtosis* describes the peaks or flats of a distribution compared with a Gaussian distribution, with positive values being more peaked. A potential problem is that skewness and kurtosis are higher-order statistics and hence are known to have larger sampling errors than others (Bulmer, 1979).

The variogram is a measure of expected differences between pairs of points as a function of the distance between them. The definition of variogram used in this study is:

$$\gamma(h) = \frac{1}{2n} \sum_{n} [(g(x+h) - g(x))^2], \tag{12}$$

where $g(x)$ is the porosity data at position x, h is the separation between points, and n is the number of pairs of points with a separation h (Clark, 1979).

An example of a variogram calculated for a particular window of FMI data is shown in Figure 5. The variogram follows the classic shape, rising steeply at first then leveling off. The variance at which the variogram levels off is termed the sill. The distance (or lag) at which the sill value is attained is termed the range. Conventional variogram analysis involves fitting a model variogram to the calculated variogram, which allows the variograms to be described and compared by sill and range values.

Most least-squares model-fitting algorithms implicitly assume that the number of points separated by short distances is far greater than the number separated by larger distances (this occurs when a variogram is calculated for data distributed over a plane such as a map). The assumption biases the model-fitting algorithms toward fitting the initial rise and plateau of the variogram curve which leads

to accurate estimates of the sill and range parameters. However, FMI data were collected around circular surfaces and hence have a greater number of pairs of points separated by large distances than small distances. Therefore, a new technique was required to estimate the range and sill parameters. The new technique, described in Appendix A involves, essentially, smoothing the variogram and taking the sill to be the average value at large distances and the range to be the distance at which the variogram first reaches this value.

In our method, any statistic that can be calculated on both sets of data can be used to improve the depth match. Another statistic considered was the correlation coefficient (Lisiecki and Lisiecki, 2002). This involved creating a vertical profile in both data sets and then matching profiles. However, this required matching average results taken from a 2D plane (horizontal CT slice) with those from a 1D line (horizontal FMI horizon). The 2D line was more likely to encounter large pores, vugs, than the 1D line making comparison difficult. We have used therefore statistics calculated using core length sections of data to minimize the effect of large pores.

RESULTS

Testing the differentiation power of statistics

To test the power of various alternative statistics independently from the FMI-CT correlation data methodology, we first tested the ability of each statistic to differentiate the FMI-derived porosity distribution, within the targeted 2 ft depth interval, from the FMI-derived porosity distribution in the rest of the borehole. Thus we avoid the circularity of argument of both testing and applying each statistic to the FMI and CT correlation problem.

Each statistic was calculated for core length (116 mm) windows of FMI-derived porosity data. After each calculation, the window was moved down the well by one row of data (2.54 mm [0.1 in]). This resulted in 80,478 windows and associated statistic values. Mean estimates calculated for windows of FMI data are shown in Figure 6 as a function of depth. The top plot shows results for available FMI data, whereas the bottom plot focuses on the region from where the core was extracted. The 2-ft target interval is marked as the CT interval. The average value of the mean, and its standard de-
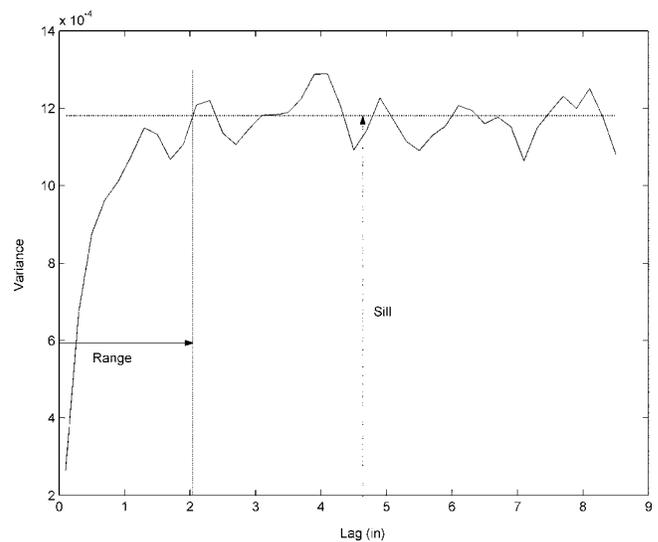

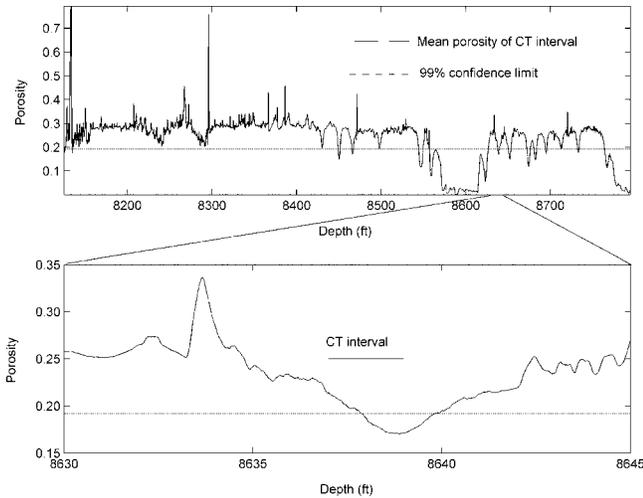Figure 5. An example of a calculated variogram with key terms used to describe its shape.

Figure 6. The calculated mean FMI-derived porosity values for windows of FMI, plotted against depth.
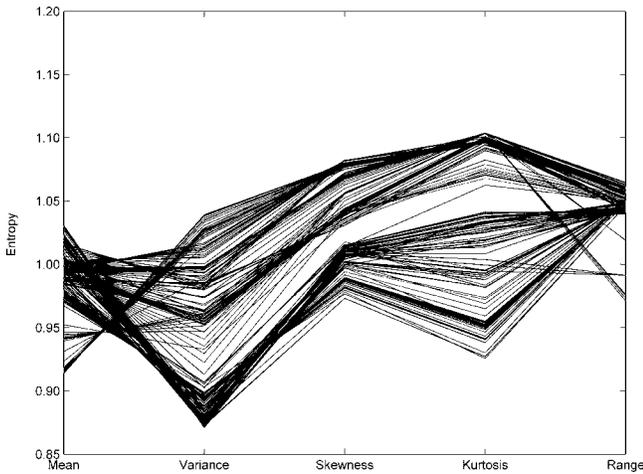

Figure 7. The entropy for each likelihood function comparing the core-sized windows within the CT interval and the rest of the core, calculated for each statistic.
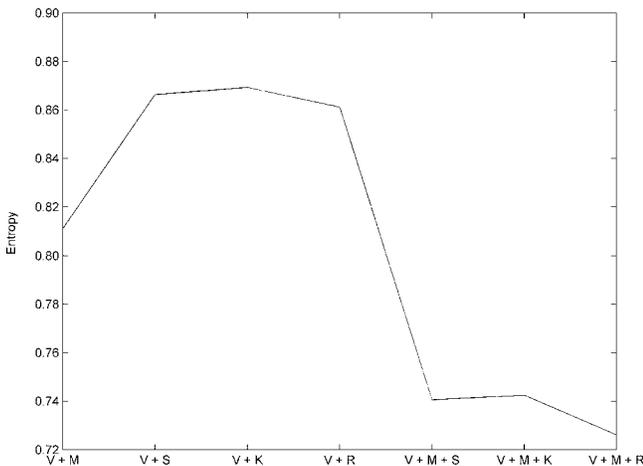

Figure 8. The entropy for each likelihood function comparing the windows within the CT interval and the rest of the core, calculated for different combinations of statistics (V = variance, M = mean, S = skewness, K = kurtosis, R = range).

viation σ within the windows of the CT interval also was calculated (Figure 6): 99.7% of the values lie within the range $\mu \pm 3\sigma$ (dotted line, Figure 6). To test the differentiation power of the mean, a set of likelihood functions (equation 5) were calculated to compare each core-length window within the CT interval against the rest of the borehole. Likelihood values of 1 indicate regions with similar mean values, whereas values close to 0 indicate regions with significantly different mean values. A selection of four likelihood functions are shown in Figure 7 for four core-sized windows within the CT interval.

To measure the depth differentiation offered by different likelihood functions, the entropy (equation 7) was calculated for each. A similar procedure was repeated for the other statistics and all calculated entropy values are shown in Figure 8 (lines connecting values using the same core-length window).

Likelihood functions with low entropy values contain more differentiation (or correlation) information than those with high entropy values. Therefore, the entropy results in Figure 8 show that the variance generally contains the greatest depth differentiation, followed by the mean. The skewness, kurtosis, and range statistics contain roughly similar amounts of information.

The appropriate order to combine the likelihood functions was determined using the entropy of the functions. A range of joint-likelihood functions were calculated using different combinations of statistics. Each combination involved the variance because it contains the most information. Resulting entropy values averaged across all windows within the CT interval are shown in Figure 9. The results show that when using two statistics, the combination of variance and mean provides the most information. When using three statistics, the addition of the range provides the most information.

CT results

The first four moments also were calculated from the averaged CT volume. The estimated value was determined using a maximum CT value of 144 in equation 2, with values of 142 and 146 being used to calculate upper and lower bounds, respectively. The estimated value for the mean was 0.214% with an upper bound of 0.224% and a low-
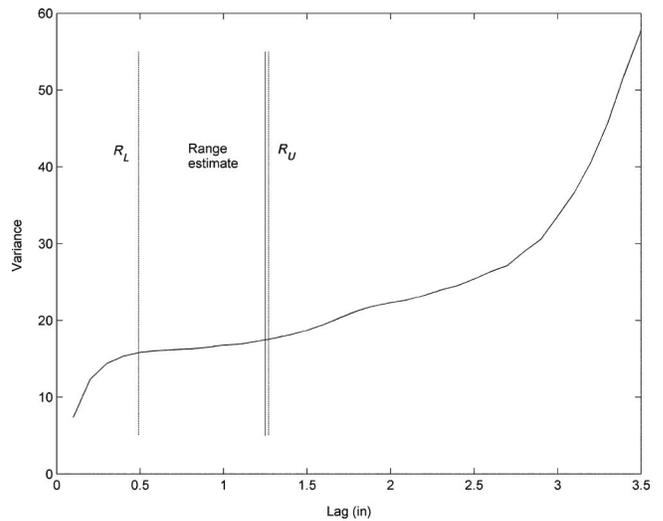

Figure 9. Horizontal variogram calculated for the averaged CT volume, with the range estimate and its upper and lower bounds marked.

er bound of 0.203%. The variance estimate was 0.00126 with an upper bound of 0.00123 and a lower bound of 0.00130. Skewness was estimated at 0.338 and kurtosis at 1.139. Skewness and kurtosis each only have one calculated value because these statistics describe the shape of the distribution and are independent of the choice of maximum CT value.

A horizontal variogram was calculated for the averaged CT data using equation 12. For CT data, the horizontal distances were assigned to 35 2.54 mm (0.1 in) bins with centers between 0.1 and 3.5 in. The calculated variogram is shown in Figure 10.

The variogram in Figure 10 is unlike the classic variogram shape shown in Figure 5. It climbs steeply, initially, then levels off at about a lag of 12.7 mm (0.5 in); it starts to climb again at a lag of 44.45 mm (1.75 in), the slope increasing with increasing lag. This behavior suggests that at least two length scale processes are responsible and this type of variogram is generally referred to as nested.

Unlike the FMI variograms previously, the averaged CT volume variogram is better determined at short lags. This is because of the larger number of pairs of points separated by short distances compared with pairs separated by larger distances within a volume. Therefore, the calculation of the range, using the same algorithm as FMI variograms, used data between 0 and 44.45 mm (1.75 in). The estimated range for the averaged CT volume was 31.75 mm (1.25 in) with a lower bound of 12.49 mm (0.49 in) and an upper bound of 32.26 mm (1.27 in).

Likelihood functions

Using the likelihood function defined in equation 5, statistics from the two data sets were compared. From the entropy results (Figure 8), the variance was shown to be the statistic containing the most information. The likelihood function for the variance values is shown in Figure 11. The three curves reflect the three estimates of porosity derived using different maximum CT values. The curves are similar and show strong likelihood values (>0.5) in the region 2632.5–2632.7 m (8637–8637.6 ft) and about 2632.88–2632.98 m (8638.05 and 8638.4 ft). These suggest that the core was extracted from within one of these regions.
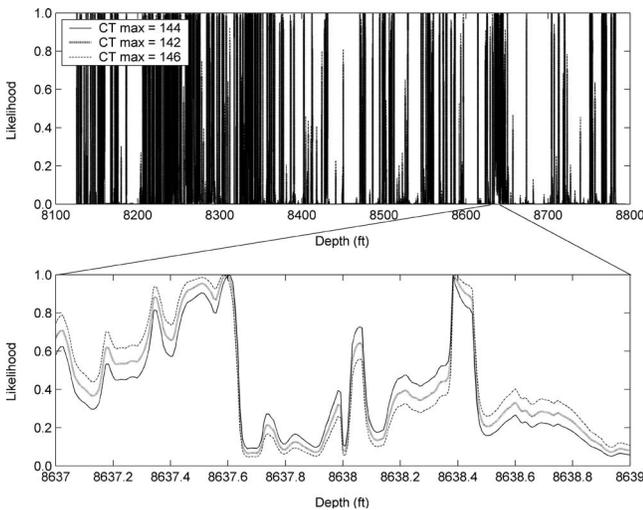

Figure 10. Comparison of CT and FMI variance values using likelihood function: (top) whole borehole; (bottom) CT interval.

The pair of statistics containing the most information was shown to be the variance and mean. The joint likelihood between these statistics was calculated and shown in Figure 12. The plots were constructed using the best estimate of CT porosity (solid line, Figure 11). The upper plot shows a great improvement in the number of possible correlation locations, with eight regions having likelihood values more than 0.5. Within the CT interval (lower plot), the joint likelihood narrows the correlation locations to the region 2632.5–2632.7 m (8637.0–8637.6 ft).

To further improve the correlation, the range statistic was added. The joint likelihood for mean, variance, and range values is shown in Figure 13. The addition of the range to the likelihood has further constrained the number of possible correlations within the well, but at the expense of a drop in likelihood strength. Within the CT interval, the addition has constrained the correlation to a single region between 2632.5 and 2632.7 m (8637.0 and 8637.06 ft).
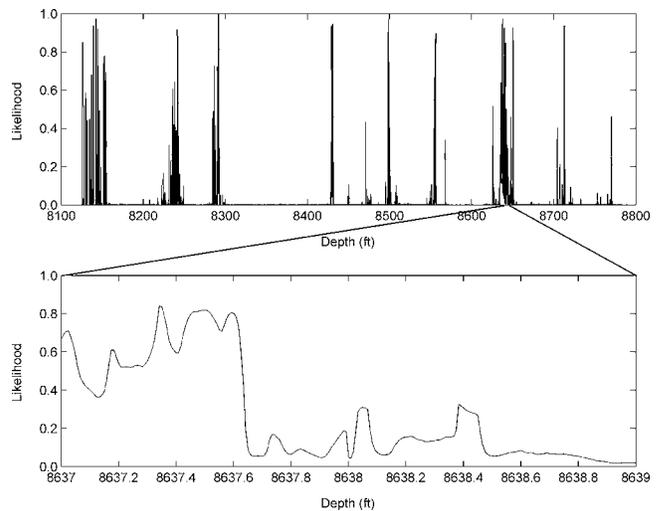

Figure 11. Comparison of CT and FMI mean and variance values using a joint-likelihood function: (top) whole borehole; (bottom) CT interval.
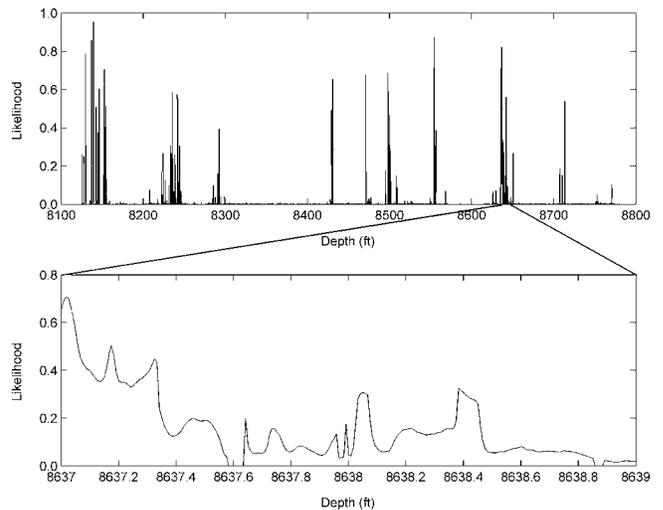

Figure 12. Comparison of CT and FMI mean, variance, and range values using a joint-likelihood function: (top) whole borehole; (bottom) CT interval.

It was hoped that skewness and kurtosis results could be added to the joint likelihood to further constrain the correlation. However, there proved to be a mismatch between the FMI and CT data sets, making these statistics unusable. For the averaged CT scan, the skewness was calculated as 0.338 and kurtosis as 1.139. These values fail to match any FMI values within the CT interval. The mismatch is attributed to large sampling errors associated with the statistics as noted earlier. The lower plot in Figure 13 therefore represents our final state of information about the depth correlation of core and FMI data because $P(x) \alpha J(x)$ within the CT interval and is 0 otherwise.

DISCUSSION

In demonstrating this method, we have used the mean, variance, and range parameter because these statistics were shown to have some predictive power within data sets. However, the method could be applied using any combination of statistics and there is nothing in the method which restricts its use to a specific rock types or formations. Hence the method can be applied to many different correlation problems.

There are a number of sources of uncertainty with the study. These mainly result from aspects of data processing, either in the choice of maximum CT value or estimation of the range parameter.

Difference in measurement conditions

The FMI data were measured at a depth of around 2600 m (8600 ft) and hence at considerable pressure. In contrast, the CT scan was carried out at the earth's surface at atmospheric pressure. It is possible that the core sample underwent decompression as it was brought to the surface, changing its porosity. To test the magnitude of this effect, the change in dimensions between a standard limestone at depth and at the earth's surface was calculated using Young's modulus (Davis and Reynolds, 1996): $E = PL/\Delta L$ where P is pressure, L is length, and ΔL is length change. We assumed that the density of limestone was 2611 kg/m³ and the overlying rocks had a similar density. The estimated value of pressure at 2600 m (8600 ft) was then estimated at 67.8 MPa. Using a value of 53 GPa for Young's
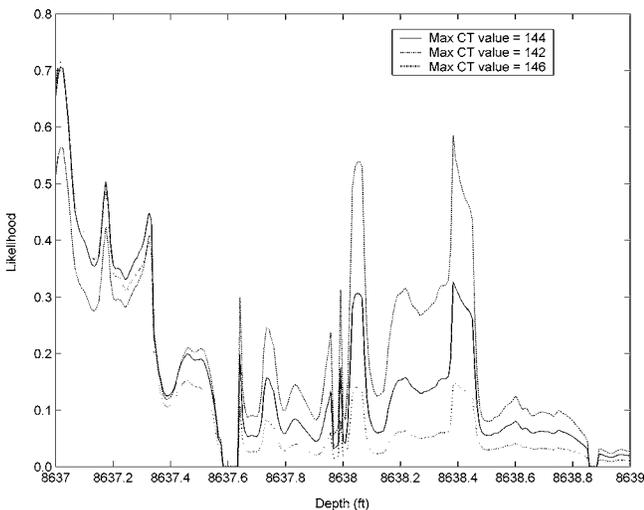
modulus in limestone, the ratio ΔL to L is calculated to be 0.0013 or 0.13%. We therefore discount decompression as a major source of error.

Error in maximum CT value

Another potential source of error is the conversion of CT values to porosity estimates. The key to the conversion process was finding a CT value representing a carbonate voxel with zero porosity. The maximum CT value in the core section was 150, but this was shown to correspond to noncarbonate features, probably pyrite. To ensure the maximum value represented a carbonate voxel, a fragment of outer rudist shell was chosen because these are known from SEM to have little porosity. The resulting maximum CT value was 144 ± 2. The range in the joint mean, variance, and range likelihood corresponding to the choice of maximum CT value is shown in Figure 14.

The single strong correlation (likelihood more than 0.5), found around 2632.5 m (8637 ft) (Figure 13), is replicated for both the upper (dotted line, Figure 14) and lower (dashed line, Figure 14) maximum CT values. The lower bound also produces two strong correlations around 2632.8 and 2632.9 m (8638 and 8638.4 ft). The insensitivity of the correlation around 2632.5 m (8637 ft) to the choice of maximum CT value makes this the most likely correlation location. The other peaks appear only for extreme values of maximum CT value.

Error in range estimate

Another source of error lies in the estimation of the range parameter from variograms. To quantify this error, an upper and lower bound also was estimated. The effect of using extreme values for the range within the joint likelihood of the mean, variance, and range are shown in Figure 15 for CT ranges and Figure A-1 for FMI ranges.

For CT range values (Figure 15), the upper bound (dotted) agrees well with the estimated value. This is expected given the similarity between range values. The lower bound (dashed) is significantly different with no strong correlations. The difference in behavior is caused by the shape of the variogram (Figure 10). The sill does not flatten completely so the range estimate and upper bound are similar. The lower bound is much smaller and lies close to the point at which

Figure 13. Range of joint mean, variance, and range likelihood values caused by choice of maximum CT value.
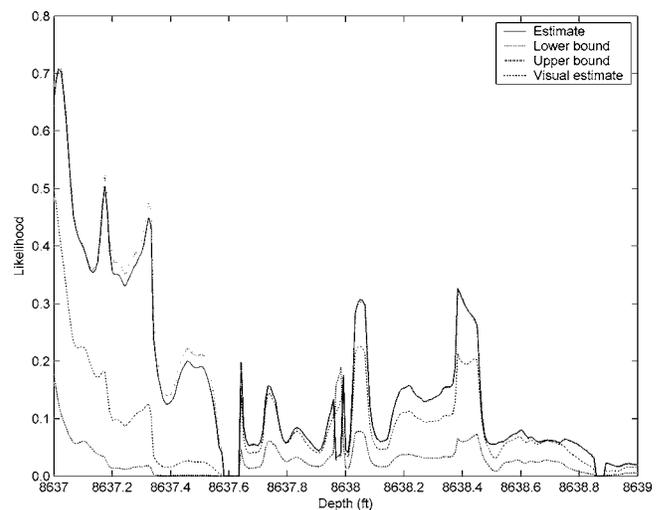

Figure 14. Range of joint mean, variance, and range likelihood values caused by estimation of CT range value.
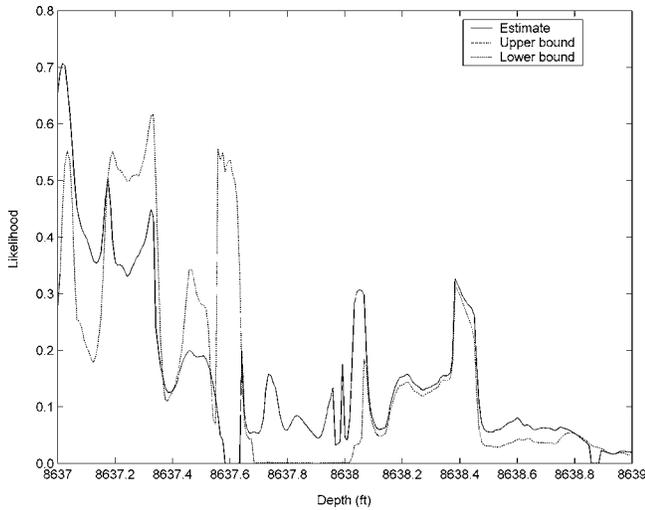
Figure 15. Range of joint mean, variance, and range likelihood values caused by estimation of FMI range values.

the variogram starts to flatten. Visual inspection of the CT variogram would suggest that the range lies midway between the calculated estimate and lower bound at 21.6 mm (0.85 in). A joint likelihood was calculated using this value (dash-dot line, Figure 15). This plots between the estimate and lower bound curves but, importantly, has its strongest values around 2632.5 m (8637 ft), consistent with the calculated estimate.

For FMI range values (Figure A-1), the upper bound (dotted line) is almost zero for the entire interval. The lower bound (dashed line) produces a number of strong correlations, one of which matches the best range estimate around 2632.5 m (8637 ft). The other strong locations occur only when using the extreme lower range value.

The likelihood plots using the extreme range values show that the choice of range value is critical. We have used the same algorithm to determine the range value for each variogram. Therefore, any bias or procedural error should be the same for each result. Results might be improved subjectively by visual inspection of each variogram, but this approach would increase greatly the time required to apply the method because of the enormous number of variograms considered.

CONCLUSIONS

A method has been proposed to constrain the correlation of geophysical logs with extracted core. A number of statistics are calculated for each data set and an entropy-based measure has been proposed to find the most informative combination. Then individual statistical correlations are combined using a joint-likelihood function and these are combined with prior information using a Bayesian methodology. The method has been demonstrated using FMI logs and CT scans of extracted core. Using mean and variance statistics proved successful in constraining the depth of FMI data relative to CT data, but failed to produce a unique depth range. Calculated higher-order conventional statistics proved unmatchable between the two data sets. A geostatistical approach using the range parameter from variograms constrained the correlation to a single depth interval of 18.29 mm (0.06 ft = 0.72 in), or seven possible FMI window positions, allowing for the subsequent development of interpretation techniques requiring accurate core-log correlation.

APPENDIX A

ESTIMATION OF VARIOGRAM PARAMETERS

Variograms calculated from FMI data are generally noisy so the data were smoothed. The smoothing was done using a 1-in moving average filter. An example of smoothed data is shown in Figure A-1 (dotted line). The sill value was then taken from smoothed data by averaging between lags 114.3–160 mm (4.5–6.3 in), an interval chosen to span part of the steady-state region and be unaffected by edge effects caused by smoothing. The range was then defined as the first crossing point of the estimated sill value with the original (nonsmoothed) variogram (Figure A-1).

The point pairs in FMI data used to find the sill are separated by large distances relative to the size of the largest significant sedimentary features in the core (rudist shell fragments). Hence the sill value should be similar to half the calculated variance within each window. The variance statistic will be used independently from this fitted model (equation 9) so the principal new statistic estimated by the above process is the range. One advantage of using the range estimate is that it does not depend on the absolute linear scaling of porosity data. Hence the range is unaffected by any errors introduced by that scaling.

To estimate the possible error in the range estimate R, the standard deviation σ_R of the original variogram from R to the maximum length scale 215.9 m [8.5 in] was calculated (horizontal dashed lines, Figure A-1). The lower bound R_L was then taken as the first crossing of the sill minus one standard deviation with the original data.
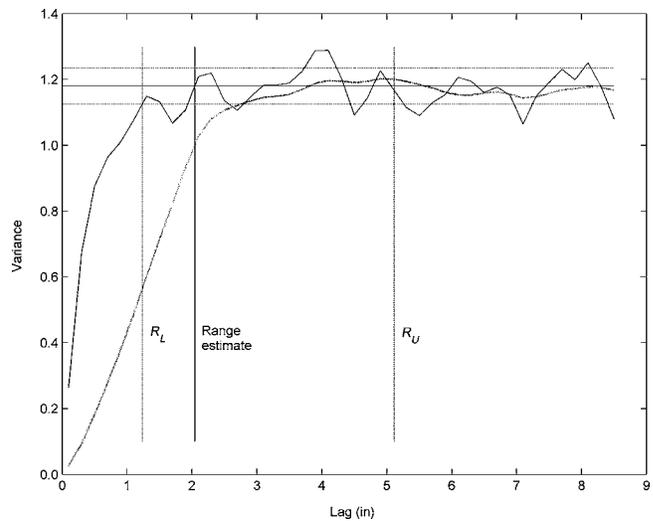

Figure A-1. Example of range fitting. The calculated variogram (solid) is smoothed (dotted) and then the average sill (solid horizontal) and its bounds (dashed horizontal) are calculated. The range estimate is the lag at which the calculated variogram and average sill first cross, R_L the lag at which the calculated variogram and the lower sill bound cross and R_U is calculated from equation A-1.

A similar approach was tried for the upper bound R_U but the results were erratic. Instead, R_U was defined as the expectation of the crossing between the sill S plus one standard deviation, and the original data, calculated

using: $R_U = \dfrac{\Sigma_x \exp\left\{-\left[\dfrac{[f(x)-(S+\sigma_R)]^2}{2\times\sigma_R^2}\right]\right\}\times x}{\Sigma_x \exp\left\{-\left[\dfrac{[f(x)-(S+\sigma_R)]^2}{2\times\sigma_R^2}\right]\right\}}$,

(A-1)

where $f(x)$ is the original variogram. An example of the lower and upper bounds R_L and R_U, respectively, are shown in Figure A-1.

REFERENCES

Alsharhan, A. S., 1987, Geology and reservoir characteristics of carbonate buildup in giant Bu Hasa oil field, Abu Dhabi, United Arab Emirates: AAPG Bulletin, **71**, 1304–1318.

Anselmetti, F. S., S. Luthi, and G. P. Eberli, 1998, Quantitative characterization of carbonate pore systems by digital image analysis: AAPG Bulletin, **82**, 1815–1836.

Arns, C. H., F. Bauget, A. Limaye, A. Sakellariou, T. J. Senden, A. P. Sheppard, R. M. Sok, W. V. Piczewski, S. Bakke, L. I. Berge, P. E. Oren, and M. A. Knackstedt, 2005, Pore-scale characterization of carbonates using X-ray microtomography: Society of Petroleum Engineers Journal, **10**, 475–484.

Buchem, F. S. P. v., B. Pittet, H. Hillgartner, J. Grotsch, A. I. A. Mansouri, I. M. Billling, H. H. J. Droste, W. H. Oterdoom, and M. V. Steenwinkel, 2002, High-resolution sequence stratigraphic architecture of Barremian/Aptian carbonate systems in Northern Oman and the United Arab Emirates (Kharaib and Shu'aiba formations): GeoArabia, **7**, 461–500.

Budd, D. A., 1989, Micro-rhombic calcite and microporosity in limestones: A geochemical study of the Lower Cretaceous Thamama group, U.A.E.: Sedimentary Geology, **63**, 293–311.

Bulmer, M. G., 1979, Principles of statistics: Dover Publishing Inc.

Choquette, P. W., and L. C. Pray, 1970, Geological nomenclature and classification of porosity in sedimentary carbonates: AAPG Bulletin, **54**, 207–250.

Clark, I., 1979, Practical geostatistics: Applied Science Publishers Ltd.

Curry, T. S., J. E. Dowdey, and R. C. Murry, 1990, Christensen's physics of diagnostic radiology: Lea and Febiger.

Davis, G. H., and S. J. Reynolds, 1996, Structural geology of rocks and regions: John Wiley & Sons Inc.

Delhomme, J. P., J. Bedford, N. M. Colley, and M. C. Kennedy, 1996, Permeability and porosity upscaling in the near-wellbore domain: The contribution of borehole electrical images: Society of Petroleum Engineers 36822, 89–101.

Duliu, O. G., 1999, Computer axial tomography in geosciences: An overview: Earth Science Reviews, **48**, 265–281.

Ketcham, R. A., and W. D. Carlson, 2001, Acquisition, optimization and interpretation of X-ray computed tomographic imagery: Applications to the geosciences: Computers and Geosciences, **27**, 381–400.

Lisiecki, L. E., and P. A. Lisiecki, 2002, Application of dynamic programming to the correlation of paleoclimate records: Paleoceanography, **17**, 1049.

Moshier, S. O., 1989, Development of microporosity in a micritic limestone reservoir, Lower Cretaceous, Middle East: Sedimentary Geology, **63**, 217–240.

Pittet, B., F. S. v. Buchem, H. Hillgartner, P. Razin, J. Grotsch, and H. Droste, 2002, Ecological succession, palaeoenvironmental change, and depositional sequences of Barremian-Aptian shallow-water carbonates in northern Oman: Sedimentology, **49**, 555–581.

Ragland, D. A., 2002, Trends in cementation exponents (m) for carbonate pore systems: Petrophysics, **43**, 434–446.

Ramakrishnan, T. S., R. Ramamoorthy, E. Fordham, L. Schwartz, H. Herron, N. Saito, and A. Rabaute, 2001, A model based interpretation methodology for evaluating carbonate reservoirs: Society of Petroleum Engineers 71704, 1–15.

Rider, M. H., 2002, The geological interpretation of well logs: Rider-French.

Safinya, K. A., P. L. Lan, M. Villegas, and P. S. Cheung, 1991, Improved formation imaging with extended microelectrical arrays: Society of Petroleum Engineers, 22726, 653–664.

Shannon, C. E., 1948, A mathematical theory of communication: Bell System Technical Journal, **27**, 379–423.

Tarantola, A., 1994, Inverse problem theory: Elsevier Science Publ. Co.

Tilke, P. G., D. Allen, and A. Gyllensten, 2006, Quantitative analysis of porosity heterogeneity: Application of geostatistics to borehole images: Mathematical Geology, **38**, 155–174.