Stochastic modeling of facies distribution in a carbonate reservoir in the Gulf of Mexico
Sanjay Srinivasan and Mrinal Sen
University of Texas at Austin, Institute for Geophysics

Abstract

We examine the basis for classification of lithofacies in the Jurassic interval of a Gulf of Mexico reservoir. Previous attempts at facies classification based on the well logs alone yielded mixed results. While some facies exhibited distinct signatures, a majority of geological facies remained unresolved using well log attributes. The correspondence between seismic facies defined on the basis of multiple seismic attributes and those observed in cores was also poor. It was therefore deemed necessary to perform seismic inversion and derive acoustic impedance volumes. We analyzed the impedance volume to define three distinct types of flow facies which were subsequently utilized within a geostatistical model for interpolating facies between wells constrained to the impedance volume. The permanence of ratios hypothesis for data integration was implemented. Our results exhibit considerable refinement over the previous models defined solely on the basis of well attributes.

Introduction

Spatial distribution of reservoir attributes such as porosity, permeability etc. in carbonate settings are controlled by the spatial distribution of vugs, dolomitization and microfractures. Seismic attributes such as impedance and amplitude can provide some indications of facies distributions in the reservoir although the resolution of seismic in many cases is much coarser than the measurement of rock types originating from other sources such as well logs, core readings etc.

A major contribution of geostatistics to reservoir modeling is the development of techniques for integrating data from various data sources. In geostatistics, the joint spatial distribution of a reservoir attribute (such as layer depth) is modeled as a spatial random function. Reservoir modeling integrating seismic and well log data has attracted a number of researchers (Xu et al, 1992; Eidsvik et al, 2001; Mukherjee et al, 2001; Larsen et al, 2006; Dubrule, 2003). The integration of seismic data can be accomplished within the random function framework using: 1) Deterministic interpolation algorithms such as, adjoint state method (Leung and Qian, 2006), Very Fast Simulated Annealing (Sen and Stoffa, 1995), Kriging with an external drift (Goovaerts, 1997) or more robustly co-Kriging (Dubrule, 2003), and 2) Stochastic imaging or simulation (Xu et al, 1992). A spatial interpolation algorithm such as kriging (or its variant kriging with an external drift) or the more robust co-kriging algorithm yields a unique map where the local estimate at every location is identified with the local conditional mean. These interpolation algorithms therefore do not provide an assessment of global uncertainty. In addition, the maps obtained from kriging are smooth and may not represent the spatial variability of the data. These issues are addressed in stochastic simulation where multiple possible realizations of the random function model are obtained, each honoring the same set of data constraints (Goovaerts, 1997). In kriging-driven sequential simulation, the local simulated value is obtained by randomly sampling from the local uncertainty distribution obtained by kriging. The reproduction of spatial uncertainty between pairs of simulated nodes is ensured by assimilating the previously simulated values in the conditioning data set for the subsequent simulation nodes. Non-kriging based simulation algorithms such as Markov Chain Monte Carlo (Larsen et al, 2006) provide realization of the reservoir model by iteratively perturbing an initial reservoir model until some target statistics of the random function model are matched. Fluctuations in the estimated model realizations provide a visual and quantifiable measure of uncertainty (Xu et al., 1992) associated with the final model.

In this paper, the use of a sequential simulation algorithm based on indicator kriging is demonstrated for modelling the spatial distribution of carbonate facies in a Gulf of Mexico reservoir. First we perform post-stack seismic inversion to obtain a 3D volume of acoustic impedance. The impedance data is calibrated against well log measurements using a probabilistic approach. The resultant probabilities are subsequently merged using a permanence of ratios hypothesis (Journel, 2002) in order to create realizations of the facies model.

Review of background information

The Jurassic age carbonate formation in the Gulf of Mexico contains an oolite reservoir facies, which is the principal exploration objective of this formation. It is divisible into an underlying, thin-bedded sandstone, shale and limestone section and an overlying limestone, dolomite and argillaceous carbonate section. The boundary between these two sections is commonly at the base of the oolite shoals. Dark, calcareous shales abruptly overlie the oolite shoals; that contact is abrupt and probably disconformable near the carbonate platform margin, while farther basinward it appears to be gradational.
Well log suites for a set of wells intersecting the Jurassic carbonate formation were available. In addition 3D seismic volumes of different attributes were also provided. Multiple seismic attributes were processed through probabilistic neural network software in order to come up with a seismic facies volume. The analysis yielded six main facies. The following correspondence between the six facies indices and the corresponding facies description based on cores is tabulated:

<table>
<thead>
<tr>
<th>Facies from Neural Network</th>
<th>Corresponding seismic Facies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mudstone/Shaly Mudstone</td>
<td>1</td>
</tr>
<tr>
<td>Mudstone-Wackestone</td>
<td>2</td>
</tr>
<tr>
<td>Wackestone-Packstone</td>
<td>3</td>
</tr>
<tr>
<td>Packstone-Grainstone</td>
<td>4</td>
</tr>
<tr>
<td>Green Lutites/Green Grayish</td>
<td>5</td>
</tr>
<tr>
<td>Dolomitic Breccia</td>
<td>6</td>
</tr>
</tbody>
</table>

Facies logs based on the seismic facies are available for most of the wells. A three-dimensional reservoir model including the facies model developed on the basis of well facies data is available. In addition, information about the reservoir horizons and location of faults are available.

**Methodology**

**Simple Statistical Analysis of Well Log data**

Prior to performing extensive multivariate analysis for modeling spatial distribution of carbonate facies, preliminary statistical analysis of the well log attributes was performed. The four attributes selected for the analysis are gamma ray, neutron porosity, bulk density, and sonic porosity. Figure 1 shows the distribution of gamma ray values for the seismic facies 1-4 and 6. As can be seen Facies 1 and 3 exhibit distinct difference in the range of gamma ray values and the shape of the distribution. However, the distributions for other facies show significant overlap.

Fig. 1 Comparison of histograms of gamma ray values in different facies.
Similar distributions were plotted for the other log attributes. These also reveal differences between facies 1 and 3, but indicate that effective classification of facies would probably require multi-attribute analysis since any one attribute may not be able to discriminate the seismic facies accurately.

Comparing facies defined using logs with those from seismic analysis

A detailed lithological description of the core obtained in the wells penetrating the formation of interest was available. These core descriptions were subsequently related to the suite of well log attributes (GR, CGR, LLD, LLS, MSFL, NPHI, Sonic TWT, RHOB) available at the same depths. Neural network analysis of the log attributes yielded well log cut-offs and subsequently these log cut-offs were used to define well facies.

The spread in log attribute values within the facies based on seismic analysis are compared to those within facies established by well log analysis in Figures 2 through 4 below.

The results of the comparison of the log-based and seismic facies models, in terms of their porosity and lithology
logs, suggest that it is very difficult to classify the seismic facies on the basis of log variations. Due to the poor correspondence between the seismic attributes and the well log based facies classification model and the impracticality of defining 6 or 7 facies on the basis of smooth seismic variations, an alternate approach based on seismic impedance analysis and integration of the impedance volume within a geostatistical model for facies distribution was developed. The various steps and results obtained by this procedure are described next.

**Seismic Inversion**

The seismic data used in post-stack inversion are listed below:

- A sub-volume of the entire field was retained for seismic inversion and subsequent model construction
- Number of samples/trace = 1251
- time sampling interval = 4ms
- Total no of traces used in inversion = 1202301
- Total no of inlines = 801
- Total no of crosslines = 1501

Two wells from within the volume were used in the inversion. The time interval for inversion was consistent with the horizons picked by interpreters. The essential steps to post-stack inversion include (1) wavelet estimation and (2) background model generation. A general flow chart for post-stack inversion is given in Figure 5 (Sen 2006).

Two wells A-1 and A-2 were used in the generation of background models. The logs were interpolated and extrapolated guided by the top and bottom horizons. A very high frequency background model and a very smooth background model were used in the inversion. Results from those will be reported here.

Wavelets were estimated at the two well locations from wells using Hampson Russell® software. An average wavelet shown in Figure 6 was developed and was used in the inversion after calibration at the well locations.

**Inversion Results and Data Integration**

The amplitude was calibrated at only one location and we used a constant scale factor in the entire data volume. Figures 7 through 9 show calibration of the results at well A-1. Note that the inverted results match the well data very well. Similar results were also derived at well A-2.

As demonstration we also include 2D sections along the two wells (Figures 9 and 10). Acoustic impedance variations within the target zone are easily identified. Similar inversions were also carried out using low frequency starting models and used in the interpretation of facies described subsequently.

---

**Fig. 5** A general flow-chart for acoustic impedance inversion using post-stack seismic data.
Fig. 6 Average wavelet used in the post-stack inversion—note the -90º phase shift of the wavelet.

Fig. 7 Seismic trace at well A-1 in blue and synthetic trace (red) using a high frequency starting model.

Fig. 8 Seismic trace at well A-2 in blue and synthetic trace (red) using final model derived by inversion: Note the excellent match between the two.
Fig. 9 Match of inverted acoustic impedance (blue) with the true log (red) at the well A-1.

Fig. 10 Acoustic impedance derived along an inline containing the well A-2. Note the target zone with high frequency acoustic impedance.
Conditioning Facies Model to Impedance Information

Once the impedance volume is available, there are several approaches to integrate the impedance data with the data inferred from well logs in order to construct reservoir models that describe the spatial variations of lithofacies. However, prior to integrating the seismic impedance data into a reservoir model it is necessary to establish the correspondence between the impedance data and the reservoir facies. Since the geological facies models are based on thresholds defined on various log attributes, the correlation between impedance and gamma ray values was checked first. Figure 12 shows the correlation between seismic impedance and gamma ray log values along wells. Since the seismic impedance values are available at 4 ms sampling interval, corresponding log values were also retrieved at the same interval. The figure indicates that the raw correlation between coincidental values is poor.

Fig. 11 Acoustic impedance derived along an inline containing the well A-1. Note the target zone with high frequency acoustic impedance.

Fig. 12 Raw point-to-point correspondence between GR and acoustic impedance values along wells.

Fig. 13 Variation in correlation between seismic impedance and gamma ray for different averaging scales.

\[ \rho = 0.056 \]
In order to understand the reason for this poor correlation between log and seismic values, volume averages over various volume supports were computed. Figure 13 shows the variations in the correlation coefficient between impedance and gamma ray for different levels of upscaling. As can be seen, maximum correlation is obtained when the time interval between the top and bottom horizons is divided into 20 sub-intervals.

Figure 14 is the correlation between the upscaled gamma ray and impedance values at the optimum averaging scale. This improved correlation was obtained after removing two outlier values from the scatter plot.

The improved correlation plot between seismic impedance and gamma ray can now be subjected to cluster analysis. The k-mean clustering algorithm was implemented within an iterative scheme where the objective was to maximize the following distance measure:

\[ d = E \left( d_{\bar{k}} - d_{\bar{k}' \bar{k}} \right)^2 + E \left( d_{\bar{k}} - d_{\bar{k}} \right)^2 \]

where \( d_{\bar{k}} \) are the data points classified in the \( k^{th} \) cluster, \( d_{\bar{k}} \) is the corresponding cluster mean. The objective is therefore to maximize the inter-cluster distance (the first term on the RHS) and minimize the intra-cluster distance (the second term). Figure 14 is the results of the cluster analysis performed using the GR-impedance cross plot and with 3 clusters. Analyses performed with 2 and 4 clusters resulted in a lower value of the distance measure \( d \).

Figure 15 Results of the cluster analysis performed assuming 3 clusters. The label \( x \) represents standardized gamma ray and \( y \) standardized impedance.

Figure 16 Results of the cluster analysis between GR-bulk density-impedance performed assuming 3 clusters. The label \( x \) represents standardized gamma ray (left) and standardized bulk density (right). The \( y \) axis represents standardized impedance in both figures.
Cluster analysis was repeated with three attributes—gamma ray, bulk density, and seismic impedance. Figure 16 shows the resultant clusters as seen in the gamma ray-impedance scatterplot (left) and the bulk-density-impedance scatterplot (right). It is evident that the multi-attribute cluster analysis reveals hidden groupings in the data, however since the clusters do exhibit some overlap in impedance values, it was deemed better to rely on the clusters defined on the basis of gamma-ray and impedance.

Figure 17 shows the distribution of gamma ray within the three clusters identified by cluster analysis.

The gamma rays in cluster 1 are clustered around low values indicating relatively low proportion of clays. The gamma ray values in cluster 2 are high indicating clay rich rock mass. Finally, the gamma rays in cluster 3 are variable and span the range from low to medium gamma ray values. The three clusters are therefore indicative of three carbonate facies—interconnected or vugular porosity (Cluster 1), clay dominated lithologies (cluster 2) and matrix dominated porosity (Cluster 3). The spatial modeling of these three facies that can be detected on the basis of seismic impedance is then the objective of the study.

Data Integration using the Permanence of Ratio Hypothesis

The fundamental paradigm in geostatistics is the sampling of realizations of spatially correlated variables from the multivariate conditional distribution $P(A|B,C,…)$, where $A$ represents the property being modeled, $B$ and $C$ are conditioning information of various kind. Cokriging and co-simulation (Deutsch and Journel, 1998; Goovaerts, 1997; Doyen, 2006) permit spatial modeling of $A$ conditioned to the available information by providing an avenue to model the multivariate conditional probability distribution utilizing spatial cross-correlation between the various attributes. Since cross-covariances and variograms are only the moments of the multivariate joint distributions, and these are sufficient to capture the variability of only symmetric (Gaussian) distributions, a more powerful approach for data integration

![Histogram of Gamma Ray (Facies 1)](image1)

![GR - Facies 1](image2)

**Fig. 17** Distribution of gamma ray values within the 3 clusters identified by cluster analysis.
would be one that is non-parametric and attempts to model the multivariate distribution using a set of selected thresholds. Such an avenue is afforded by the permanence of ratio hypothesis (Journel, 2002). In this hypothesis, the multivariate conditional distribution is modeled by merging the univariate distributions obtained by conditioning to individual conditioning data. Thus,

\[ x = \left( \frac{b}{a} \right)^{\tau_1} \left( \frac{c}{a} \right)^{\tau_2} \]

where the measure \( x \) is defined as the discriminant

\[ x = \frac{1 - \Pr \{ A \mid B, C \}}{\Pr \{ A \mid B, C \}} \]

i.e., \( x \) is the relative distance to the occurrence of \( A \) due to information in \( B \) and \( C \). The other measures are similarly defined:

\[ a = \frac{1 - \Pr \{ A \}}{\Pr \{ A \}}, \quad b = \frac{1 - \Pr \{ A \mid B \}}{\Pr \{ A \mid B \}}, \quad c = \frac{1 - \Pr \{ A \mid C \}}{\Pr \{ A \mid C \}} \]

The exponents \( \tau_1 \) and \( \tau_2 \) are regulated by the redundancy between the information in \( B \) and \( C \) (Krishnan, 2007). In the case of unit exponents, the expression above reduces to:

\[ x = \left( \frac{c}{a} \right) \]

i.e., the relative distance to the occurrence of \( A \) due to \( C \) remains unaffected by the occurrence of event \( B \).

The above expression provides a powerful approach to combining elemental conditional probabilities derived from various data sources and this is exploited for the task of facies modeling. The conditional probability \( \Pr(A|B) \) reflects the uncertainty in the facies at a location due to the facies information available at wells \( B \). This conditional probability distribution is that derived by indicator kriging and utilizes the prior geological information in the form of indicator variograms. The information available in the seismic data is represented by the term \( \Pr(A|C) \). Since the relationship between seismic and well facies is imprecise, the information in seismic is quantified in the form of probabilities.

Bayes’ rule is used to derive the conditional probability \( \Pr(A|C) \):

\[ \Pr(A|C) = \frac{\Pr(C|A) \cdot \Pr(A)}{\Pr(C)} \]

where the likelihood is inferred from data by plotting the histogram of seismic impedance within each cluster.

Figure 18 shows the histogram of seismic impedance within each cluster. In order to overcome the problem in inferring the likelihood \( \Pr(C|A) \) due to sparse samples belonging to some thresholds, smoothing of the experimental histogram was performed. This smoothing procedure is non-parametric and utilizes simulated annealing in order to ensure that the
smoothed histogram reproduces key statistics. This procedure for smoothing is akin to the process of modeling experimental distributions using a mixture of Gaussian kernels which is a key step in a probabilistic neural network. The process for modeling the conditional distribution P(A|C) described here is therefore analogous to the process using a probabilistic neural network except that it is non-parametric and does not require the use of Gaussian probabilities etc.

Figure 18 indicates high impedance values corresponding to cluster 2, which already has been identified as a clay-rich facies. Cluster 1 has lower impedance values and corresponds to the facies with inter-connected porosity. The impedance values in cluster 3 span the range from low to high and represent the matrix-dominated facies.

Using the proportion of facies computed along wells, P(A) and the prior histogram of impedance P(B), the conditional probabilities P(A|C) for various thresholds of impedance were calculated. These are tabulated in Table 3 below. As can be seen, there is clear discrimination of facies corresponding to some impedance thresholds. However, that discrimination is ambiguous at some other seismic thresholds.

Table 2: Tabulation of conditional probability P(A|C) for different facies conditioned to impedance threshold values.

<table>
<thead>
<tr>
<th>Impedance Threshold</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>7400.00</td>
<td>2.9131303E-02</td>
<td>1.4415875E-04</td>
<td>0.9707246</td>
</tr>
<tr>
<td>7800.00</td>
<td>0.1033083</td>
<td>7.0227310E-05</td>
<td>0.8966215</td>
</tr>
<tr>
<td>8200.00</td>
<td>0.3152358</td>
<td>7.8684359E-05</td>
<td>0.6846855</td>
</tr>
<tr>
<td>8600.00</td>
<td>0.4852211</td>
<td>1.0350616E-04</td>
<td>0.5146754</td>
</tr>
<tr>
<td>9000.00</td>
<td>0.6833298</td>
<td>1.9665471E-04</td>
<td>0.3164736</td>
</tr>
<tr>
<td>9400.00</td>
<td>0.7949884</td>
<td>6.4031931E-04</td>
<td>0.2047313</td>
</tr>
<tr>
<td>9800.00</td>
<td>0.7810345</td>
<td>4.1399468E-03</td>
<td>0.2148256</td>
</tr>
<tr>
<td>10200.00</td>
<td>0.6066645</td>
<td>2.0404505E-02</td>
<td>0.3729310</td>
</tr>
<tr>
<td>10600.00</td>
<td>0.4060703</td>
<td>4.2640947E-02</td>
<td>0.5512888</td>
</tr>
<tr>
<td>11000.00</td>
<td>0.3533477</td>
<td>6.6490807E-02</td>
<td>0.5801615</td>
</tr>
<tr>
<td>11400.00</td>
<td>0.3196448</td>
<td>1.096322</td>
<td>0.5070229</td>
</tr>
<tr>
<td>11800.00</td>
<td>0.3126995</td>
<td>0.1552853</td>
<td>0.5320152</td>
</tr>
<tr>
<td>12200.00</td>
<td>0.5317208</td>
<td>8.7327652E-02</td>
<td>0.3809516</td>
</tr>
<tr>
<td>12600.00</td>
<td>0.6786323</td>
<td>8.6726435E-02</td>
<td>0.2346413</td>
</tr>
<tr>
<td>13000.00</td>
<td>0.6863726</td>
<td>0.1542912</td>
<td>0.1593362</td>
</tr>
<tr>
<td>13400.00</td>
<td>0.6346956</td>
<td>0.1865951</td>
<td>0.1787092</td>
</tr>
<tr>
<td>13800.00</td>
<td>0.2572277</td>
<td>0.1199410</td>
<td>0.6228313</td>
</tr>
<tr>
<td>14200.00</td>
<td>3.0672975E-02</td>
<td>1.5284929E-03</td>
<td>0.9677985</td>
</tr>
<tr>
<td>14600.00</td>
<td>2.5825975E-02</td>
<td>2.4579307E-03</td>
<td>0.9717161</td>
</tr>
<tr>
<td>15000.00</td>
<td>2.2859737E-02</td>
<td>2.1596279E-03</td>
<td>0.9749807</td>
</tr>
</tbody>
</table>

Once the conditional probabilities on the basis of seismic are available, the permanence of ratio hypothesis can be applied. The remaining parameters are the factors 1 and 2 as discussed before, reflect the redundancy between the different pieces of information. In this case 1 was set to be 1, thereby maximizing the influence of the primary well data, while 2 was set equal to 0.4 which is approximately the correlation coefficient between GR-Impedance reflected in Figure 13.

Figure 19 depicts the experimental indicator variograms corresponding to Facies 1 and Facies 2.
Table 3 Indicator variogram model parameters for the three facies.

<table>
<thead>
<tr>
<th>Facies</th>
<th>Structure 1</th>
<th>Structure 2</th>
<th>Final Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Isotropic</td>
<td>135° dn.</td>
<td>The facies model that resulted from utilizing the permanence of ratio hypothesis is shown in Figure 19. The simulation grid is of dimension 1501 × 801 × 20. This reflects the seismic grid dimension of 1501 cross lines for each inline and a total of 801 inline values. For comparison, a corresponding slice from an indicator simulation model obtained unconstrained to the seismic data is also shown. The seismic impedance slice corresponding to the facies slice is also shown in Figure 19.</td>
</tr>
<tr>
<td>Facies 1</td>
<td>100 units, 8 units vert. 350/200, 8 units vert.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facies 2</td>
<td>100 units, 8 units vert. 350/200, 8 units vert.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facies 3</td>
<td>70 units, 6 units vert. 200/150, 8 units vert.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 20 A slice through the 3D facies model (top left) obtained constrained to the seismic data. A corresponding slice through the model unconstrained to seismic data is shown in the top right figure. The corresponding seismic impedance slice is shown in the bottom. The regions that have been updated significantly have been identified.
From Figure 20 it is evident that regions that were originally simulated to be facies 3 (in yellow - the matrix dominated facies) are updated to facies 1 (in red-the interconnected vugular facies). These changes are brought about because of lower seismic impedance values in those areas (see within the marked areas). Figure 21 shows the same comparison in the form of XZ slices. The seismic impedance data serves to make the good (interconnected vugular) facies to exhibit better continuity in the vertical direction.

Conclusions

We performed a thorough analysis of carbonate facies definition for the Jurassic interval of a Gulf of Mexico reservoir. The analysis revealed that facies definition on the basis of log attributes is difficult since the subtle variations in carbonate facies cannot be resolved by the logs. We also analyzed seismic facies defined on the basis of multi-attribute seismic. However, there is poor correspondence between the seismic facies and those defined on the basis of log attributes. For this reason we felt that it was necessary to perform a post-stack seismic inversion in order to derive acoustic impedance volume that was subsequently linked to facies variations as indicated by log attribute variations. Cluster analysis of the scatter plot between gamma ray log values along wells and the corresponding seismic impedance values was performed in order to define effective flow facies. It is implicit in this analysis that the seismic data will be insufficient to resolve subtle variations in carbonate facies as defined on the basis of geology. In lieu of such a detailed classification, three effective facies were defined-vugular interconnected facies, matrix dominated poorly connected facies and clay dominated facies.

Once the relationship between the inverted impedance values and these effective flow facies at well locations were established, we performed geostatistical interpolation to populate the entire reservoir model with effective flow facies. The permanence of ratio hypothesis was used to combine the information calibrated from seismic impedance with the facies data at the wells. The permanence of ratio hypothesis is based on the calibration of conditional probability information on the basis of the seismic data and also deriving the conditional probability on the basis of well data and the assumed spatial continuity model and subsequently merging those two distributions accounting for redundancy in the information. The resultant facies model was compared against that obtained using only the well information. It was observed that the seismic data helps resolve the effective flow facies much better and helps
represent the spatial continuity of facies much better.

References


Larsen, A., M. Ulvmoen, H. Omre and A. Buland, 2006, Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model: Geophysics, 71, R69-R78


Sen, M. K., 2006, Seismic Inversion, SPE publications, Dallas, USA.

