Mining the database

My kingdom for a trend: An analysis is building regional geologic models for use in basin modeling.

AUTHORS

Brian Locke, *Samuel Mentemeier, and*
Brian O'Neill, *Anadarko Petroleum and*
H. Roice Nelson Jr., *Geophysical Development Corp. (GDC)*

As the speed of computational time exponentially accelerates, databases seem similarly to be getting larger and, unfortunately, more unwieldy. With almost unlimited disk space, more and more data is getting pulled into the potential black hole called a database. To navigate this universe of data, it is crucial to be able to scrutinize the contents and quality of a database to effectively identify key trends in the data. It is important to start with a high quality, well quality-controlled (QCd) database such as the Geophysical Development Corp. (GDC) rock property database, but it is equally important to have a capable vehicle which is able to sort through the data swiftly. Spotfire Decisionsite is data mining software that allows the user to simultaneously plot multiple parameters in a variety of 2-D and 3-D linked views. Spotfire, combined with the GDC database, generates a virtual playground for discovering rock property trends. The purpose of this article is to demonstrate a few of the ways Anadarko has recently used Spotfire DecisionSite data mining software to identify rock property trends from the GDC database. These insights have proven to be extremely valuable for evaluation of pore pressure, average density, velocity, sand footage and compaction trends.

The database

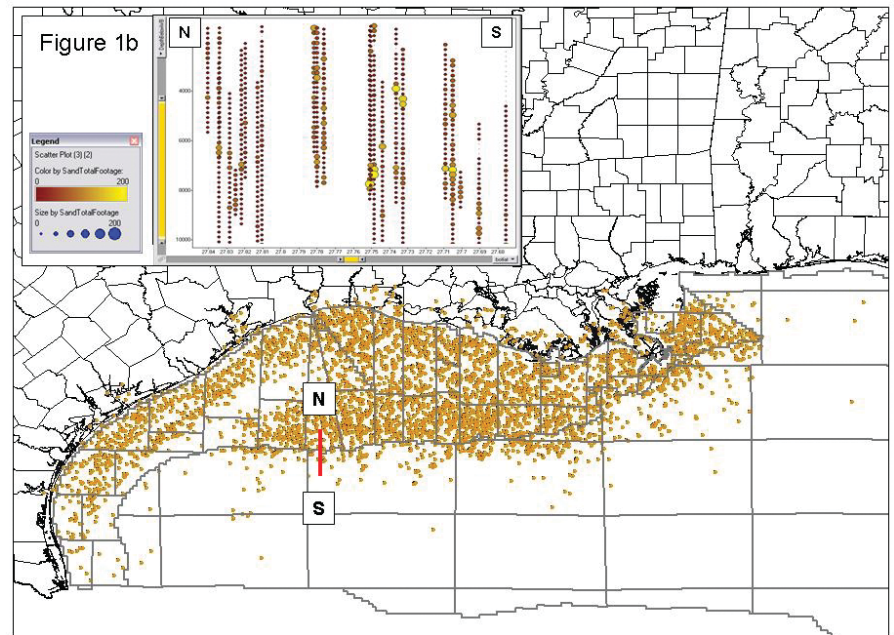GDC's geophysical rock property

Figure 1a

Figure 1. Figure 1a shows the extent of the GDC Geophysical Rock Properties Database for the northern Gulf of Mexico overlain on a boundary map. Each circle represents a well. The North/South cross-section line depicts the line shown in Figure 1b. Figure 1b depicts total sand percentage on a log cross-section. The size and color of the circles are a function of the total sand footage per 200-ft (61-m) sample interval. (All images courtesy of Geophysical Development Corp.)

database contains 7,000 wells that span across the northern Gulf of Mexico (Figure 1). The data from each well is binned in 200-ft (61-m) sample intervals. GDC calculates geophysical rock properties for each sample interval where there are sufficient well logs, including sand and shale density, compressional or P-wave velocity, sand and shale footage, mud weight, and temperature. A subset of the GDC database has been viewed in Spotfire for this project.

One of the key values of viewing the GDC database in Spotfire is the ease of plotting multiple parameters against each other to quickly evaluate the data. Spotfire's multi-variable plotting allows

for visualization in 2-D and 3-D and at the same time has the ability to independently display variables based on symbol shape, size, color and rotation, making for essentially an eight-variable plotting tool. The ease at which multiple plots can be made allows for the quick identification and evaluation of trends within the GDC dataset. Data points can easily be highlighted on any one plot and will automatically be highlighted on all previously generated plots. This ability can be used to simply and quickly QC the data and identify trends or outlier points. These points can then be omitted from the database.

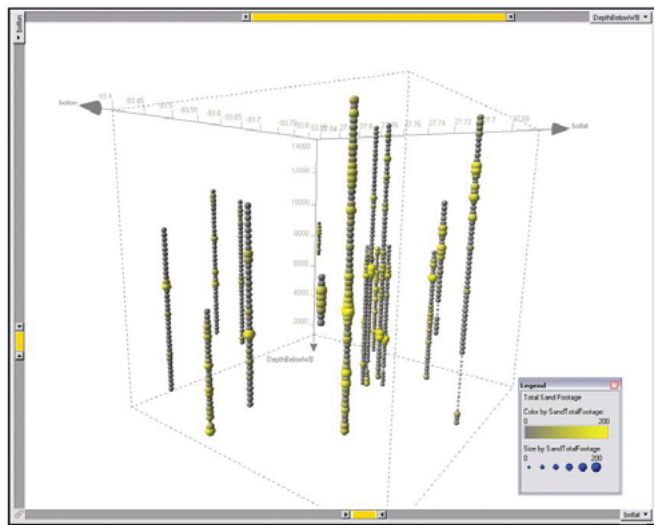Spotfire can also display spatially

Figure 2. This figure shows a 3-D plot of sand footage vs. latitude, longitude and depth below water bottom. Sand thickness is plotted as a function of color and size. The bigger and yellower the interval is, the more sand is present in the 200-ft (61-m) sample interval.
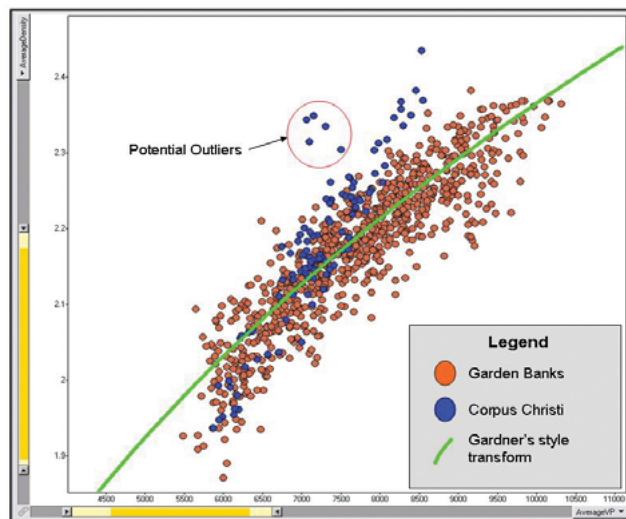

Figure 3. Cross plots of Average Density vs. Average P-wave velocity were generated in Garden Banks and Corpus Christi. The Garden Banks data (orange points) match a Gardner-style transform, while the Corpus Christi data (blue points) have a different trend.

rectified maps as backgrounds on the plots (Figure 1). GDC has provided Anadarko with several useful maps, such as a bathymetry map and an initial onset of geopressure map, both generated from the database. Anadarko also uses MapConnect (an added feature of Spotfire) to import background geo-referenced maps to be used as a backdrop for trend analysis.

Spatial understanding

In addition to 2-D and 3-D plots, Spotfire DecisionSite enables spatial review of the GDC data as maps, cross-sections or volumes. Plotting attributes of interest against depth, limiting the latitude or longitude range on the slider-bars, and moving the small interval through the volume generates a powerful interactive view of the data as a log cross-section (Figure 1 inset). Plotting an attribute such as sand footage using a color, size or shape against latitude, longitude and depth creates 3-D displays which provide a powerful understanding of three-dimensional spatial relationships (Figure 2).

Determining porosity

Seismically predicting porosity with depth and relating it to permeability is a key challenge facing many geoscientists. One method to understanding porosity with depth is to understand density with depth. Typically, the deeper the sediments, the more dense and unfortunately less porous they become. Seismic data, which is a function of velocity, can be converted into a density volume using a derivation of the Gardner's equation, which relates average density to average P-wave velocity. The density volume can then be converted into a porosity volume in a similar manner as density wireline log data are converted into porosity units. Both density and velocity are commonly recorded in offshore wells as wireline data, which has been incorporated into GDC's database. The end product is a seismic-based porosity volume which becomes very useful in assessing exploration prospects.

Figure 3 depicts a cross-plot of average density versus average P-wave velocity of well data for a portion of the Garden Banks (orange dots) and Corpus Christi (blue dots) areas in the northern Gulf of Mexico. A regional Gardner's style transform (Huston, et al., 2004) is also plotted on the data as the green line. The Gardner-style transform fits the Garden Banks data but does not fit the Corpus Christi data. The rock properties in the Corpus Christi area have much lower average velocities corresponding to higher average densities (lower porosities). The lower velocities could be tied to a volcanic or carbonate component unique to the lithology of this area. Also, note the potentially questionable data (higher density, lower velocity) that does not lie on the Corpus Christi trend. This data can be highlighted and viewed on other plots to attempt to verify its validity or can be deleted before generating a Corpus Christi velocity to density transform. After recognizing this trend for the Corpus Christi Area, other areas can be examined for expressions of this same component. Different velocity-to-density transforms can easily be generated for each area to better predict density and porosity using seismic data.

The authors have found this to be a very useful tool to create spatial views of the database such as map and cross-section displays and use these displays to select wells and areas for additional study. Multiple attribute cross-plots such as Density vs. Depth, Sand Distribution vs. Depth, Temperature Gradients, Geopressure Relationships and others are exceptionally easy to create. This ongoing analysis is integral to building regional geologic models and trend analyses for use in basin modeling and validation of geophysical modeling. **E&P**