Abstract
Reservoir characterization requires a good understanding of the spatial distribution of reservoir properties. Statistical regression techniques such as multiple regression and neural networks have been used extensively for this purpose. This paper presents a novel technique to understand the limitations of the reservoir data and to quantify the uncertainty of the subsequent mapping results. We use the Pinedale Anticline dataset in Wyoming to illustrate the workflow and the usefulness of the technique. A set of structural and seismic attributes is to map the cumulative gas production, a proxy for fracture intensity, and the degree of extrapolation is numerically quantified.

Introduction
Mapping reservoir properties is a highly nonlinear problem. Statistical regression techniques such as multiple regression and neural networks have been used extensively for this purpose. These techniques require a set of calibration data (known input-output data pairs), commonly known as the “training set.” The training set is built using the densely sampled data (i.e. input drivers, e.g. seismic attributes) collocated with the isolated known data (i.e. target/output, e.g. well data). The selection of the representative input drivers is also a critical step in applying regression-based techniques because possible dependent inputs may exist that add no additional information content to the solution, and the removal of redundant or irrelevant data would avoid over-constraining the model and also reduce the CPU time. This topic however is beyond the scope of this paper.

In this paper, we will introduce a novel technique to perform a “reality check” of the training set that clearly shows the potential limitations of the reservoir data. The concept of attribute extrapolation will be discussed in the next section, followed by the introduction of coefficient of extrapolation. The workflow will be demonstrated with a dataset in the Pinedale Anticline in Wyoming.

Attribute Extrapolation
The problem of spatial extrapolation has been well addressed through the practice of geostatistics with the concept of “kriging variance.” Many regression practitioners however have overlooked the problem of “attribute extrapolation.” This issue relates to the extent of extrapolation beyond the coverage of the training data for reservoir mapping purposes. It helps us to understand the limitation of the available data in relation to its predictive power at the unsampled regions (Wong and Boerner 2003).
The study presented by Gauthier et al. (2000) was one of the few published works that revealed the actual problem. The authors examined the extent of extrapolation based on the coverage of each of the drivers in the training data set and unsampled regions. Fig. 1 illustrates the limitation of the training data. As shown, the span of the input data does not cover the span of the input in the unsampled regions, and therefore extrapolation occurs in any statistical regression techniques. This paper attempts to quantify the extent of extrapolation and provide an indication of the “danger zones.”

![Fig. 1 Indication of extrapolation of an input driver (modified after Gauthier et al. 2000).](image)

**Coefficient of Extrapolation**

In this paper we introduce “coefficient of extrapolation.” The idea was first presented in Wong and Boerner (2003). We first examined the individual driver and calculated the relative distance of departure from the extreme value (i.e. minimum or maximum) in the training data: the farther away from the extreme value, the larger the coefficient. We then weighed the relative distances of all the input data types and an average value was calculated for each cell. For the data within the coverage of the training data, the coefficient was set to zero (i.e. no extrapolation). The remaining problem with this method is that we looked at each input driver only independently. Fig. 2 displays the problem of extrapolation even though the two data points have no extrapolation for either individual driver. When the data is posted on the two-dimensional input space, extrapolation clearly reveals. Therefore it is necessary to quantify the degree of extrapolation in the multidimensional input space.

![Fig. 2 Extrapolation in two input dimensions.](image)
In essence, the new technique looks for the closest training data point for each unsampled data point. The minimum (normalized) Euclidean distance is used to calculate the coefficient of extrapolation. We rank all the input drivers according to the “OPRAC” coefficients (Wong and Boerner, 2004) and weigh the significance of each driver in the distance calculation. This procedure solves the problem presented in Fig. 2. Note that if the unsampled data point were a training data point, the distance to itself would be 0.0 and therefore the calculated coefficient of extrapolation would also be 0.0. If the coefficient were 1.0, the unsampled data point would be the farthest from all the available training data. Since the coefficient is available at each cell, the results can be shown as a map for any reservoir property of interest.

Example: Pinedale Anticline

The Pinedale Anticline is located in Sublette County, southwestern Wyoming. A 3D seismic survey was acquired over the anticline and processed using Amplitude Versus Angle and Azimuth (AVAZ) analysis that estimates the local intensity and orientation of fracturing (Gray et al. 2003). The seismic attributes offer the advantage of dense lateral coverage and can be used to correlate with well data. In this paper we selected four significant drivers (crack density, coherency of the zero-incident shear wave, azimuth of anisotropic gradient, and top structure depth) to map the four-month cumulative gas production at 19 wells, a proxy for fracture intensity (Boerner et al. 2003). The study was done in 2D with a total of 28,640 cells in the map area. The well locations are shown in Fig. 3 together with a top structure map.

---

Fig. 3. Well locations and top structure map.

Fig. 4. Coefficient of extrapolation map.

Fig. 5. S-wave coherency vs crack density.

Fig. 6. Coefficient of extrapolation vs S-wave coherency
Fig. 4 shows the resulting map of the coefficient of extrapolation. The red regions represent regions with significant extrapolation, while the blue regions have low extrapolation. Fig. 5 shows a cross-plot of the top two drivers, S-wave coherency and crack density, for both training and unsampled data. As shown, the training data covers only a small part of the overall data coverage. Fig. 6 shows a plot of the corresponding coefficient of extrapolation with the S-wave coherency. The S-wave coherency with data greater than 2 experienced significant extrapolation. Note that the coefficients were zero at the well locations.

Fig. 4 provides a good indication of the inherent predictive power of the data at hand. Note particularly that this uncertainty map is constructed prior to the selection of the statistical regression techniques, and therefore it is not model-based uncertainty estimates. It can be used to cross-check the results derived from any mapping techniques. One must become cautious if a mapping technique gives confident result (e.g. low variance) in a particular cell that has a high coefficient of extrapolation.

Conclusions

In this paper, we introduced a novel technique to understand the limitations of the reservoir data and to quantify the uncertainty of the subsequent mapping results. A new attribute for reality check, “coefficient of extrapolation,” was proposed and used to understand the predictive power of the reservoir data at hand. The workflow was demonstrated with the data from the Pinedale Anticline in Wyoming. The study shows the usefulness of the uncertainty map and demonstrates how it can be applied in practice.

References


Wong P. and Boerner S. 2003. Fracture intensity modeling using soft computing approaches. SEG Summer Research Workshop on “Quantifying Uncertainty in Reservoir Properties Prediction,” Galveston, TX, abstract only.