P 293

Total organic carbon prediction for shale gas exploration using statistical clustering, multiple regression analysis

Abhishek Kumar and Dr. Satish K. Sinha*

Summary

In shale gas exploration, prospectivity of a shale formation is analyzed based on several geochemical parameters of source rock like – Total Organic Carbon (TOC), Thermal maturity (from Vitrinite Reflectance and Rock Eval Pyrolysis) etc. Such data have not been acquired uniformly in the past in conventional exploration for obvious reasons. Now, in case of shale gas exploration it has become important estimate these geochemical parameters from available information. Here we propose to use a geostatistical approach in which alternating conditional expectation (ACE) algorithm is used for electrofacies analysis in conjunction with the known TOC values at various well locations in a shale layer.

Keywords: *Reservoir characterization, Shale Gas, Geostatistics, ACE Algorithm*

Introduction

TOC content present in reservoir rocks is one of the important geochemical parameters, which could be used for evaluating residual production potential as well as the geochemical characterization of hydrocarbon bearing units. In general, organic rich shales are characterized by higher porosity, higher sonic transit time, lower density, higher gamma ray, and higher resistivity compared to other shales.

A three-step approach has been proposed in predicting TOC content using well log data. Initially, the well log data are classified into a set of electrofacies. This classification does not require any further subdivision of the dataset but follows naturally based on the unique characteristics of well log measurements reflecting mineral and lithofacies responses within the logged intervals. Cluster analysis of petrophysical data are done to put them into different electrofacies. Current study suggests an improved and optimal model for TOC estimation by integration of Alternating Conditional Expectation (ACE) algorithm using well log data. This optimal model is then used to predict TOC values at unsampled locations and a three dimensional model of TOC is generated using geostatistical methods.

Theory

Electrofacies can be described as a rock type or sediment exhibiting similar set of well log responses. We classify the well log data into electrofacies based on the similarity of their response. This electrofacies classification does not require any artificial subdivision of the data population but follows naturally based on the unique data values reflecting minerals and lithofacies within the interval. A combination of principal component analysis (PCA), model-based cluster analysis and discriminant analysis is used to describe and identify the electrofacies types. In PCA, principal components are ranked according to their contribution to the total variance of the dataset and their contribution can be determined by descending eigenvalues. Initial three to five principal components which defines around 90% variation will be used to do the cluster analysis for electrofacies identification. Cluster analysis is used to classify a dataset into groups that are internally homogeneous and externally isolated on the basis of a measure of similarity or dissimilarity between groups. Clustering allows separating volumes of discrete classes that are computed directly, using all significant principal components.

Further ACE is used with each identified electrofacies to predict a geochemical parameter. The method of ACE,

Rajiv Gandhi institute of Petroleum Technology (RGIPT), Rae Bareli, India,
E-mail: abhikumar@rgipt.ac.in

originally proposed by Breiman and Friedman (1985), is based on non-parametric regression. In non-parametric regression apriori knowledge of the functional relationship between a dependent variable Y and multiple independent variables $X_1, \ldots X_p$ is not required. The non-parametric transformation technique generate regression relations in a flexible data using conditional expectations or scatter plot smoothers and in doing so, let the data itself suggest functional forms or detect inherent nonlinearities. Optimum non-parametric transformations produce maximum correlation in the transformed space (Wang and Murphy, 2004; Xue et al., 1997).

Further, we use the correlation to predict the TOC values and the concept of geostatistics has been used to prepare a model of the TOC distribution. Geostatistics allows us to account for a spatial relationship of a variable. In many natural phenomena, measurements of a variable reveal that the variable values measured close to each other are similar. As the distance between the measured values increases, the similarity between the two measurements decreases. Geostatistics takes advantage of this similarity and captures the spatial relationship through certain correlation functions. Using these functions, different methods in geostatistics like Kriging, conditional simulation can be used to estimate values of a variable at unsampled locations (Kelkar and Perez, 2002).

Electrofacies Analysis

The data presented in this analysis have been gathered from 4 wells. With the consideration of the data quality and field-wide availability, a suite of well logs is selected for the analysis. In this field, we have seven well logs namely the spontaneous potential (SP), gamma ray (GR), deep and shallow Laterolog (LLD, LLS), sonic (DT), neutron porosity (NPHI), and density (RHOB) logs. These seven logs are chosen for characterizing the electrofacies groups.

Initially PCA is used to summarize the data effectively and reduce the dimensionality of the data without significant loss of information. PCA is applied to obtain the principal components PCj (j=1… 7) from the well log data after normalization. Figure 1 shows the scree plot, a bar plot of the variance contribution of the principal components which often provides a convenient visual method of identifying the important components. Only 4 principal components explain around 90% variation of the whole data set. Now these four principal components will be used to do the cluster analysis for electrofacies identification.

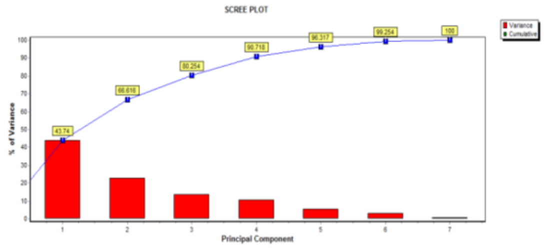

Figure 1: Relationship between the principal components and the variance they explain

Now model based cluster analysis is used to define eight distinct groups based on the unique characteristics of the well log measurements. We have used first four principal components from PCA for cluster analysis which defines more than 90% variance of the dataset. In Figure 2, each cluster can be treated as an electrofacies that reflects the hydrologic, lithological, and digenetic characteristics.

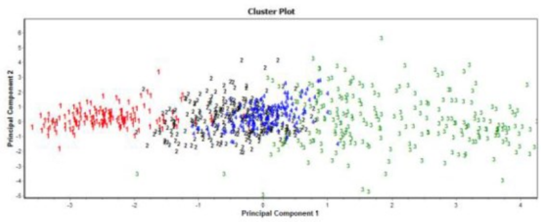

Figure 2: Cluster analysis showing four distinct clusters.

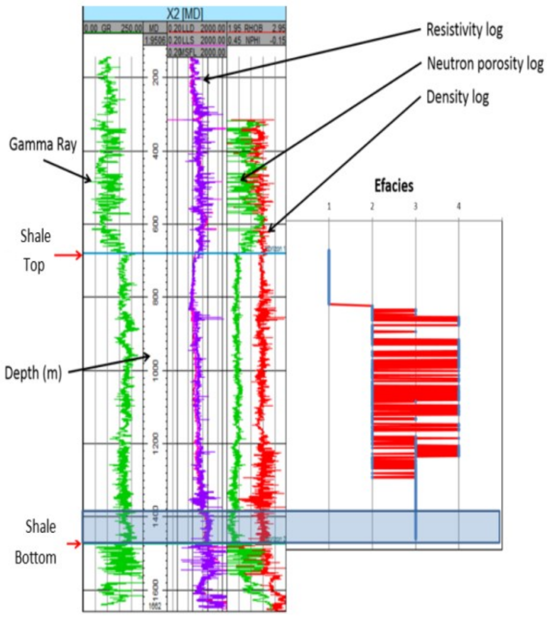

Figure 3: Comparison of the well logs with the identified facies from E-FACIES analysis for the well X2.

Geochemical characterization using ACE

Seven log parameters for Shale zone have been taken into consideration to work out the TOC (dependent variable). We subjected these parameters to non-parametric regression by invoking ACE. The optimal transformations between TOC and different combination of select seven independent variables were obtained through ACE. The best possible results are obtained by using three independent variables namely DT, LnRT (logarithmic LLD) and NPHI. The sum transformation of these three independent variables has been constructed. Now TOC can be estimated from these variables for a point where TOC is unknown (uncored intervals) using the regression analysis shown in Figure 4 derived by ACE. We have tried different sets of combination to find out this correlation from X4 well, which is the best possible match for the purpose of deriving TOC data from well logs at uncored wells or unsampled locations.

For the purpose of validation, we have compared the TOC values derived from the data of X3 (cored well) well using above correlation and with the TOC values available from the core study as shown in Figure 5.

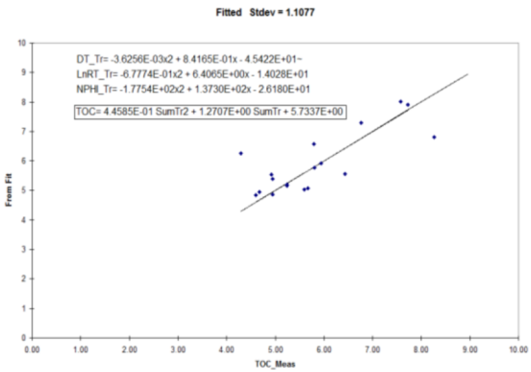

Figure 4: Fitted standard deviation with the equation used in the correlation.

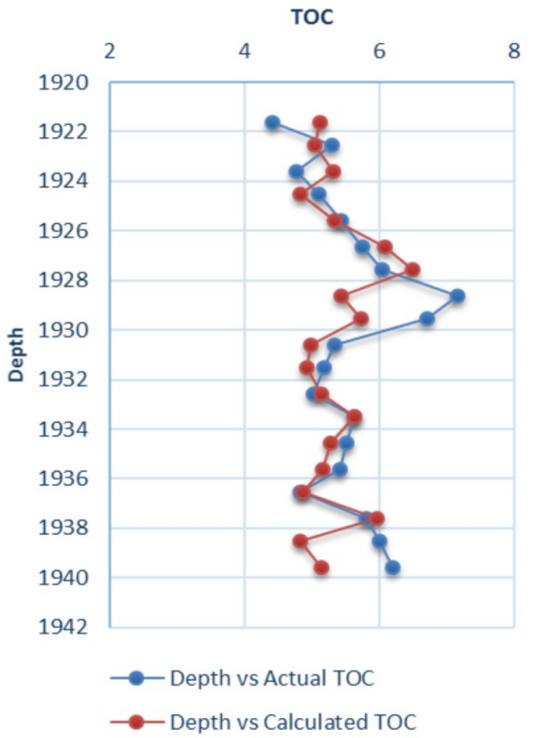

Figure 5: Comparison of core derived TOC data with the calculated TOC for shale interval in the well X3.

Three dimensional model generation for TOC

A 3-D reservoir model was generated by conditional simulation at unsampled locations using the TOC values generated from the well log data using above correlation. The resulting reservoir TOC model provided insights into assessing the geochemical characterization of pay zone.

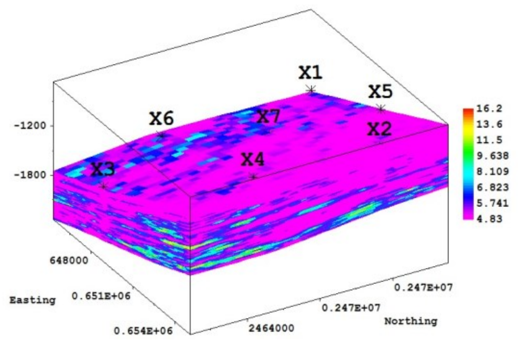

Figure 6: Three dimensional (3D) grid view for the distribution of total organic carbon (TOC).

Conclusions

The present work has been carried out with the objective to analyze key geochemical, and petrophysical parameters which should be considered to the select shale formations for shale gas exploration and developing a model for estimation of these parameters by applying non-parametric methodology. A three step approach was used in predicting TOC values from petrophysical data.

Present study has been on pre-classifying the well log responses into several distinct clusters, or electrofacies. Cluster Analyses were applied to determining the electrofacies from well log data. Then non-parametric regression methodology was used for estimating the geoscientific parameters. The numerical model leads to a mathematical construct which is able to provide values of geochemical parameter TOC in terms of the employed variables (predictors). The values are reasonable as well as acceptable when examined against the available TOC values from the cored interval.

In this work we have dealt with the geostatistical methodology to develop a 2D and 3D model for shale facies. In order to carry out facies modeling the value of TOC calculated from logs has been used for selected wells. Variogram model to obtain the spatial continuity of modeled data has been developed. Using the predicted correlation, a three dimensional TOC model has been developed.

The integrated study and analysis of the 3D and 2D models of petrophysical characteristics, geophysical and geochemical data will facilitate the identification of sweet spot for shale gas exploration in the field under consideration. The drilling of wells for shale gas exploration require the well profile to be placed horizontally within identified shale section to be hydro fractured. The completion of well and its proper placement in required shale section can be reasonably taken care of by referring to the 3D and 2D models.

Acknowledgements

This work was carried out as a part of master's thesis of the first author. We would like to thank ONGC for providing the dataset for this study. We are grateful to RGIPT for supporting this research.

References

Breiman, L., and Friedman, J. H., 1985, Estimating Optimal Transformations for Multiple Regression and Correlation; Journal of the American Statistical Association, 80(391), 580-598.

Kelkar, M., & Perez, G., 2002, Applied Geostatistics for Reservoir Characterization; Society of Petroleum Engineers Inc.

Wang, D., & Murphy, M., 2004, Estimating Optimal Transformations for Multiple Regression Using the ACE Algorithm; Journal of Data Science, 2, 329-346.

Xue, G., Datta-Gupta, A., Valko, P., & Blasingame, T., 1997, Optimal Transformations for Multiple Regression: Application to Permeability Estimation from Well Logs; SPE Formation Evaluation, 12(2), 85-94.