Journal of Petroleum Science and Engineering 170 (2018) 664–674

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

Automatic lithology prediction from well logging using kernel density estimation

A.N. Corina*, S. Hovda

Norwegian University of Science and Technology (NTNU), Trondheim, Norway

ARTICLE INFO

Keywords:
Real-time drilling data
Gamma ray
Statistical classification
Kernel density estimation
Non-parametric data
Lithology prediction

ABSTRACT

Technologies of real-time data measurement during drilling operation have kept the attention of petroleum industries in the past years, especially with the benefit of real-time formation evaluation through logging-while-drilling technology. It is expected that most of the logging data will be recorded in real-time operation. Hence, application of automated lithology prediction tool will be essential.

An automatic method to predict lithology from borehole geophysical data was developed. It was solved as a multivariate classification problem with multidimensional explanatory variables. The learning algorithm combines kernel density estimates and a classification rule that is based on these estimates. The goal of this work is to test the method on a univariate variable and validate the prediction accuracy by calculating the misclassification rates. In addition, the results will be established as a baseline for application in practice and future developments for multivariate variables analysis.

Gamma-ray from wireline logging is selected as the variable to describe two lithology groups of shale and not-shale. Data from six wells in the Norwegian Continental Shelf were extracted and examined with aids of explorative data analysis and hypothesis testing, and then divided into a training- and test data set. The selected algorithm processed the training data into models, and later each element of test data was assigned to the models to get the prediction. The results were validated with cutting data, and it was proved that the models predicted the lithology effectively with misclassification rates less than 15% at its lowest and average of ± 31%. Moreover, the results confirmed that the method has a promising prospect as lithology prediction tool, especially in real-time operation, because the non-parametric approach allows real-time modeling with fewer data assumptions required.

1. Introduction

The process of lithology identification is traditionally executed using data from cutting visualization, core inspection, or wireline logging. And today, many new technologies are advancing and replacing the manual process into a more automated process, such as high-speed telemetry. This development means that more types of borehole geophysical data are measured in the real-time operation, and hence lithology identification methods are expected to be more straightforward and precise than the traditional methods. This motivates the development of an automated method of lithology prediction.

The early technique of lithology interpretation was accomplished using qualitative approach through identification of log separations or unique trends between several well log curves visually without the requirement of calculations. In practice, this technique provides quick evaluations, especially over a depth of interval which is consistent. However, the application becomes demanding for complex lithologies identification that requires large dataset analysis and depends on the geological history of the area (Ellis and Singer, 2007).

The advanced progress of modern computers has stimulated the development of quantitative methods of lithology identification with improved speed and accuracy. There are wide variations of mathematical techniques adapted as lithology identification tool, such as clustering (Wolf and Pelissier-Combescure, 1982; Ye and Rabiller, 2000), fuzzy logic (Cuddy et al., 1997; Saggaf and Nebrija, 2003), and neural networks (Benaouda et al., 1999; Maiti, 2007). One of the early studies that implemets statistical probability method with combination of clustering and classification technique for lithofacies determination was accomplished by Delfiner et al. (1987). Since then, many other studies were carried out in similar manners, including studies by Busch et al. (1987) and Coudert et al. (1994). Those studies came in conclusion that the classification technique based on probability density was promising

* Corresponding author.
E-mail addresses: anisa.corina@ntnu.no (A.N. Corina), sigve.hovda@ntnu.no (S. Hovda).

https://doi.org/10.1016/j.petrol.2018.06.012
Received 8 November 2017; Received in revised form 10 May 2018; Accepted 4 June 2018
0920-4105/

for lithology prediction and the statistical methods were suitable for handling large databases. However, the assumption of normal (Gaussian) distribution for the density probability function was believed to be strict for modeling non-parametric data.

Modeling the non-parametric data that are infinite-dimensional is best approached using non-parametric statistic technique. The application is convenient for dataset that grows in size – i.e. a dataset whose final structure of data distribution is yet unknown–, such as model from real-time dataset. In statistic probability, the estimation of probability density function of non-parametric data is usually accomplished using kernel density estimator. It is also an excellent tool for estimating univariate, bivariate, or trivariate data, even when the number of data points is relatively low (Silverman, 1986). Kernel density estimator has also been applied to solve geophysical and geologicals problem in the past (Mwenifumbo, 1993; Mwenifumbo et al., 2004). Mwenifumbo (1993) specifically applied the estimator on well logging data and proved that the results of probability density function were precise in showing the major features of each lithofacies.

Until recently, the automated lithology predictions that based on statistical probability density did not take account of non-parametric modeling, meaning that the assumptions were not practicable on real-time dataset. Therefore, in this study we attempted to develop a lithology prediction method using a classification technique based on probability density function of explanatory variables, which was estimated using kernel density estimator. The selected classification technique implemented a classification rule, or classifier, to generate the final classification models. Two types of classifiers were presented in this study, one of which implemented prior probability value.

To give a brief overview of the proposed method, we presented a set of two-dimensional data with 30 points (black, red, and blue points) as a contour plot of the probability density functions, estimated by kernel density estimator in Fig. 1a. Fig. 1b describes a trinary classification rule, which neglected the prior probability, based on two-dimensional data, dividing the data into three different classes marked with Region 1, 2, and 3. If the classification rule was modified, by taking prior probability into account, some regions expanded or shrunk depends on the probability value of the particular region (see Fig. 1c). Notice that there are several points now classified into the Region 1 after the classification rule was modified.

One of the principal aims in this study is to test the proposed learning algorithm by using a univariate data, which is gamma ray log, and acquire the accuracy given by the models from classifying new observations to lithology groups of shale and not-shale. Our methods to select the data and how to employ them into the learning algorithm are described in detail prior the test. Another of our aims is to present the application of the proposed method in practice as a baseline for petroleum engineers to implement, especially in real-time operation.

2. Dataset description

The data used in this study was from six wells located in Norwegian Continental Shelf. The wells are situated at the eastern part of the South Viking Graben with three wells from Block 15, situated at Gina Krog field within Ve sub-basin, and three wells from Block 16, situated at Ivar Aasen field within the Gudrun Terrace (Fig. 2). The configuration of the South Viking Graben is mainly due to the Callovian-Ryazinian rift event. The South Viking Graben has a steep bounding with a small terrace to the east (The Gudrun Terrace). The Gudrun Terrace is dominated with shallow marine deposition on the basin flanks, with terrace topography. The fault bounding the graben to the west was active during the regressive phase of Lower Oxfordian, while sediment gravity flowed to the grabenal area. The Ve sub-basin is located at the grabenal area with a thick section of Cretaceous (Steel et al., 1995).

The available data included gamma-ray logs, well schematic, geological descriptions, and mud logging. In this study, we chose gamma ray log as the explanatory variable to distinguish shale and not-shale

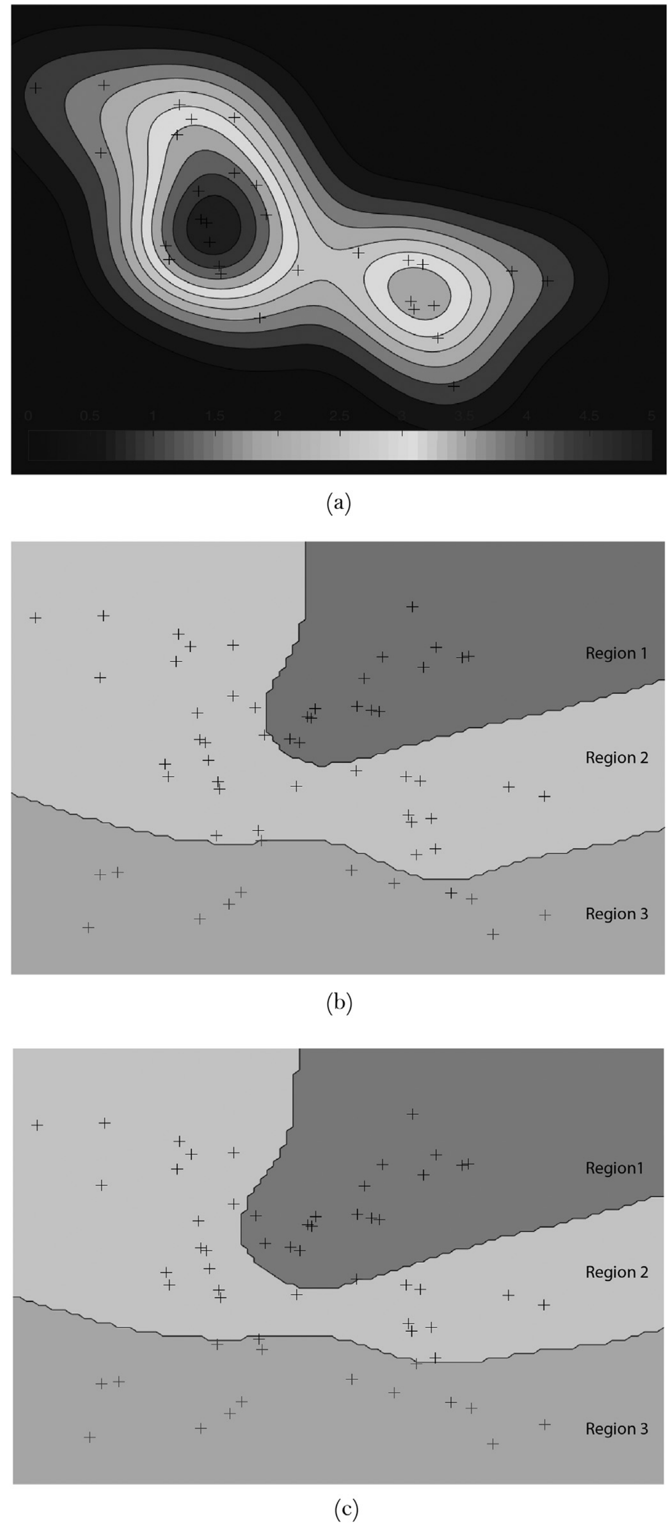

Fig. 1. The 2-dimensional multivariate analysis: (a) probability density function from kernel density estimation, (b) group region based on classification rule without prior probability, and (c) group region based on classification rule with prior probability.

lithology because it is a reliable shale detector and the tool is commonly run in combination with high pulse telemetry. Gamma ray tool measures the composition of the natural-occurring isotopes contained in the rocks, such as potassium, uranium, and thorium (Ellis and Singer, 2007). Due to high content of radioactive mineral in shale, the tool is effective to identify shale (Schlumberger Educational Services, 1989).