International Journal of Coal Geology 181 (2017) 11–22

Contents lists available at ScienceDirect

International Journal of Coal Geology

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

Coal ash content estimation using fuzzy curves and ensemble neural networks for well log analysis

Irdian Siregar[1], Yufu Niu[1], Peyman Mostaghimi, Ryan T. Armstrong*

School of Petroleum Engineering, The University of New South Wales, NSW 2052, Australia

ARTICLE INFO

Keywords:
Ash content
Fuzzy curves
Neural network
Well log analysis

ABSTRACT

Many important variables for reservoir development and production cannot be derived analytically from continuous well logs. Empirical regression and classification techniques have been widely used to predict these variables from well logs. This approach generally uses data from core analysis and well logs to train a model, which then can be used to estimate a variable where core analysis data are not available. In formation evaluation, the amount of training data is limited or costly to acquire. This may result in regression models having limited predictability. This paper addresses the problem of sparse data by using fuzzy logic and ensemble neural networks to estimate coal ash content from a collection of sparse data. Ash content is a significant parameter to evaluate coal quality and it is usually measured from proximate analysis in the laboratory. Ash content is estimated based on the components of six major oxides (Al_2O_3, SiO_2, K_2O, CaO, Fe_2O_3 and TiO_2) by using an X-ray fluorescence technique. We first use fuzzy curve analysis to rank the relationships between well log and ash content data to determine input parameters for estimating ash content. The data sets were then sampled with a bootstrap-aggregating algorithm to create a number of training sets for building ensemble neural networks. The neural networks in the ensembles were trained individually and the outputs were combined to estimate ash content. In total 20 core samples were collected from a New South Wales (Australia) coal bed methane well in the Gloucester Basin and analyzed for ash content. The well was analyzed using density, photoelectric, gamma ray, neutron, acoustic, resistivity, spontaneous potential, and resistivity imaging logging techniques. The tested algorithm produces repeatable ash content prediction (standard deviation of repeated predictions is 0.43%) and effectively reduces the prediction variance and bias compared to the single neural network with early stopping algorithm. The workflow is data-driven and could be used to estimate other complex variables that are required when evaluating coal beds.

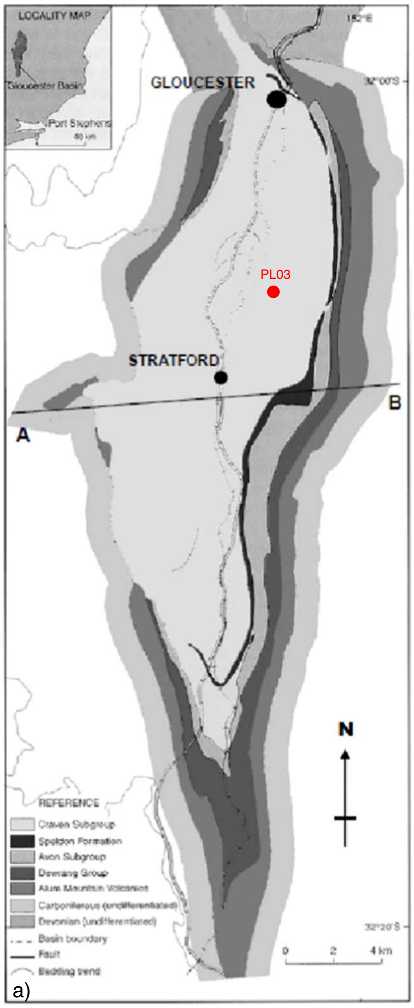
1. Introduction

The petrophysical properties of rocks are required for the evaluation of hydrocarbon reservoirs and are often inferred from well logging data. Well logging provides continuous evaluation of formation properties that can be interpreted as rock properties, mineralogy and sedimentary. Basic logging data can be interpreted to determine lithology and facies of a reservoir (Serra, 1984). Furthermore, spectral natural gamma ray adds details to mineralogy evaluations, while imaging tool data adds details to sedimentary studies. Interpretations of well logging data are often based on empirical relations of well log data to core sample analysis in laboratory. An interpretation model is trained/calibrated with the core sample analysis data, then used to evaluate intervals where core samples are not available. Calibrating reservoir properties such as permeability and porosity to well logging data by using statistical methods such as fuzzy logic and neural networks have been used in traditional and unconventional reservoir exploration for more than thirty years and the quality and range of measurement capabilities and methodology are improving (Lin and Cunningham, 1995; Wong et al., 1998; Helle et al., 2001; Saggaf and Nebrija, 2003; Ilkhchi et al., 2006; Abdulraheem et al., 2007; Nashawi and Malallah, 2010; Zerrouki et al., 2014; Ghosh et al., 2016).

Coalbed methane (CBM) reservoirs are complex and classified as an unconventional resource of energy (Mostaghimi et al., 2017). Current log evaluations of coal ash content are based on density log data by regression approaches to determine an empirical relationship between well log and lab measurements (Agostini, 1977; Daniels et al., 1983; Mullen, 1989; Fu et al., 2009). Agostini (1977) built a linear relationship between ash content from proximate analysis and high resolution density logs. However, it only had ± 39% accuracy for predicted ash

* Corresponding author.
E-mail address: ryan.armstrong@unsw.edu.au (R.T. Armstrong).
[1] Both authors contributed equally to this publication.

http://dx.doi.org/10.1016/j.coal.2017.08.003
Received 8 June 2017; Received in revised form 4 August 2017; Accepted 6 August 2017
Available online 07 August 2017
0166-5162/



Group	Subgroup	Formation	Coal Seams	Top-Base(m)
Gloucester Coal	Craven Subgroup	Leloma or Woods Road Formation	Bindaboo	134.97-180.63
			Deards	191.14-204.46
		Jilleon or Bucketts Way Formation	Cloverdale	229.32-259.76
			Roseville	291.40-329.34
			Fairbairns Lane	344.23-454.25
		Wenham Formation	Bowens Road	494.03-497.53
	Avon Subgroup	Dog Trap Creek Formation	Glenview	553.91-618.61
		Waukivory Creek Formation	Avon	659.86-669.56
			Triple	679.88-699.79
			Rombo	710.53-718.11
			Valley View	776.61-795.68
			Parkers Road	902.22-961.50

b)

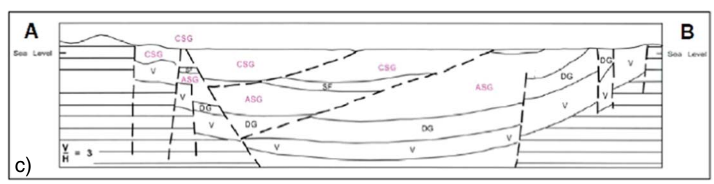


Fig. 1. Location and surface geology of the gloucester basin.
Modified after AGL geology of the gloucester basin.

content at 90% confidence level. Mullen (1989) developed another linear relationship between ash content from core analysis and bulk density log with variable results. Because of the variability of coals, the relationships are only valid regionally. Finding empirical relationships between well log and coal properties requires an adequate amount of data and measurement ranges. However, data sets available for evaluation are often small and sparse, due to the high cost and other technical factors involved in obtaining core samples. In addition, log evaluations of thin beds are challenging, especially when the properties of the adjacent beds are in high contrast, such as in coal and silt sequences that are very common in CBM reservoirs. For example, in the CBM reservoir evaluated in this work, single coal seams are usually interlayered with siltstone and/or sandstone. Bed boundary effects on log readings are inevitable and thus add to the complexity to the analysis.

Ash content or proximate analysis is an important measure for coal since it determines the quality of the coal. The relationship between ash and well logging data is often built based on proximate analysis in the laboratory (Agostini, 1977; Daniels et al., 1983; Mullen, 1989; Ghosh et al., 2016). These papers built a relationship between well logging and coal ash, which is measured by X-ray fluorescence (XRF) technique. Kiss (1966) was the first to determine the major inorganic elements of coal by XRF. Afterwards, Kuhn et al. (1975) demonstrated the trace elements of coal. He found that some variations occur at higher trace element concentrations especially in the more roughly ground coals. Evans et al. (1990) examined major elements by WD-XRF and trace elements by ED-XRF in coal ash. In particular they examined Chlorine and phosphorus in coal core plugs. Kimura (1998) calculated ash content from the sum of major element and also found the relationship between the inorganic elements and related minerals in coals by combining XRF and XRD. Kelloway et al. (2014) examined coal element characteristics and calculated the major element oxides related to the total ash and relative density by Itrax XRF techniques that can scan entire cores. With the aid of XRF, coal ash and relative density can be estimated and can improve the accuracy of well log analysis in CBM reservoirs.

Artificial neural network (ANN) approaches have been extensively used to interpret petrophysical properties in hydrocarbon reservoirs (Huang et al., 1996; Wong et al., 1998; Weiss et al., 2001). Compared with traditional empirical and multiple regression equations, ANN is a non-linear method that does not require *a priori* selection of the mathematical model (Huang et al., 1996). It is particularly useful when the analytical model for well log evaluation has not been defined, due to the complexity of the rock. Applying ANN to small data sets tends to produce output with high variance, which depends heavily on how partitioning of the available data into training, validation and testing subsets is conducted (Bui et al., 2008). To improve estimation from small data sets, instead of using a single ANN model, several models with randomly sampled input, random weight and biases are trained individually and then combined into an ensemble neural network. Each network model produces generalisation errors on different subsets of input data; however, the collective decisions by the ensemble tend to