Computers & Geosciences 101 (2017) 38–47

Contents lists available at ScienceDirect

Computers & Geosciences

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

Case study

# Quantitative thickness prediction of tectonically deformed coal using Extreme Learning Machine and Principal Component Analysis: a case study

Xin Wang[a], Yan Li[b], Tongjun Chen[c,e,*], Qiuyan Yan[a], Li Ma[d]

[a] School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, Jiangsu, China
[b] School of Agricultural, Computational and Environmental Sciences, University of Southern Queensland, Toowoomba, Queenland, Australia
[c] School of Resource and Earth Science, China University of Mining and Technology, Xuzhou, Jiangsu, China
[d] Key Laboratory of Coal Resources Exploration and Comprehensive Utilization, Ministry of Land and Resources, Xian, Shanxi, China
[e] Key Laboratory of CBM Resource & Reservoir Formation Process, Ministry of Education, Xuzhou, Jiangsu, China

ARTICLE INFO

*Keywords:*
Thickness prediction
Tectonically deformed coal
Extreme learning machine
Seismic attribute
Principal component analysis
Cross validation

ABSTRACT

The thickness of tectonically deformed coal (TDC) has positive correlation associations with gas outbursts. In order to predict the TDC thickness of coal beds, we propose a new quantitative predicting method using an extreme learning machine (ELM) algorithm, a principal component analysis (PCA) algorithm, and seismic attributes. At first, we build an ELM prediction model using the PCA attributes of a synthetic seismic section. The results suggest that the ELM model can produce a reliable and accurate prediction of the TDC thickness for synthetic data, preferring Sigmoid activation function and 20 hidden nodes. Then, we analyze the applicability of the ELM model on the thickness prediction of the TDC with real application data. Through the cross validation of near-well traces, the results suggest that the ELM model can produce a reliable and accurate prediction of the TDC. After that, we use 250 near-well traces from 10 wells to build an ELM predicting model and use the model to forecast the TDC thickness of the No. 15 coal in the study area using the PCA attributes as the inputs. Comparing the predicted results, it is noted that the trained ELM model with two selected PCA attributes yields better predication results than those from the other combinations of the attributes. Finally, the trained ELM model with real seismic data have a different number of hidden nodes (10) than the trained ELM model with synthetic seismic data. In summary, it is feasible to use an ELM model to predict the TDC thickness using the calculated PCA attributes as the inputs. However, the input attributes, the activation function and the number of hidden nodes in the ELM model should be selected and tested carefully based on individual application.

## 1. Introduction

Tectonically deformed coal (TDC) is a kind of coal which their composition had been physically and chemically deformed under the movement of tectonic stress in the previously geological period (Cao et al., 2003; Frodsham and Gayer, 1999). In the present research, the occurrences of gas outbursts have direct associations with the TDC. The thicker the TDC thickness, the higher the probability of gas outbursts (Cao et al., 2003; Xue et al., 2012). Mining unpredicted thick TDC areas would set miners in very high risks (Hackley and Martinez, 2007; Ju and Li, 2009; Li et al., 2003; Pan et al., 2012, 2015). If the TDC thickness can be predicted quantitatively and accurately, safe coal mining would be easier to achieve. Currently most of the research in the literature are qualitative and focus on the prediction distribution and seismic characterization of the TDC.

Extreme Learning Machine (ELM) method, proposed by Huang et al. (2006), is an improvement of single-hidden layer feed-forward neural networks (SLFNs). The learning speed of the ELM can be thousands of times faster than the traditional learning algorithms, like artificial neural networks (ANNs), while obtaining better generalization performance (Huang, 2014). In addition, the ELM has many other advantages, such as easy to implement, quick to converge to the smallest training error, small norms of weights and good generalization performance (Huang et al., 2006). Therefore, it has been widely used in regression, multiclass classification, data analysis of non-linear time series, environmental data analysis, water level forecasting of streamflow and pattern recognition (Benoît et al., 2013; Butcher et al., 2013; De Lima et al., 2016; Deo and Sahin, 2016; Leuenberger and Kanevski, 2015; Yang and Zhang, 2016).

Seismic is a main reliable method to forecast the characteristics of coal beds. The most used seismic data in coal beds characterization are seismic attributes which are mathematically or geometrically derivate values of coal beds reflection (Chopra and Marfurt, 2007; Ge et al.,

* Corresponding author at: School of Resource and Earth Science, China University of Mining and Technology, Xuzhou, Jiangsu, China.
*E-mail addresses:* wxgrin@163.com (X. Wang), Yan.Li@usq.edu.au (Y. Li), tjchen@cumt.edu.cn (T. Chen), mary248@163.com (L. Ma).

http://dx.doi.org/10.1016/j.cageo.2017.02.001
Received 26 July 2016; Received in revised form 28 January 2017; Accepted 1 February 2017
Available online 03 February 2017

2008; Kadkhodaie et al., 2009). In general, those attributes have two characteristics: (1) the associations among different seismic attributes are non-linear; (2) the number of seismic attributes is huge. Those characteristics hinder the wide implementation of feed-forward neural networks (FNNs) as all the parameters of FNNs are tuned iteratively with a very slow convergence. In real applications, the learning speed of FNNs is too slow to achieve an acceptable prediction for coal mining. Unlike the conventional FNNs, ELMs can overcome the problem. It can generate the node parameters in a hidden layer and converge in much faster speed (Huang et al., 2015). In addition, the non-linear processing ability of ELMs is strong and adaptive. So ELMs are popular for quantitatively predicting the TDC thickness in coal-bed mining zones.

In this paper, firstly we use an ELM method to train and build a prediction model using derivate seismic attributes from a synthetic seismic section. Secondly, we use this model to quantitatively predict the thickness distribution of the TDC for a Xinjing operational coal mine in western China.

## 2. Geological settings

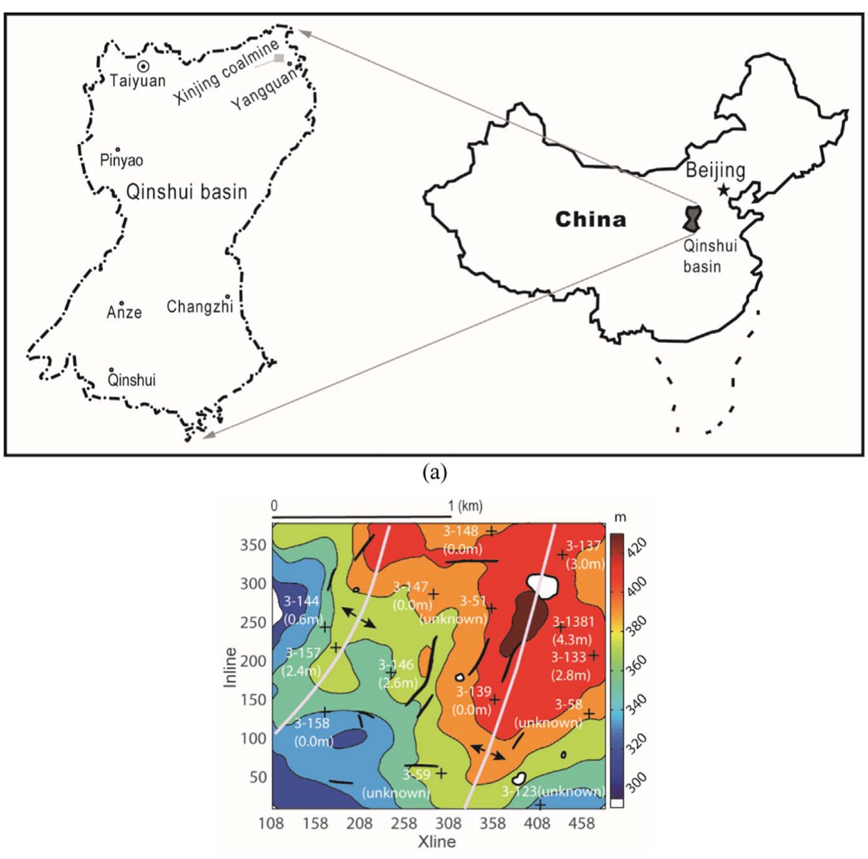
### 2.1. Locations and geological history

The Xinjing coal mine is a gas-outburst prone mine, located in the northeast corner of the Qinshui basin, western China, as shown in Fig. 1(a). In general, the Qinshui basin is a large synclinorium basin with bilateral symmetry (Teng et al., 2015). During the past geological ages, the Xinjing coal mine experienced three main geological movements, i.e. the Indosinian period (250 Ma), the Yanshan period (208 Ma) and the Himalayan period (65 Ma). Among the three, the Yanshan period with the NWW-SEE compression stress (the stress is along North-North-West direction or South-East-East direction) has the most impact on the formation of local structures. As a result, the Xinjing coal mine formed a NW plunging syncline with a few NE-NEE secondary folds, where the strata's dip angle is less than 11°.

In this coal mine, there are two main minable coal beds, No. 3 and No. 15 coalbeds. The No. 3 coalbed belongs to Permian Shanxi formation while the No. 15 coalbed belongs to Carboniferous Taiyuan formation. Since the No. 3 coal has been well studied, we focus on the No.15 coal in this study. The study area is about 3.45 $km^2$, located in the middle zone of the coal mine. In the study area, there are 14 wells uncovered the No.15 coal. According to the records of those wells and 3D seismic interpretations, the elevations of the No. 15 coal are shown in Fig. 1(b). In the figure, the No. 15 coal has formed two main anticlines, including the NNE-SSW trend (the axis of anticline is along North-North-East direction or South-South-West direction), and formed two sets of faults (a displacement of rocks along a shear surface) which are along the NNE-SSW trend and the East-West trend, respectively. The faults with the NNE-SSW trend are the reverse faults, and the faults with the E-W trend are normal faults. All the interpreted faults are small faults, which their throws are less than 12 m and their lengths are less than 486 m.

### 2.2. Coalbed characteristics

Fig. 2 shows a stratigraphic column around the No. 15 coal in Xingjin coal mine, where the No. 15 coal is sandwiched between




(a)

(b)

**Fig. 1.** Coal mine location (a) and the overview map of No. 15 coal's elevation in the study area (b). In Figure (b), the gray curves are the axes of anticlines, the black curves are fault traces, and the white ellipses are collapse columns. The '+' markers indicate the locations of the wells. Those beside the '+' markers are the well names and their corresponding TDC thickness.