Evaluation of coalbed methane potential in Parvadeh IV coal deposit in central Iran using a combination of MARS modeling and Kriging

H. Molayemat and F. Mohammad Torab*

Department of Mining and Metallurgical Engineering, Yazd University, Yazd, Iran

Received 14 August 2016; received in revised form 28 November 2016; accepted 12 February 2017

*Corresponding author: fmtorab@yazd.ac.ir (F. Mohammad Torab).

Abstract
Coalbed methane (CBM) plays an important role in coal mining safety and natural gas production. In this work, the CBM potential of B2 seam in Parvadeh IV coal deposit, in central Iran, was evaluated using a combination of local regression and geostatistical methods. As there were 30 sparse methane sampling points in the Parvadeh IV coal deposit, no valid variogram was achieved for the methane content. A multivariate adaptive regression splines (MARS) model was used to reproduce the methane content data based on seam depth, thickness, and ash content. The MARS model results were used in ordinary kriging to estimate the methane content in all mine blocks. A combination of MARS modeling and ordinary kriging in CBM studies is introduced, for the first time, in this paper. The results obtained show that high methane zones are located in the central and south western parts of the deposit. The in situ CBM potential varies from 6.0 to 16.1 m³/t, and it was estimated to be 1.39 billion m³ at the average depth of 267 m in an area of 86.55 km². Although this volume is remarkable, little is known as how much of this resource is actually producible. Consequently, high methane-bearing zones are highly recommended for further studies as a source of natural gas for extraction and reducing the hazards and explosion risks of underground coal mining.

Keywords: Coalbed Methane (CBM), MARS Modeling, Geostatistics, Kriging, Ash, Parvadeh IV Tabas.

1. Introduction
Natural gas is Iran’s primary fuel source to generate electricity. In addition, it is the most common source for warming and cooking in Iran. It is predicted that consumption of natural gas in Iran will increase up to 2.3 times of the current consumption in 2025. Regarding Iran’s 20-year vision plan (2006–2025) and government plans to decrease the dependency on crude oil and having a more variety of fossil fuel resources, coalbed methane (CBM) reserves are one of the targets for natural gas potential studies [1].

CBM exploration and development are important due to mining safety, greenhouse gas emissions, and demand for natural gas [2]. CBM studies are mainly carried out in three fields: methane genesis and depositional factors that affect CBM formation, CBM production and the effective reservoir parameters, and CBM reserve evaluations.

Lazar et al. (2014) have investigated the distribution, composition, and origin of coalbed gases with an analytical procedure in order to reduce the outburst risk in a coal mine [3]. They concluded that in shallow excavation fields, carbon dioxide (CO₂) is the major gas component, while at deeper sites, higher methane (CH₄) values were found. Most CBM researches have focused on methane production [4-8]. Some new CBM production techniques have been suggested by Shi et al. (2014), Keshavaraz et al. (2014), and Ren et al. (2014) [9-11].

All CBM production plans need a precise and reliable evaluation study to model variation in methane across the deposit. A research work has been carried out by Beaton et al. (2006) for CBM resources and reservoir characteristics [12]. They introduced a favorable local area for more detailed studies about production plans. In another study,
Salmachi et al. (2014) have identified potential locations for well placement in a developed CBM reservoir using a geostatistical simulation modeling [13]. This paper presents an approach to estimate the CBM potential of a coal deposit in central Iran based on the structural and non-structural estimation methods. Regarding the discrete nature of the methane content of the Parvadeh IV deposit, a multivariate adaptive regression splines (MARS) model was used to reproduce the methane content of all boreholes based on seam depth, ash, and thickness. The results of the MARS model were used to interpolate the methane content in all mine blocks of the Parvadeh IV deposit using ordinary Kriging. Although there are high gas bearing coal deposits in Iran, previous studies have concentrated on introducing and selecting coal deposits for further CBM detailed studies [14]. However, no CBM reserve estimation has been carried out in Iran. Parvadeh coalfield is the most important coking coal reserve in Iran regarding its vast and high quality coal [15], and also it has been reported as a high gas-bearing coal deposit [14] (Figure 1). These facts making Parvadeh coal deposit as a proper case for CBM estimation.

2. Geology of studied area
Coal seams of Tabas coalfield was formed in a synclinal basin with shale, sandstone, silt, and carbonate rocks in upper Triassic. This depositional stratum is a part of Nayband formation. Parvadeh deposit is bounded by faults including the Rostam fault in the north, the Quri-Chay fault in the south, and the Zenowghan fault in the middle to western parts. These faults have affected the structural and tectonic conditions of the area, and have formed discontinuities and secondary minor faults. Based on these faulting activities, Parvadeh area can be divided into 6 sub-regions, Parvadeh IV being the most favorable part because of its coal quantity and quality [17]. Sedimentary and stratigraphy studies have shown that the coal depositional basin was formed from the north-western parts of Parvadeh area, and was spread to the eastern and south-eastern parts [18]. Seam thickness increases from west to east. Also coal impurities (sulfur and phosphorous) in the western parts are higher than the eastern parts, and some coal seams completely disappear in the south-eastern boundaries [19]. Important coal seams of Parvadeh IV deposit are named B1, B2, C1, C2, and D (Figure 2 A and B). Regarding the extent, continuity, thickness, and quality of coal seams, only C1 and especially B2 are minable. In addition, B2 seam has more gas contents. Thus it is reasonable to choose B2 for gas studies to prevent explosion and also as a probable source of CBM. In the present study, the CBM reserve of B2 seam was evaluated as the most favorable coal seam in Parvadeh IV deposit.

Figure 1. A: Main structural and geological units of Iran. B: Location of Parvadeh coal deposit in central Iran (Figure. 2A area indicated by red rectangle) [16].
3. Multivariate Adaptive Regression Splines (MARS) method

Friedman (1991) has introduced MARS as a statistical method for fitting the relationship between a set of input variables and dependent variables [20]. MARS is a non-linear non-parametric method based on the divide and conquer strategy in which the training data sets are partitioned into separate piecewise linear segments (linear splines) of differing gradients (slopes) [21]. No specific assumption about the underlying functional relationship between the input variables and the output ones is required [20]. The end points of the segments are called knots. A knot marks the end of one region of data and the beginning of another. The resulting piecewise curves (known as Basis Functions (BFs)) give a greater flexibility to the model, allowing for bends, thresholds, and other departures from linear functions [22].

MARS modeling is based on searching and generating BFs in a stepwise procedure. Based on the defined interaction levels of independent variables, MARS searches and finds optimal knot locations. It is possible by using an adaptive regression algorithm. MARS models are constructed in two phases of forward and backward procedure. Adding functions and finding potential knot locations as the forward phase (which results in over-fitting) and pruning the least effective terms as the backward phase (which results in controlling the over-fitting problem) [23].

If Y is the target output and X = (X_1, ..., X_p) is a matrix of p input variables, and it is assumed that the data is generated from an unknown “true” model, in the case of a continuous response, this would be:

\[ Y = f(x_1, ..., x_p) + e = f(x) + e \]  

where e is the distribution of the error. MARS approximates the function f by applying basis functions (BFs). BFs are linear splines (smooth polynomials) including piecewise linear and piecewise cubic functions. For simplicity, only the piecewise linear function is expressed. Piecewise linear functions are of the form \( \max(0, x - t) \) with a knot occurring at value t. The equation \( \max(0, x - t) \) means that only the positive part of (x - t) is used, otherwise it is given a zero value. Formally,

\[ \max(0, x - t) = \begin{cases} 
  x - t, & \text{if } x > t \\
  0, & \text{otherwise} 
\end{cases} \]

The MARS model f(X) is constructed as a linear combination of BFs and their interactions, and is expressed as:

\[ f(X) = \beta_0 + \sum_{m=1}^{M} \beta_m \lambda_m(X) \]
The MARS modeling is a data-driven process. To fit the model in Eq. (3), first a forward selection procedure is performed on the training data. A model is initially constructed with only the intercept $\beta_0$. In each step, the basis pair that produces the largest decrease in the training error is added. Considering a current model with $M$ basis functions, the next pair is added to the model in the form of:

$$\hat{\beta}_{M+1}\mathcal{B}_i(X) \max \left(0, X_j - t\right) + \hat{\beta}_{M+2}\mathcal{B}_i(X) \max \left(0, t - X_j\right)$$

with each $\beta$ being estimated by the method of least squares. As a basis function is added to the model space, interactions between BFs that are already in the model are also considered. This continues until the model reaches some pre-determined maximum number of terms leading to a purposely over-fitted model [23].

To eliminate the over-fitting data and other inconsistencies in the data, a backward deletion sequence follows. The aim of the backward deletion procedure is to find a close to optimal model by removing extraneous variables. The backward pass prunes the model by removing terms one by one, deleting the least effective term at each step until it finds the best sub-model. The model subsets are compared using the less computationally expensive method of Generalized Cross-Validation (GCV). The GCV equation is a goodness of fit test that penalizes large numbers of BFs and serves to reduce the chance of over-fitting. For the training data with $N$ observations, GCV for a model is calculated as follows [20]:

$$GCV = \frac{1}{N} \sum_{i=1}^{N} \left[y_i - f(x_i)\right]^2 \left[1 - \frac{\sum_{i=1}^{N} \left[y_i - f(x_i)\right]^2}{\frac{M + d (M - 1)/2}{N}}\right]$$

where $M$ is the number of BFs, $d$ is the penalizing parameter, $N$ is the number of datasets, and $f(x_i)$ denotes the predicted values of the MARS model. GCV can be described as mean squared error of MARS model based on training data divided by a factor that increases with increase in the model complexity. Note that $(M - 1)/2$ is the number of hinge function knots. GCV penalizes both the number of BFs and the number of knots. Afterwards, a procedure of step by step deletion of BFs is performed to minimize Eq. 3 until an adequate model is found. Since knot locations and BFs are selected based on the input data, MARS is an adaptive method that specified itself to the
problem at hand. After determination of the optimal MARS model, analysis of variance (ANOVA) can be used to determine the relative importance of independent variables in the final model. It is possible by grouping BF s that involve one variable and another grouping of BF s that involve a pairwise interaction and higher interaction levels [19].

In practice, MARS models are simple to interpret and more flexible than linear regression models. Building MARS models often requires little or no data preparation. The hinge function automatically partitions the input data. Thus the effect of outliers is contained. Nevertheless, as with most statistical modeling techniques, known outliers should be considered for removal before training a MARS model [24].

4. CBM Study of Parvardeh IV

Coalbed gas is mainly formed by methane and other hydrocarbons, carbon monoxide, nitrogen, hydrogen sulfide, oxygen, and a little amount of argon and neon. The gas content increases by depth (Figure 4) in Parvadeh IV, and methane forms about 86% of the total gas content (Table 1).

There are 38 gas study boreholes in the Parvadeh IV deposit. As it can be seen in Figure 5, four boreholes are located on or near a folding axis. Also another four boreholes are peripherally located on major and minor faults. Because of the existence of low pressure zones in these fractured and faulted areas, the gas is normally drained out of coal seams, and as a result, the measured methane content would be abnormally low.

Figure 6 illustrates that the methane content of boreholes, which is peripherally located in the faulted and folded zones, demonstrates two separated populations with lower methane contents rather than the expected values. As this study concentrates on CBM of intact parts of Parvadeh IV, it was decided to ignore the data for these boreholes in modeling to prevent inaccuracy in estimations.

The Box and Whisker outlier test and the Johnson transformation function (arcsinh(x)) were used to prepare the data for the 30 remained CBM study boreholes in order to perform the methane content variography. However, due to the small number of sampling points in the deposit, no variogram model was obtained. In order to increase the density of CBM data, a MARS model was applied to all the available data for 154 coal study boreholes all around the deposit (Table 2). MARS model was designed to use depth (measured up to the bottom of the seam), thickness, and ash content as three input independent variables to reproduce methane content in all the 154 sampling points.

The MARS modeling results showed that the correlation coefficient between the actual and predicted methane content values was 0.943 (Figure 7). The mean absolute percentage error (MAPE) and normalized mean squared error (NMSE) of the MARS results were calculated as 5.4% and 0.682%, respectively. It can be concluded that the proposed MARS model is precise in CBM prediction via secondary variables of B2 seam.
Table 1. Gas components in Parvadeh IV deposit [18].

<table>
<thead>
<tr>
<th>Chemical component</th>
<th>Partial percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH₄</td>
<td>86.2</td>
</tr>
<tr>
<td>C₂H₆</td>
<td>6.3</td>
</tr>
<tr>
<td>C₃H₈</td>
<td>3.1</td>
</tr>
<tr>
<td>C₄H₁₀</td>
<td>1.0</td>
</tr>
<tr>
<td>CO₂</td>
<td>1.1</td>
</tr>
<tr>
<td>N₂</td>
<td>2.1</td>
</tr>
<tr>
<td>Other</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Figure 5. Location of 38 gas study boreholes in Parvadeh IV deposit.

Figure 6. Three populations on the basis of methane contents in Parvadeh IV deposit divided into intact, fractured, and folded zones.
Table 2. Statistics of data used in MARS modeling.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Methane content (m³/t)</th>
<th>Depth (m)</th>
<th>Ash content (%)</th>
<th>Seam thickness (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data count</td>
<td>30</td>
<td>154</td>
<td>154</td>
<td>154</td>
</tr>
<tr>
<td>Minimum</td>
<td>5.34</td>
<td>34.86</td>
<td>12.9</td>
<td>0.3</td>
</tr>
<tr>
<td>Maximum</td>
<td>16.62</td>
<td>719.55</td>
<td>39.8</td>
<td>1.5</td>
</tr>
<tr>
<td>average</td>
<td>12.31</td>
<td>266.8</td>
<td>25.2</td>
<td>0.830</td>
</tr>
<tr>
<td>median</td>
<td>12.91</td>
<td>233.5</td>
<td>24.9</td>
<td>0.8</td>
</tr>
<tr>
<td>STD</td>
<td>2.53</td>
<td>156.3</td>
<td>5.17</td>
<td>0.192</td>
</tr>
</tbody>
</table>

Optimum MARS model is achieved by minimum GCV. The model is defined as:

\[
\text{Methane Content} = 12.4971 - 0.0842081 \times BF2 + 0.0211938 \times BF3 + 0.0682841 \times BF4 - 19.5164 \times BF7 + 0.829387 \times BF9 + 0.700052 \times BF10
\]  

where,

\[
BF1 = \max (0, DEPTH - 136.15);
BF2 = \max (0, 136.15 - DEPTH);
BF3 = \max (0, THICKNESS - 0.8) \times BF1;
BF4 = \max (0, 0.8 - THICKNESS) \times BF1;
BF5 = \max (0, ASH - 17.5);
BF7 = \max (0, THICKNESS - 1);
BF9 = \max (0, 0.85 - THICKNESS - 0.85) \times BF5;
BF10 = \max (0, 0.85 - THICKNESS) \times BF5;
\]

Relative effects of independent predictors (depth, thickness, and ash content) with the response methane content are illustrated in Figure 8. It is easy to directly read and compare how the changes in the three input variables influence the behavior of the response methane content in pure ordinal units. The plot of Figure 8A shows that when the depth increases, the methane content also increases but the influence of ash and thickness on methane content are more complicated with partially negative and positive influences (Figure 8B). Also it is notable that the depth knot was calculated as 136.15 m, which means that the methane content behavior in shallow parts of the deposit (depths less than 136.15 m) is different from the deep parts (Figure 9). This separation limit completely conforms to the separation of the oxidized and intact gas zones in Parvadeh IV.

A horizontal block model was designed for the CBM evaluation and coking coal reserve estimation. Regarding the Parvadeh IV long wall mine designs, the block size had to be considered as 100×100 m. The block model was finalized by applying structural boundaries, and it encompassed 8655 blocks, which means that the Parvadeh IV deposit area is 86.55 km².

The Box and whisker outlier recognition test and the normal distribution Johnson transformation function (arcsinh(x)) were used to prepare 154 methane content data to form variograms in multiple directions, and as a result, the spatial
heterogeneity of methane content in Parvadeh IV was extracted in the form of a variogram models and anisotropy map (Figures 10 and 11). It was determined that most of the continuity of methane content values (major axis of anisotropy ellipsoid) was in 120° azimuth and related maximum range distance was 6155 m. Minor axis of anisotropy ellipsoid is along 30° azimuth, and the relative range distance was 1805 m. Afterward, by considering these heterogeneity factors, the Ordinary Kriging (O.K.) and the Johnson inverse transformation were used to estimate the methane content for every single block in the Parvadeh IV deposit (Figure 12). This combination of MARS modeling (as a non-structural estimation method) and O.K (as the best unbiased structural estimator) provides a tool to model the methane content fluctuation precisely.

For the CBM potential evaluation and the coking coal reserve estimation, the thickness and ash content data was transformed into a normal distribution. The spatial heterogeneity of the thickness and ash contents was specified using the variogram models and variogram maps (Figures 10 and 13). These two variables were estimated for 8655 blocks by O.K. (Figures 14 and 15). A regression equation was used to calculate the specific gravity (SpG) of coal in every block based on the estimated ash content (Eq. 7)

\[
SpG = 0.01(\text{ash} \%) + 1.131; \quad R = 0.80
\] (7)
Figure 10. Major experimental variogram and fitted variogram models of A: methane content, B: seam thickness, and C: ash content in Parvadeh IV deposit.
Figure 11. Variogram map of methane content of B₂ seam in Parvadeh IV deposit.

Figure 12. Distribution of methane contents in Parvadeh IV deposit on the basis of combination of MARS model (Equation 6) and O.K.

Figure 13. Variogram map of (A) ash content and (B) thickness of B₂ seam in Parvadeh IV deposit.
Figure 14. Distribution of ash content in Parvadeh IV deposit.

Figure 15. Distribution of B₂ seam thickness in Parvadeh IV deposit.
5. Results
The high methane-bearing blocks are located in the central part of the deposit with the methane content up to 16.1 m$^3$/t. In the other side, the low methane zone is located in the shallow northern part of the deposit with a minimum methane content of 6 m$^3$/t. The high methane content values form a 4000 m wide strip, which is inclined from the north-west to the south-east of the deposit (Figure 12). This strip has a highly local confirmation with the placement of carbonate rocks in the surface (Figure 16). Because of the high weathering-resistance of carbonate rocks (limestone) in the studied area, it formed elevated grounds and hilltop morphological structures. This phenomenon caused an increase in the burial depth beneath the carbonate rocks of Parvadeh IV, and as a result, trapped gas and methane contents increased in the central part of the deposit. Also as the seam dip directed to south west, a high methane content zone was formed in deep levels in south-western parts of the deposit (Figure 12).

![Figure 16. Placement of carbonate rocks in central parts of Parvadeh IV deposit.](image)
Coal reserve of B₂ seam in the Parvadeh IV deposit was calculated as 103.2 million tons with an average 24.2% of ash. This high average ash content is due to a lamination form of thin silt layers in B₂ coal seam in many parts of Parvadeh area, especially in Parvadeh IV. 68.7 million tons of coking coal in Parvadeh IV has 25% or lower ash content, and 34.5 million tons of the reserve has 25% to 40% of ash.

The in situ CBM potential of the Parvadeh IV deposit was estimated as 1.39 billion cubic meter, which lies at a depth of 34.9-719.6 m with an average depth of 267.0 m. The most favorable parts of the deposit for CBM potential were designated by multiplying the seam thickness by the methane content of the blocks. The obtained results are illustrated in Figure 17. It is recommended to drill gas exploratory boreholes and pilot production wells in the discussed favorable areas. Further studies must concentrate on the CBM production factors such as the permeability, porosity, and gas productivity of B₂ seam and its hanging wall and footwall rocks.

![Figure 17. In situ CBM reserve throughout Parvadeh IV deposit.](image)

6. Conclusions
In the present work, the CBM potential of B₂ seam in the Parvadeh IV coalfield in the central Iran was investigated using a combination of local regression (i.e. MARS) and geostatistical methods (i.e. ordinary Kriging). A MARS model was used to reproduce the gas data. Its results were used in ordinary Kriging to estimate the methane content for all mine blocks. This combination resulted in the following conclusions.

Structural factors of the coal deposit such as the burial depth, faulting, and folding have vital effects on CBM studies, and they must be considered as critical factors. The CBM study of the Parvadeh IV deposit showed that different gas zones existed in the area. These zones must be identified clearly, and numerical models must be applied to each zone separately. MARS modeling is a precise and reliable non-structural estimation method when the data has a discrete nature (like CBM data of different gas zones). MARS can be run with a small number of samples, and is a simple model to understand and interpret. It was concluded that MARS could be significantly useful to bind with Kriging to enhance precision of geostatistical estimations.

Parvadeh IV has a high potential of CBM (1.39 billion cubic meters of in-situ methane in an area of 86.55 km² and an average depth of 267.0 m), whereas there is no conventional natural gas
resource in central Iran, and especially, in the surrounding area of the Parvadeh IV deposit, CBM can be proposed as a local source for natural gas production. Therefore, it is highly recommended to study the CBM potential of C1 and other coal seams in the whole Parvadeh area. As methane forms most of the coalbed gas in Parvadeh IV, considering the average thickness of less than 1 m in B2 seam of the deposit, excavation will yield a high-pressure gas. Related studies on the parameters that affect gas eruption are necessary in the Parvadeh IV deposit, and gas effect must be taken into account in outburst and explosion risks.

References


ارزیابی پتانسیل گاز م탄 در کانسار زغال سنگ پروده 4 با استفاده از تلفیق مدل سازی مارس و کریجینگ

حسین ملایمت و فرهاد محمدتراب

دانشکده مهندسی معدن و متالورژی، دانشگاه یزد، ایران

ارسال: 30/8/6152، پذیرش: 56/6/6152

fmtorab@yazd.ac.ir

چکیده:

مطالعه گاز زغال سنگ نقش مهمی در ایمنی معدنکاری و تولید گاز م탄 دارد. در این پژوهش، پتانسیل متان لایه B2 در کانسار زغال سنگ سنگ پروده 1 در ایلام مرکزی به وسیله تلفیقی از رگرسیون محلی و زمین آمار ارزیابی شده است. از آنجایی که داده‌های مطالعه گاز در تنه‌های 91 غواصی انتهایی با لایه‌های متغیر از بین رفته بودند، این مدل رگرسیون چند متغیره تطبیقی (مارس) بر اساس درصد وزنی خاکستر، ضخامت و عمق باید به کار گرفته شد. نتایج این مدل برای تعیین میزان نهایی روش کریجینگ معنی‌داری در نتایج دیگر از تخمین پتانسیل متان در کانسار استفاده شده است. به منظور بازسازی و ارزیابی گاز سنگ در کانسار پروده 1 تلفیق مدل مارس و کریجینگ برای ارزیابی گاز زغال سنگ پروده 1 با استفاده از داده‌های گاز سیل استفاده شد. نتایج نشان داد که مناطق متان خیز در مرکز و جنوب غربی کانسار قرار دارند. پتانسیل میزان با رابطه 4 تا 13/81 میلیارد مترمکعب با عمق نسبی 467 متر، در سطحی به وسعت 88/5 کیلومتر مربع تولید می‌شود. با وجود حجم قابل توجه میزان براورد شده، اطلاعات قابل تولید آن در دست نیست؛ لذا مطالعات پژوهشی و بررسی نرخ تولید میزان در یک بخش مستند کانسار بخش‌های می‌شود. همچنین این مناطق به عنوان تا پایان ریسک بالا انتخاب و بر پایه سنگ معنی‌دار می‌شوند تا ملاحظات طراحی و نهایی مناسب برای استخراج زیرزمینی در پروده 4 به درستی و با ضریب اطمینان بالا لاحق شوند.

کلمات کلیدی: گاز متان زغال سنگ، مدل مارس، زمین آمار، کریجینگ، خاکستر، پروده 4 غواصی.