

# Case-Based reasoning evaluation model of coal resource mining rights

SHAOHUI ZOU<sup>2,3</sup>, ZHILI DING<sup>2,3</sup>, JINSUO ZHANG<sup>3</sup>

**Abstract.** Coal resources will still play a significant role in the evolution of China's energy structure in a long period of time. With the rapid development of coal resource mining rights market, case-based reasoning (CBR) evaluation method becomes more and more important. Currently CBR evaluation method of coal resource mining rights (EM-CRMR) can't solve some key problems including the criteria of case selecting, the set-up of similarity model. Based on CBR and modern statistical theories and methods, this paper reveals the basic principle of case-based reasoning EM-CRMR, defines some important concepts such as attribute index, attribute strength function, underlying asset, case assets, and case similarity. Then it establishes the criteria for selecting cases and the similarity model under the known attribute strength function. Finally, it sets up case-based reasoning EM-CRMR and its application process. In addition, a real case of coal resource mining rights was used to study the application of this model. The results indicate that the price got by using case-based reasoning EM-CRMR is very close to the final bargain price.

**Key words.** Coal resources, mining right, case-based reasoning, case similarity, attribute strength function.

## 1. Introduction

The key to building the internal driving mechanism of optimal development and utilization of Chinese coal resource is that coal producing enterprises can be encouraged to get resources in public market through fair competition [1]. Coal resource mining rights refer to the rights to exploit coal resources and gain exploited coal products within the range permitted by a lawful mining license in China. Case-based reasoning EM-CRMR refers to a method in which the final deal price of similar coal resource mining rights which was dealt or evaluated is to be adjusted to determine the value of the coal resource mining rights awaiting for evaluation. The

---

<sup>1</sup>Acknowledgement - This paper is supported by the National Natural Science Foundation of China (No. 71273207) and the Science and Technology Research and Development Program of Shaanxi(No.2011kjxx54).

<sup>2</sup>Workshop 1 - School of Management, Xi'an University of Science and Technology, Xi'an, China

<sup>3</sup>Workshop 2 - Energy Economy and Management Research Centre, Xi'an University of Science and Technology, Xi'an, China

theoretical basis of case-based reasoning EM-CRMR is CBR theories. Each time when a new problem is solved, new experience can be gained and applied to next new problem [2]. The present researches on CBR focused on case retrieval technique and case revision method, in which constructing reasonable case similarity model is the key to improve case retrieval quality and case revision [3, 4]. Kaoru Hirota and Hajime Yoshino used fuzzy theory to represent fuzzy membership function of case attribute index and then built case similarity function, which laid foundation for the solution of uncertain problems in CBR [5]. These researches [6-8] promoted the application of CBR theories to the decision-making of coal enterprises' operation and management.

Compared with the researches on management decision-making based on CBR, the researches on assets evaluation method based on CBR are mainly focused on real estate evaluation and building cost prediction [9-11]. However, there are so many fundamental differences between real estate and coal resource mining rights in many aspects, such as value attribute, value content, value factors, case description and so on, that real estate evaluation method based on CBR can not be applied to the evaluation of coal resource mining rights directly [12]. It is the precondition of case retrieval to correctly describe the relation between attribute index of coal resource mining rights and bargain price of coal resource mining rights under case-based reasoning. Previous studies have shown that the distance retrieval algorithms based on membership function of attribute index often have some inherent shortcomings, such as large data requirements and long retrieval time [13, 14]. Therefore, this paper firstly sets up case retrieval method and case similarity measuring method based on the attribute strength function of coal resource mining rights which can minimize above deviation, and then puts forward the case-based reasoning EM-CRMR.

## 2. Methodology

### 2.1. concepts definition

Suppose  $P$  is the bargain price of coal resource mining rights,  $y_i (i = 1, 2, \dots, n)$  is attribute indexes,  $n$  is the number of attribute indexes,  $Y$  is vector of attribute indexes and is represented by vector  $Y = (y_1, y_2, \dots, y_n)^T$ . The function  $P = f(Y) = f(y_1, y_2, \dots, y_n)$  is called attribute strength function of  $P$  and  $Y$ . According to the statistical test results of practical application, the relation of  $P$  and  $Y$  shows obvious linear relationship. Thus, the relation of  $P$  and  $Y$  can be represented by:

$$\hat{P} = \hat{a}_0 + \sum_{j=1}^n \hat{a}_j y_j \quad (1)$$

In this equation,  $\hat{P}$  is the estimation of  $P$ ,  $\hat{a}_j (j=1 \sim n)$  is attribute coefficient or attribute factor, and  $\hat{a}_j$  is the unit attribute value of mining resource rights price. SM is used to represent the coal resource mining rights waiting for evaluation,  $P_s$  is the market price of SM,  $Y_s$  represents attribute index vector of SM. The vector consisted by  $n$  attribute indexes of SM is  $Y_s = (y_{s1}, y_{s2}, \dots, y_{sn})^T$ . CM is

used to represent the case coal resource mining rights which are similar to SM and their prices are predictable or available.  $P_c$  is market price or evaluation price.  $Y_c$  is attribute index vector of CM. The vector consisted by  $n$  attribute indexes of CM is  $Y_c = (y_{c1}, y_{c2}, \dots, y_{cn})^T$ . Attribute indexes used to describe case in case base is consistent, the number of SM and CM attribute indexes should be the same. Suppose  $C_a = \{CM_i \mid i=1 \sim m\}$  is the set of alternative coal mining rights observed and gained by evaluators in transacted or evaluated coal mining rights,  $m$  is the quantity of case mining rights, the observed and evaluated price corresponding to  $CM_i$  is  $P_{ci}$ .  $Y_{ci} = (y_{i1}, y_{i2}, \dots, y_{in})^T$ . The matrix of  $m$  case mining rights attribute index values is  $Y_{cm} = [(Y_{ci})^T]_{m \times 1} = [y_{ij}]_{m \times n}$ .  $S$  is the case similarity of SM and CM,  $S \geq 0, \sum_i^m S(CM_i, SM) = 1$ .

## 2.2. model description

Suppose  $P = f(Y) = f(y_1, y_2, \dots, y_n)$  in equation (1) is a known function, which means it can be obtained by linear regression method. Substitute the attribute value of SM into equation (1) and then the estimated price of SM is obtained:

$$\hat{P}_s = \hat{a}_0 + \sum_{j=1}^n \hat{a}_j y_{sj} \quad (2)$$

It's the simplest model. But since  $\hat{a}_j$  is based on similar cases, the evaluation price by equation (2) will cause bigger error. The price of SM should be a weighted means with case similarity. At last, the final estimated price of SM can be obtained the following model.

$$\hat{P}_s = \sum_{i=1}^M S_i P_{si} \quad (3)$$

In this equation,  $M$  is the number of CM in optimal or quasi-optimal case mining rights set  $C_b$  of  $C_a$ .  $C_b$  is the subset of  $C_a$ .  $P_{si}$  ( $i=1, 2, \dots, M$ ) is the adjusted price of SM which is estimated by means of case mining rights  $CM_i$ . If  $CM_i$  is identical with SM ( $Y_s = Y_{ci}$ ),  $P_s$  can be represented by  $P_{ci}$  ( $P_s = P_{ci}$ ). Nevertheless, since  $CM_i$  is not completely identical with SM but just similar to it ( $Y_s \neq Y_{ci}$ ),  $P_s$  is possibly greater or less than  $P_{ci}$ .  $\Delta P_i$  is the deviation between  $P_{ci}$  and  $P_s$ ,  $P_s = P_{ci} + \Delta P_i$ .  $\Delta P_i$  is caused by the difference between attribute indexes ( $Y_s$  and  $Y_{ci}$ ). So  $\Delta P_i$  is actually the price difference between two mining rights caused by the difference of two mining rights attribute indexes. No matter which one it is SM or  $CM_i$  ( $i=1 \sim m$ ), their attribute strength function both have the same regression coefficients  $\hat{a}_j$  ( $j=0 \sim n$ ). Therefore,  $P_{si}$  can be determined by the following equation:

$$P_{si} = P_{ci} + \sum_{j=1}^n \hat{a}_j (y_{sj} - y_{ij}) \quad i = 1, 2, \dots, m \quad (4)$$

In this equation,  $P_{ci}$  is the price of  $CM_i$ , and  $\sum_{j=1}^n \hat{a}_j (y_{sj} - y_{ij})$  is equal to  $\Delta P_i$ .

### 2.3. case selection criterion

Due to the differences between CM and SM, the final price of SM estimated by any subset in alternative case set  $C_a = \{CA_i \mid i=1 \sim m\}$  according to equation (4) will have some deviation from its expected value. Any rational evaluators would expect that such a deviation can be as smaller as possible. Therefore, the smallest deviation between the final estimated price of SM and its expected value can be the case selection criterion. Because  $P_{ci}$  and  $\hat{a}_j$  are random variables,  $P_{si}$  and  $\hat{P}_s$  are also random variables according to equation (3) and (4). Variance  $\hat{\sigma}^2(P_s)$  can be used to measure the deviation between the final estimated price of SM and its expected value. So the case selection criterion is simplified as the smallest variance criterion of the final price of SM. Since  $\hat{P}_s$  is a random variable,  $\hat{\sigma}^2(P_s)$  can be used to represent the deviation of  $\hat{P}_s$  from its expected value. According to the definition of random variable variance, we can get  $\hat{\sigma}^2(\hat{P}_s) = E[\hat{P}_s - E(\hat{P}_s)]^2$ . Substitute equation (??) into it and suppose  $M = m$ , and we can get  $\hat{\sigma}^2(\hat{P}_s) = E[\sum_{i=1}^m S_i P_{si} - E(\sum_{i=1}^m S_i P_{si})]^2$  which can be further evolved into

$$\hat{\sigma}^2(\hat{P}_s) = \sum_{i=1}^m S_i^2 \hat{\sigma}^2(P_{si}) + \sum_{i=1}^m \sum_{j=1}^m S_i S_j \hat{c}ov(P_{si}, P_{sj}) \quad (i \neq j) \quad (5)$$

In this equation,  $S_i (i = 1, 2, \dots, m)$  is the similarity of  $CM_i$ ,  $\hat{\sigma}^2(P_{si})$  is the estimated value of variance  $\sigma^2(P_{si})$  after SM adjusts the price  $P_{si}$ ,  $\hat{c}ov(P_{si}, P_{sj}) (i \neq j)$  is the estimated value of the covariance  $cov(P_{si}, P_{sj}) (i \neq j)$  of  $P_{si}$  and  $P_{sj}$ . When  $i=j$ , the estimated value of covariance of  $P_{si}$  and  $P_{sj}$  is their variance estimated value ( $\hat{c}ov(P_{si}, P_{sj}) = \hat{\sigma}^2(P_{si}) = \hat{\sigma}^2(P_{sj})$ ). Equation (??) can be represented by the following matrix form

$$\hat{\sigma}^2(\hat{P}_s) = S^T D(P) S \quad (6)$$

in which  $S = (S_1, S_2, \dots, S_m)^T$  is the vector of case similarity,  $D(P)$  is the matrix composed of variance and covariance, also

$$D(P) = [D_{ij}]_{m \times m} = \begin{bmatrix} \hat{\sigma}^2(P_{s1}) & \hat{c}ov(P_{s1}, P_{s2}) & \dots & \hat{c}ov(P_{s1}, P_{sm}) \\ \hat{c}ov(P_{s2}, P_{s1}) & \hat{\sigma}^2(P_{s2}) & \dots & \hat{c}ov(P_{s2}, P_{sm}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{c}ov(P_{sm}, P_{s1}) & \hat{c}ov(P_{sm}, P_{s2}) & \dots & \hat{\sigma}^2(P_{sm}) \end{bmatrix} \quad (7)$$

Obviously,  $S_i (i = 1, 2, \dots, m)$  is an unknown variable. Since the adjusted price  $P_{si}$  of SM is definitely somewhat deviant from its expected value  $E(P_{si})$ , which can be measured by the variance of  $P_{si}$ ,  $\sigma^2(P_{si})$ . Then we can get  $\sigma^2(P_{si}) = E[P_{si} - E(P_{si})]^2$ . Substitute equation (??) into  $\sigma^2(P_{si})$ , we can get

$$\sigma^2(P_{si}) = E \left[ \left( P_{ci} + \sum_{j=1}^n a_j (y_{sj} - y_{ij}) \right) - E \left( P_{ci} + \sum_{j=1}^n a_j (y_{sj} - y_{ij}) \right) \right]^2 \quad (8)$$

which can be further introduced and simplified into

$$\hat{\sigma}^2(P_{si}) = \sum_{k=1}^n \sum_{l=1}^n (y_{sk} - y_{ik})(y_{sl} - y_{il})cov(\hat{a}_k, \hat{a}_l) + 2 \sum_{l=1}^n (y_{sl} - y_{il})cov(P_{ci}, \hat{a}_l) + \sigma^2(P_{ci}) \quad (9)$$

If the estimated values of these three parameters,  $\sigma^2(P_{ci})$ ,  $cov(\hat{a}_k, \hat{a}_l)$  and  $cov(P_{ci}, \hat{a}_l)$  can be obtained, we can use equation (9) to calculate the estimated value of  $\sigma^2(P_{si})$ . The calculation of  $\sigma^2(P_{ci})$ ,  $cov(\hat{a}_k, \hat{a}_l)$  and  $cov(P_{ci}, \hat{a}_l)$  can be solved by means of mathematic statistics. Because  $P_{si}$  and  $P_{sj}$  are estimated by equation (9), the estimation  $\hat{cov}(P_{ki}, P_{sj})$  of  $cov(P_{ki}, P_{sj})$  is

$$\hat{cov}(P_{ki}, P_{sj}) = \hat{cov}(P_{ci} + \sum_{k=1}^n \hat{a}_k (y_{sk} - y_{ik}), P_{cj} + \sum_{l=1}^n \hat{a}_l (y_{sl} - y_{jl})) \quad (10)$$

The observed price  $P_{ci}$  ( $i = 1, 2, \dots, m$ ) of each CM is distributed independently, so the covariance between them is zero ( $\hat{cov}(P_{ci}, P_{cj}) = 0$  ( $i \neq j$ )). Because on the linear assumption every case asset attribute factor  $\hat{a}_k$  ( $k = 1, 2, \dots, n$ ) is identical, and  $\hat{a}_k$  and  $P_{ci}$  are mutually independent,  $\hat{cov}(P_{ci}, \hat{a}_l) = \hat{cov}(P_{cj}, \hat{a}_k) = 0$ . Substitute them and equation (4) into equation (10) and then simplify them, we can get

$$\hat{cov}(P_{si}, P_{sj}) = \sum_{k=1}^n \sum_{l=1}^n (y_{sk} - y_{ik})(y_{sl} - y_{jl})\hat{cov}(\hat{a}_k, \hat{a}_l) \quad (i \neq j) \quad (11)$$

Equation (11) is the estimated formula of the final estimated price variance of SM. On the basis of equation (11),

this paper proposed the following case selection criterion:

Stage one: according to the estimated value  $\hat{\sigma}^2(P_{si})$  of variance  $P_{si}$ , evaluators should select the case mining rights with smaller  $\hat{\sigma}^2(P_{si})$  from alternative case set  $C_a = \{CM_i \mid i=1 \sim m\}$  to form an optimal case mining rights set  $C_b = \{CA_i \mid i=1 \sim M\}$ . Specific principles are:

a. if  $\hat{\sigma}^2(P_{si})$  are all smaller, all case mining rights in alternative case set  $C_a = \{CM_i \mid i=1 \sim m\}$  can be selected in principle. But due to the cost, both the dates of case mining rights transaction and mining rights evaluation should be taken into account for selection. In another word, the CM with smaller  $\hat{\sigma}^2(P_{si})$  and closer date to the evaluation date of SM should be selected.

b. if  $\hat{\sigma}^2(P_{si})$  has clear sizes, the case mining rights with bigger  $\hat{\sigma}^2(P_{si})$  should be removed and the rest ones should be selected. If the rest CM has a large quantity, we can select the CM with smaller  $\hat{\sigma}^2(P_{si})$  and closer date to the evaluation date of SM.

c. if  $\hat{\sigma}^2(P_{si})$  are all bigger, it's necessary to readjust and select alternative CM set.

Stage two: select the optimal case mining rights set from quasi-optimal case mining rights set  $C_b = \{CA_i \mid i=1 \sim M\}$ . It means if there is a case subset  $C_a^{(*)}$  corre-

sponding to its case similarity vector  $S^{(*)}$  which makes

$$\hat{\sigma}^2(P_s^{(*)}) = \min_l \{ (S^{(l)})^T D(P^{(l)}) S^{(l)} \} \tag{12}$$

the case subset  $C_a^{(*)}$  is the optimal case set in which  $l=1 \sim \sum_{i=1}^M (M!/i!(M-i)!)$ .

**2.4. case similarity model**

According to equation (12), if the similarity of the cases in  $C_b^{(l)}$  makes  $\hat{\sigma}^2(P_s^{(l)})$  the smallest the case subset  $C_b^{(*)}$  corresponding to  $S^{(*)}$  is the optimal case set,  $S^{(*)}$  is the similarity of optimal case set. In order to get the solution of  $S^{(*)}$  suppose the number of case subset  $C_b^{(l)}$  in quasi-optimal case mining rights set  $C_b$  is  $M^{(l)}$   $M^{(l)} \leq M$ ,  $C_b^{(l)} = \{CM_1, CM_2, \dots, CM_{M^{(l)}}\}$  the case similarity vector corresponding to  $C_b^{(l)}$  is  $S^{(l)} = (S_1, S_2, \dots, S_{M^{(l)}})^T$  the adjusted price of underlying mining rights corresponding to  $C_b^{(l)}$  is  $P^{(l)} = (P_{s1}, P_{s2}, \dots, P_{s_{M^{(l)}}})^T$ , and the matrix of variance and covariance is  $D(P^{(l)})$ . On the condition that the last case subset  $C_b^{(l)}$  of  $C_b$  satisfies equation(12), the optimized model can be set up as follows:

$$\begin{aligned} & \text{Min } (M^T \Lambda) \\ \text{st. } & \begin{cases} S^{(l)} - [I - \dot{S}^{(l)} \theta^T] D(P)^{-1} \Lambda = \dot{S}^{(l)} \\ D(P^{(l)}) S^{(l)} - \vec{\theta} \lambda - \Lambda = 0 \\ S^{(l)} \geq 0 \\ \Lambda \geq 0 \end{cases} \end{aligned} \tag{13}$$

In this equation,  $\lambda$  and  $\Lambda$  are Lagrange multiplier and multiplier vector restraining  $(S^{(l)})^T \theta = 1$  and  $S^{(l)} \geq 0$ ,  $\vec{\theta}$  is the unit vector of  $M^{(l)} \times 1$ ,  $\Lambda = (\Lambda_1, \Lambda_2, \dots, \Lambda_{M^{(l)}})^T$ ,  $\vec{\theta}^T B^{(l)} = 0$ ,  $\dot{S}^{(l)} = D(P)^{-1} \theta / \theta^T D(P)^{-1} \theta$  and  $B^{(l)} = [I - \dot{S}^{(l)} \theta^T] D(P)^{-1}$ .

$I$  is the unit matrix of  $M^{(l)} \times M^{(l)}$ . The optimal solution of this model  $\hat{S}^{(l)} = (S_1^*, S_2^*, \dots, S_{M^{(l)}}^*)^T$  is the optimal similarity of the cases in case subset  $C_b^{(l)}$ . For the last case subset of  $C_b$ , the variance subset of corresponding case similarity vector and final SM estimated price can be obtained by solving equation (??). Among these variance subsets, there is the smallest variance  $\hat{\sigma}^2(\hat{P}_s^{(*)})$  definitely. So the case similarity vector subset corresponding to it is the similarity vector of the optimal case set  $\hat{S}^{(*)} = \{S_1^*, S_2^*, \dots, S_{M^{(*)}}^*\}$ , and the case subset corresponding to  $\hat{S}^{(*)}$  is the optimal case set  $C_b^{(*)} = \{CM_1^*, CM_2^*, \dots, CM_{M^{(*)}}^*\}$ . After solving  $\hat{S}^{(*)} = \{S_1^*, S_2^*, \dots, S_{M^{(*)}}^*\}$ , we can substitute them and  $P_{si}$ ,  $i=1 \sim M^{(*)}$  corresponding to it into equation (4) to calculate the final estimated price of SM.

### 2.5. evaluation model

The final estimated price of underlying mining rights  $\hat{P}_s$  can be figured out by means of equation (4) after solving the case similarity of optimal CM. Because  $P_{ci}$  ( $i=1,2,\dots,M$ ) is unit price (the coal resource mining right price of every ton of coal), the calculation model of the total value of the SM is

$$V = Z \cdot \sum_{i=1}^M S_i^* [P_{ci} + \sum_{j=1}^n \hat{a}_j (y_{sj} - y_{ij})] \quad (14)$$

in which  $Z$  is coal resources reserves, and equation (14) is the case-based reasoning EM-CRMR.

### 2.6. case application

Coal mine A is located in the western Yuyang District of Yulin of Shaanxi Province, close to Inner Mongolia Autonomous Region. According to the attribute properties of coal mine A, this paper selected 10 coal mines of Shenhua Shendong Group, and 2 local coal mines in Yulin as the research objects of case mining rights. This paper used expert assessment method (establishing specific grading system first, and then inviting chief engineers of coal mine to assess it). The results are seen in table 1.

Table 1. The attribute indexes value of underlying mining rights and case mining rights  $\lambda$  for a trapezoidal plate for different values of taper constant  $\beta_1$  and constant aspect ratios  $a/b = 1.0$ ,  $c/b = 0.5$

Cases <i>SM</i> or <i>CM</i>	Case Prices	Attribute Indexes					
		Coal Seam Thickness $y_1$	Coal Seam Stablility $y_2$	Ash Con- tent $y_3$	Calorific Value $y_4$	Mining Depth $y_5$	Tons of Coal Profits $y_6$
Case0		7.021	7.180	6.361	6.440	6.819	6.230
Case1	6.35	7.454	7.577	7.589	6.987	6.897	6.586
Case2	9.53	6.324	5.637	6.023	6.213	6.321	5.547
Case3	8.623	6.324	6.857	6.551	6.241	6.324	5.547
Case4	8.345	7.114	6.745	6.546	6.234	6.211	5.745
Case5	9.234	5.324	5540	5.513	6.030	5.730	7.390
Case6	8.203	7.740	7.101	6.891	6.509	6.853	5.922
Case7	10.021	7.120	6.612	7.249	6.870	6.340	5.680
Case8	4.115	7.049	6.522	6.470	6.190	6.550	5.201
Case9	3.874	6.060	6.470	6.230	6.081	6.790	5.010
Case10	4.356	6.578	5.687	6.374	6.897	6.532	5.623
Case11	3.456	7.324	7.547	7.856	4.214	5.241	6.321
Case12	3.014	5.241	4.254	4.325	5.324	6.321	4.214

Note: Before inviting experts to assess every attribute index value, we established standards in advance. The upper limit of every attribute index value is 10.

According to the data in table 1, the attribute function of coal resource assets (regression equation) is regressed, which is

$$\hat{P} = 21.72 + 0.71 y_1 + 2.93 y_2 - 3.68 y_3 + 5.19 y_4 - 7.99 y_5 + 0.25 y_6 + 0.84 y_7$$

in which  $\hat{a}_0 = 21.72, \hat{a}_1 = 0.71, \hat{a}_2 = 2.93, \hat{a}_3 = 0.71, \hat{a}_4 = -3.68, \hat{a}_5 = 5.19, \hat{a}_6 = -7.99, \hat{a}_7 = 0.25, \hat{a}_8 = 0.84$ . The significance of regression equation is tested by means of statistics  $F$ . When  $\alpha = 0.01$ , we can check the distribution list of  $F$  and get critical value  $F_{0.01}(7, 4) = 14.97$  ( $F = 80.22 > 14.97$ ). So the regression equation above is very significant. Because when  $\alpha = 0.01$ , we can check the distribution list of  $t$  and get critical value  $t_{1-\alpha/2}(12-7-1) = t_{0.995}(4) = 5.441$ ,  $t_1 = 44.324$ ,  $t_2 = 50.152$ ,  $t_3 = 44.571$ ,  $t_4 = 34.810$ ,  $t_5 = 6.006$ ,  $t_6 = 45.206$ ,  $t_7 = 4.206$ , variables  $y_1, y_2, y_3, y_4, y_5, y_6$  and  $\hat{P}$  are very evident or relevant to high evidence. It has something to do with underlying mining rights and case mining rights being in the same region. Since attribute factor matrix is  $\hat{a}_j = [(Y^T Y)^{-1} Y^T P]_j$ .

The variance of residual and the variance of regression coefficient—covariance matrix are



$$\hat{\sigma}^2(\varepsilon) = \bar{\varepsilon}^T \bar{\varepsilon} / (K - n - 1) \text{ and } D(\hat{A}_a) = \left[ [\hat{\sigma}^2(\varepsilon) (Y^T Y)^{-1}]_{ij} \right]_{n \times n} \quad (i, j = 1, \dots, n)$$

in which  $K = 12, n = 7$ , the estimated value of residual variance is  $\hat{\sigma}^2(\varepsilon) = \bar{\varepsilon}^T \bar{\varepsilon} / (K - n - 1) = 0.0024$ . According to these data, variance—covariance matrix  $D(\hat{A}_a)$  of attribute factor

$$\begin{aligned} \hat{A}_a &= (\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{a}_5, \hat{a}_6, \hat{a}_7)^T \\ &= (-42.824, 4.845, 4.112, 3.338, 1.275, 3.214, 1.225, 1.252)^T \end{aligned}$$

is shown in table 2.

Table 2. variance—covariance matrix of attribute factors  $D(\hat{A}_a)$

$D(\hat{A}_a)$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$	$\hat{a}_7$
$\hat{a}_1$	2.86E-02	-5.05E-01	-2.39E-01	1.60E-01	9.81E-02	-2.96E-01	-2.22E-01
$\hat{a}_2$	-5.05E-01	8.90E+00	4.22E+00	-2.82E+00	-1.73E+00	5.22E+00	3.91E+00
$\hat{a}_3$	-2.39E-01	4.22E+00	2.00E+00	-1.34E+00	-8.21E-01	2.48E+00	1.85E+00
$\hat{a}_4$	1.60E-01	-2.82E+00	-1.34E+00	8.95E-01	5.48E-01	-1.66E+00	-1.24E+00
$\hat{a}_5$	9.81E-02	-1.73E+00	-8.21E-01	5.48E-01	3.36E-01	-1.02E+00	-7.59E-01
$\hat{a}_6$	-2.96E-01	5.22E+00	2.48E+00	-1.66E+00	-1.02E+00	3.07E+00	2.29E+00
$\hat{a}_7$	-2.22E-01	3.91E+00	1.85E+00	-1.24E+00	-7.59E-01	2.29E+00	1.72E+00

Second, the optimal case assets set is selected according to the first criterion in case selection criteria. By means of formula  $\hat{\sigma}^2(P_{ci}) = \hat{\sigma}^2(\varepsilon)(1 + 1/K)$  we can get the variance estimation of case mining rights price  $P_{ci}$ . Based on the data above, we can calculate the adjusted price of this underlying mining rights and its variance estimated value as follows (seeing in table 3).

Table 3. The estimated price of underlying assets and its variance estimation

Cases $CA_i$	The adjusted price estimation of underlying mining rights $P_{si}$	The variance estimation of adjusted price $\hat{\sigma}^2(P_{si})$	Standard deviation $\hat{\sigma}(P_{si})$	Ordering according to the criterion of smallest-smaller variance
1	8.32	0.045	0.193	7
2	15.36	0.0268	0.157	5
3	14.25	0.0222	0.146	1
4	12.31	0.0242	0.159	4
5	12.54	0.0244	0.152	3
6	16.74	0.0267	0.169	6
7	14.45	0.0245	0.152	2
8	6.32	0.038	0.174	8
9	17.21	0.039	0.197	9
10	18.24	0.0360	0.196	10
11	6.32	0.042	0.189	11
12	7.25	0.046	0.201	12

From table 3 we can discover that the adjusted price variances of underlying mining rights corresponding to case assets 1, 8, 9, 10, 11, 12 are relatively bigger (above 0.030). According to the first criterion of case selection criteria, these 6 cases can be omitted and the rest case mining rights constitute optimal case set  $C_b = \{CA_2, CA_3, CA_4, CA_5, CA_6, CA_7\} = \{2, 3, 4, 5, 6, 7\}$

Third, optimal case similarity is figured out. According to the data of  $Y_s = (y_{s1}, y_{s2}, \dots, y_{s7})^T$ ,  $Y_{cm} = [(Y_{ci})^T]_{7 \times 1} = [y_{ij}]_{7 \times 7}$  and ,the variance—covariance matrix  $D(P) = [D_{ij}]_{6 \times 6}$  of the adjusted price of underlying mining rights is calculated as follows (table 4).

Table 4. variance—covariance matrix of underlying assets adjusted price

$D_{ij}$	$P_{s2}$	$P_{s3}$	$P_{s4}$	$P_{s5}$	$P_{s6}$	$P_{s7}$
$P_{s2}$	0.0291	0.0242	0.0293	0.0142	0.0147	0.0178
$P_{s3}$	0.0242	0.0142	0.0262	0.0145	0.0172	0.0290
$P_{s4}$	0.0293	0.0262	0.0217	0.0240	0.0220	0.0329
$P_{s5}$	0.0142	0.0145	0.0240	0.0242	0.0240	0.0317
$P_{s6}$	0.0147	0.0172	0.0220	0.0240	0.0114	0.0410
$P_{s7}$	0.0178	0.0290	0.0329	0.0317	0.0410	0.0268

In all the subsets of quasi-optimal case set

$C_b = \{CA_2, CA_3, CA_4, CA_5, CA_6, CA_7\} = \{2, 3, 4, 5, 6, 7\}$ , there are 6 subsets consisted

of 1 case, 15 subsets  $l=7\sim 21$  consisted of 2 cases, 20 subsets  $l=22\sim 41$  consisted of 3 cases, 15 subsets  $l=42\sim 56$  consisted of 4 cases, 6 subsets  $l=57\sim 62$  consisted of 5 cases, and 1 subset  $l=63$  consisted of 6 cases  $l=63$ . Similarity models are built for these subsets (seen in equation (16)). This paper set up corresponding calculation procedures, taking the example of case subset  $C_b^{(53)} = \{CA_3, CA_4, CA_5, CA_7\} = \{3, 4, 5, 7\}$  to illustrate the determination of similarity. From table 4 we can know that the variance—covariance matrix of underlying mining rights adjusted price

$P^{(53)} = (P_{s3}, P_{s4}, P_{s5}, P_{s7})^T$  corresponding to case subset  $C_b^{(53)}$  is

$$D(P^{(53)}) = \begin{bmatrix} 0.0232 & 0.0293 & 0.0142 & 0.0147 \\ 0.0293 & 0.0262 & 0.0217 & 0.0224 \\ 0.0142 & 0.0217 & 0.0240 & 0.0220 \\ 0.0147 & 0.0224 & 0.0220 & 0.0240 \end{bmatrix}$$

Its inverse matrix is

$$D(P^{(53)})^{-1} = \begin{bmatrix} 202.291 & -95.930 & -15.623 & -67.862 \\ -95.930 & 202.567 & -114.147 & 6.156 \\ -15.623 & -114.147 & 184.155 & -38.023 \\ -67.862 & 6.156 & -38.023 & 116.175 \end{bmatrix}$$

The optimized model of similarity is built according to equation (16) as follows

$$\begin{aligned} \underset{S^{(53)}}{\text{Min}} \hat{\sigma}^2(\hat{P}_{S^{(53)}}) / 2 &= (0.0212 S_3^2 + 0.0252 S_4^2 + 0.0231 S_5^2 + 0.0231 S_7^2 \\ &+ 0.0384 S_3 S_4 + 0.0344 S_3 S_5 + 0.0340 S_3 S_7 \\ &+ 0.0414 S_4 S_5 + 0.0334 S_4 S_7 + 0.0330 S_5 S_7) / 2 \\ \text{st.} \quad &\begin{cases} S_3 + S_4 + S_5 + S_7 = 0 \\ S_3, S_4, S_5, S_7 \geq 0 \end{cases} \end{aligned}$$

Substitute  $D(P^{(53)})$  and  $D(P^{(53)})^{-1}$  into equation (10), and then we can get

$$\dot{S}^{(53)} = (\dot{S}_3, \dot{S}_4, \dot{S}_5, \dot{S}_7)^T = (0.320, 0.180, 0.317, 0.183)^T$$

Likewise, the variance between the case similarity of case subsets and its final estimated price can be calculated. Table 5 shows the smallest variance case subset of those subsets consisted of one, two, three, four, five or six cases and its corresponding similarity.

Table 5. The smallest variance case subset of those subsets consisted of 1~6 cases and its similarity

Subcase	Case similarity	Min variance
$C_b^{(l)} = \{CM_i\} (i=3, 4, 5, 6, 7)$ F3,5,7 F3,4,5,7 F3,4,5,6,7 F2,3,4,5,6,7)	$S_2^* S_3^* S_4^* S_5^* S_6^* S_7^*$	
$CM_3$	1	0.0212
$CM_3 CM_7$	0.592 0.408	0.0195
$CM_3 CM_5 CM_7$	0.398 0.293 0.309	0.0185
$CM_3 CM_4 CM_5 CM_7$	0.320 0.180 0.317 0.183	0.0165
$CM_3 CM_4 CM_5 CM_6 CM_7$	0.398 0 0.293 0 0.309	0.0172
$CM_2 CM_3 CM_4 CM_5 CM_6 CM_7$	0.412 0.288 0 0.293 0 0.307	0.0175

From table 5 we can see that there are 4 smaller variance case subsets of estimated prices. So the optimal case set is  $C_b^{(*)} = \{CM_3^*, CM_4^*, CM_5^*, CM_7^*\} = \{CM_3, CM_4, CM_5, CM_7\}$ . Its similarity is

$$\hat{S}^{(*)} = (0.320, 0.180, 0.317, 0.183)^T .$$

Four, the final estimated price  $\hat{P}_s$  of underlying mining rights and the total value of coal mining rights waiting for evaluation are calculated. Substitute  $\hat{S}^{(*)} = (0.320, 0.180, 0.317, 0.183)^T$  and its adjusted price (seeing in table 3) into equation (15) to get the final estimated price of underlying mining rights which is

$$\hat{P}_s = 0.320 \times 11.25 + 0.180 \times 10.31 + 0.317 \times 9.54 + 0.183 \times 12.45 = 10.76(\text{Yuan/t})$$

The total value of coal mining rights waiting for evaluation is  $V_1 = 31.07 \times 10.76 = 33.43(\text{billion Yuan})$ .

### 3. Results

The mining rights price of coal mine **A** is 4.45 Yuan/t, which was estimated by some assets evaluation organization in June of 2012 by means of discount cash flow method. However, if using real option method, the evaluation result is 7.85 Yuan/t. In fact, the market price of coal mining rights in 2012 in the region where coal mine **A** is located is about 12 Yuan/t. Therefore, the results of coal resource mining rights using case-based reasoning EM-CRMR is very close to the market price.

### 4. Conclusion

Through defining key concepts such as attribute index, attribute strength function, underlying asset, case assets, and case similarity, this paper solves the key problems about case selection criteria, and case similarity model of coal mining rights. The application of a real case also indicates that the price got by using CBR coal mining rights evaluation method is very close to the final deal price.

## References

- [1] S. H. ZOU, J. S. ZHANG, N. C. JAIN: *Study on the direction of reforming coal resource development tax and fee*[J]. CHINA MINING MAGAZINE 18 (2009), No. 8, 17–21.
- [2] S. H. LIAO: *Expert system methodologies and applications-A decade review from 1995 to 2004*[J]. Expert Systems with Applications 28 (2005), No. 1, 93–103.
- [3] P. JUELL, P. PAULSON: *Case-based systems*[J]. IEEE Intel. Syst. Applic 18 (2003), No. 4, 60–67.
- [4] H. KAORU, Y. HAJIME, M. Q. XU, ET AL: *An Application of Fuzzy Theory to the Case-based Reasoning of the CISG*[J]. Journal of Advanced Computational Intelligence 1 (1997), No. 2, 86–93.
- [5] Y. C. ZHOU, X. Y. LI: *Relationship Value Evaluation by Difference Relation Case-based Reasoning*. OPERATIONS RESEARCH AND MANAGEMENT SCIENCE 19 (2010), No. 3, 70–74.
- [6] Q. GUO, Y. H. GUO, C. X. GUO: *Alternative Ranking Based on Fuzzy Similarity in Group Decision-Making*. JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY 45 (2010), No. 2, 307–311.
- [7] L. C. GU, Q. Y. GUO, M. R. CAO, D. L. ZHANG: *Case retrieval algorithm based on structure matrix*. Journal of Multimedia 8 (2013), No. 6, 712–720.
- [8] X. K. ZHANG, W. Q. CHEN, W. S. CHENG, S. C. LI : *The Market Comparison Approach Based on the Fuzzy Mathematics*. Chinese Journal of Management Science 9 (2001), No. A203, 37–42.
- [9] S. H. JI, M. PARK, H. S. LEE: *Case Adaptation Method of Case-Based Reasoning for Construction Cost Estimation in Korea*[J]. Journal of Construction Engineering and Management 138 (2012), No. 1, 43–52.
- [10] S. Y. KIM, J. H. SHIM: *Combining case-based reasoning with genetic algorithm optimization for preliminary cost estimation in construction industry*[J]. Canadian Journal of Civil Engineering 41, (2014), No. 1, 65–73.
- [11] S. H. ZOU, J. H. XU, B. DONG: *The Case-Based Reasoning Evaluating Model of Coal Resource Mining Right With Unknown Attribute Strength Function* [J]. CHINA MINING MAGAZINE 25 (2016), No. 5, 42–47.
- [12] H. Z. SU, J. Y. LI, J. P. CAO, Z. P. WEN: *Macro-comprehensive evaluation method of high rock slope stability in hydropower projects*[J]. Stochastic Environmental Research and Risk Assessment 28 (2014), No. 2, 213–224.

Received November 16, 2016

