

巴伦诺尔一矿煤发热量的综合预测

陈晓华¹, 王举龙²

(1. 中国煤炭科工集团有限公司 沈阳设计研究院 辽宁 沈阳 110015;

2. 申克(天津)工业技术有限公司 北京 100010)

摘要:通过分析巴伦诺尔一矿煤质得到了 17 组煤质分析数据,包括 M_{ad} 、 A_d 、 H_{daf} 、 $Q_{gr,d}$ 。利用多元回归分析的原理,建立煤的发热量关于煤中水分和灰分含量的多元回归方程,并通过 R 检验、 F 检验、 t 检验,证明了回归方程的作用显著,即具有实用价值。但是线性回归分析仅能对煤的发热量进行估算,并不能精确预测。因此,采用了支持向量机(SVM)算法对多元线性回归的初步预测结果进行小范围修正,修正结果显示总体预测精度明显提高,这 2 种方法的结合效果优于常用方法。

关键词: 煤发热量; 多元线性回归; 支持向量机; 预测模型

中图分类号:TQ533

文献标识码:A

文章编号:1006-6772(2012)03-0067-04

Forecast of calorific value of coal in Balunnuoer NO. 1 coal mine

CHEN Xiao-hua¹, WANG Ju-long²

(1. Shenyang Design and Research Institute, China Coal Technology and Engineering Group Co., Ltd., Shenyang 110015, China;

2. Shenck (Tianjin) Industrial Technology Co., Ltd., Beijing 100010, China)

Abstract: Get seventeen groups coal analysis data of Balunnuoer NO. 1 coal mine, involving M_{ad} , A_d , H_{daf} and $Q_{gr,d}$. Deduce the multiple regression equation about the calorific value of coal and M_{ad} , A_d based on the principle of multiple regression analysis. The R test, F test, t test show that the regression equation has practical value. However, the results are not accurate. So the support vector machine (SVM) is adopted to revise the preliminary forecasting results on a small scale. The results show that, through revising, the overall forecasting accuracy is greatly improved and it's more efficient than conventional methods.

Key words: calorific value of coal; multiple linear regression; support vector machine; prediction model

煤发热量是评价煤炭质量的一项重要指标,也是评价燃煤特性及经济性的重要参数。煤发热量受水分、灰分含量影响显著,因此,通过研究煤中水分和灰分的含量及对煤发热量的相对作用可以预测煤的发热量。

通过多元回归分析建立煤的发热量关于煤中水分和灰分含量的多元回归方程,利用 EXCEL 计算

待定系数,并经过 R 检验、 F 检验、 t 检验,表明该方程的可行性和可靠性。但是煤的发热量与水分、灰分之间的关系十分复杂,不仅在总体趋势上呈现一定的相关性,而且细节上也具有较强的非线性关系。因此,用线性回归分析仅能对煤发热量进行估算,难以精确预测。应用支持向量机(SVM)算法对多元线性回归的初步预测结果进行小范围修正,以

减小预测过程中的随机误差。

1 巴伦诺尔一矿煤质分析

巴伦诺尔一矿位于巴彦胡硕盆地西南部,可采煤层为 2-3 2-6 2-8 号煤层,首采区主要开采 2-8

号煤层,煤种为褐煤。采用锡林郭勒盟华鼎矿业有限公司编制的《内蒙古自治区西乌珠穆沁旗巴彦胡硕煤田巴伦诺尔一区煤炭勘探报告》的钻孔资料。从中提取了 17 个有代表性的钻孔,表 1 为巴伦诺尔一矿煤质分析。

表 1 巴伦诺尔一矿煤质分析

样本序号	原始数据				折算后数据			
	$M_{ad}/\%$	$A_d/\%$	$H_{daf}/\%$	$Q_{gr,d}/(\text{MJ}\cdot\text{kg}^{-1})$	$M_t/\%$	$H_{ad}/\%$	$Q_{gr,ad}/(\text{MJ}\cdot\text{kg}^{-1})$	$Q_{net,ar}/(\text{MJ}\cdot\text{kg}^{-1})$
1	13.03	8.95	4.21	26.48	25.03	3.62	23.03	18.63
2	13.61	9.57	3.99	26.02	25.61	3.41	22.48	18.16
3	11.22	13.87	4.09	24.11	23.22	3.59	21.40	17.34
4	14.06	6.25	4.26	27.65	26.06	3.62	23.76	19.20
5	13.79	10.08	4.24	25.91	25.79	3.62	22.34	17.99
6	11.73	8.46	4.22	26.11	23.73	3.69	23.05	18.71
7	10.19	8.08	4.29	26.38	22.19	3.81	23.69	19.34
8	14.76	8.05	3.76	25.99	26.76	3.17	22.15	17.86
9	10.96	12.95	4.28	24.35	22.96	3.77	21.68	17.56
10	18.63	7.80	3.37	27.48	30.63	2.71	22.36	17.88
11	16.14	9.27	4.09	26.52	28.14	3.40	22.24	17.81
12	11.65	12.37	4.42	25.16	23.65	3.87	22.23	17.98
13	12.26	20.34	4.45	22.74	24.26	3.87	19.95	15.98
14	11.35	7.12	4.34	26.89	23.35	3.81	23.84	19.40
15	12.72	17.79	4.14	23.25	24.72	3.58	20.29	16.30
16	12.56	23.60	3.98	21.00	24.56	3.45	18.36	14.67
17	15.64	12.35	3.79	25.25	27.64	3.17	21.30	17.08

表 1 中 M_t 根据井田水文地质情况及地质报告中 M_{ad} 预计。通过表中参数,根据公式(1)^[1] 计算出 $Q_{net,ar}$ 。

$$Q_{net,ar} = \frac{(Q_{gr,ad} - 206H_{ad})(100 - M_t)}{100 - M_{ad}} - 23M_t \quad (1)$$

将 M_t 、 A_d 作为独立的自变量, $Q_{net,ar}$ 作为受这 2 个自变量影响的因变量,设 $x_1 = M_t$, $x_2 = A_d$, $y = Q_{net,ar}$, 选择二元线性回归分析的数学模型为 $y = a + bx_1 + cx_2$ 。

2 煤发热量模型建立

2.1 回归模型的建立

EXCEL 电子制表系统中函数的语法分为函数名和参数两部分,参数用圆括号括起,之间用逗号隔开。进行回归分析时,主要采用线性回归函数 Linest 辅以使用索引取值 Index 与四舍五入 Round 函数^[2]。表 2 为 EXCEL 回归结果。表 2 中,首先在 A20 中输入公式: = ROUND(LINEST(D2: D18 ,B2: C18 ,TRUE ,TRUE) ,2) ,然后进行数组操作:选中 A20: C24 按 F2 再按 Shift+Ctrl+Enter。

表 2 EXCEL 回归结果

	A	B	C	D
1	样本序号	M_t	A_d	$Q_{net,ar}$
2	1	25.03	8.95	18.63
3	2	25.61	9.57	18.16
4	3	23.22	13.87	17.34
5	4	26.06	6.25	19.20
6	5	25.79	10.08	17.99
7	6	23.73	8.46	18.71
8	7	22.19	8.08	19.34
9	8	26.76	8.05	17.86
10	9	22.96	12.95	17.56
11	10	30.63	7.80	17.88
12	11	28.14	9.27	17.81
13	12	23.65	12.37	17.98
14	13	24.26	20.34	15.98
15	14	23.35	7.12	19.40
16	15	24.72	17.79	16.30
17	16	24.56	23.60	14.67
18	17	27.64	12.35	17.08
19	回归值:			
20	-0.26	-0.18	25.35	
21	0.01	0.03	0.70	
22	0.97	0.22	#N/A	
23	254.58	14.00	#N/A	
24			#N/A	

回归结果: $a=25.35$ $b=-0.18$ $c=-0.26$

则发热量回归方程为

$$Q_{\text{net,ar}} = 25.35 - 0.18M_1 - 0.26A_d$$

2.2 全相关系数 R 检验

回归模型对数据拟合度的好坏,通常通过全相关系数衡量。全相关系数是指因变量与多个自变量之间的相关关系。一般在 $0 \sim 1$ 变化,全相关系数越接近 1,表明回归方程对样本数据的拟合度越好。表 3 为 EXCEL 全相关系数计算结果。

表 3 EXCEL 全相关系数计算结果

	A	B	C	D
1	样本序号	M_1	A_d	$Q_{\text{net,ar}}$
2	1	25.03	8.95	18.63
3	2	25.61	9.57	18.16
4	3	23.22	13.87	17.34
5	4	26.06	6.25	19.20
6	5	25.79	10.08	17.99
7	6	23.73	8.46	18.71
8	7	22.19	8.08	19.34
9	8	26.76	8.05	17.86
10	9	22.96	12.95	17.56
11	10	30.63	7.80	17.88
12	11	28.14	9.27	17.81
13	12	23.65	12.37	17.98
14	13	24.26	20.34	15.98
15	14	23.35	7.12	19.40
16	15	24.72	17.79	16.30
17	16	24.56	23.60	14.67
18	17	27.64	12.35	17.08
19	全相关系数:			
20	$R^2 =$	0.9732		

在 B20 中输入公式: =INDEX(LINESIT(D2: D18 , B2: C18 ,TRUE ,TRUE) ,3 ,1) ,可得出 $R^2 = 0.9732$,即全相关系数 $R = 0.9865$ 。

查相关系数临界值表^[3] ($\alpha = 0.05$) $R_{0.05(17-3)} = 0.4973$ $R = 0.9865 > 0.4973$,因此 M_1 A_d 和 $Q_{\text{net,ar}}$ 之间有很显著的线性相关性。

2.3 F 检验

F 检验又称方差齐性检验,主要通过比较两组数据的方差,来确定它们的精密度是否有显著性差异。其值按式(2)计算

$$F = \frac{S_R / m}{S_E / (n - m - 1)} \quad (2)$$

式中 S_R 为回归平方和; S_E 为残差平方和; m 为变量

个数; n 为样品个数。

表 4 为方差分析结果^[4]。

表 4 方差分析

项目	df(自由度)	SS(平方和)	MS(均方)	F(统计值)
回归分析	2	24.10968	12.05484	254.5813
残差	14	0.662923	0.047352	
总计	16	24.77261		

在给定的显著水平 ($\alpha = 0.05$) 时,查 F 分布表^[3] 得 $F_{0.05(2,14)} = 3.74$,因为 $F = 254.5813 > 3.74$,说明 $\alpha = 0.05$ 水平上求得的回归方程是有意义的,具有较高的置信度, M_1 A_d $Q_{\text{net,ar}}$ 之间的共同影响显著^[5]。

2.4 t 检验

F 检验只是说明煤发热量与水分、灰分整体之间有较好的线性关系,但并不能说明单一的水分或单一的灰分对煤发热量影响是显著的。因此需要通过 t 检验来证明这两者对煤发热量的单独影响程度。

在给定的显著水平 ($\alpha = 0.05$) 时,通过查询 t 分布表可以得到 $t_{0.05/2(17-3)} = 2.1448$,通过计算得到水分和灰分对煤发热量的 t 值分别为

$$t_1 = -0.1829 / 0.02584 = -7.078$$

$$t_2 = -0.25765 / 0.01145 = -22.502$$

水分和灰分 t 值的绝对值均大于临界值 2.1448,说明水分和灰分对于预测煤发热量均是重要变量。

3 发热量预测模型的修正

再精确的预测模型都可能因为对其它影响因素的忽略产生预测误差。预测误差通常是指预测值与实际值的偏差,一般预测模型只考虑主要影响因素,忽略次要影响因素。而在这些次要因素的长期作用下,预测误差可能会形成一定趋势。误差修正就是考虑了这些次要因素,拟合误差趋势的一个过程,其主要目的是为了结合不同模型的优势,最终提高模型预测精度的过程。预测误差修正的基本方法:用一种预测方法建立预测模型,通过预测模型计算出初始值,与原始值进行比较,就能得到该预测方法的预测误差,再使用合适的预测方法对误差进行预测,得到预测误差修正量,因而就能得到更加准确的预测结果。

支持向量机(SVM)是当前应用比较广泛的非

线性预测建模工具,与人工神经网络方法相比,SVM模型的预测精度更高,收敛速度更快,泛化能力更强。SVM回归算法可通过专用工具箱实现,现在通用的SVM的工具箱有2种:一种为SVM通用工具箱,另一种工具箱为SVM OSU_3.00工具箱,都要在Matalab软件的界面下进行使用。设平衡因子取500,拟合精度取0.01,核函数选用RBF径向基函数^[6]: $K(X_i, X_j) = \exp(-\|X_i - X_j\|^2 / 2\sigma^2)$,径向基参数取 $\sigma = 0.2$ 。表5为发热量原始值与预测值比较。

表5 发热量原始值与预测值比较

样本 序号	$Q_{net,ar} / (\text{MJ} \cdot \text{kg}^{-1})$			
	原始值	线性模型 预测初始值	SVM 误差修正 模型结果	综合模型 预测结果
1	18.63	18.52	0.085	18.60
2	18.16	18.25	-0.021	18.23
3	17.34	17.56	-0.093	17.47
4	19.20	19.03	0.059	19.09
5	17.99	18.09	-0.020	18.07
6	18.71	18.88	-0.088	18.79
7	19.34	19.26	0.059	19.31
8	17.86	18.44	-0.214	18.23
9	17.56	17.85	-0.136	17.71
10	17.88	17.81	0.078	17.89
11	17.81	17.87	-0.003	17.87
12	17.98	17.88	0.086	17.96
13	15.98	15.69	0.225	15.92
14	19.40	19.30	0.144	19.44
15	16.30	16.28	0.043	16.32
16	14.67	14.79	-0.081	14.71
17	17.08	17.16	-0.040	17.12

由表5可以看出,用SVM修正子模型后,线性回归模型的预测误差得到了较好的修正,预测精度明显提高。

4 结 语

煤的发热量是煤质研究和煤质评价的一项重要指标。通过多元回归方程来预测煤的发热量,可以节约测试时间和测试费用。但由于煤的发热量与工业分析各数据之间的关系十分复杂,通过传统的多元线性回归方法很难达到预期的效果。为此,运用多元线性回归方法减小预测模型的系统误差,用支持向量机(SVM)算法对多元线性回归的初步预测结果进行小范围修正,上述测试结果表明这2种方法的结合应用,能充分发挥各自的特点,可以进行优势互补,预测效果很好。

参考文献:

- [1] 戴少康. 选煤工艺设计使用手册[M]. 北京: 煤炭工业出版社, 2010.
- [2] 陈晓华, 邵金元. 利用 EXCEL 进行煤炭发热量的快速回归分析[J]. 煤炭技术, 2006(9): 120-122.
- [3] 盛骤, 谢式干, 潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 1997.
- [4] 刘晓国, 赵志根, 黄文辉. 汝淇沟煤矿煤炭发热量的回归分析研究[J]. 洁净煤技术, 2011, 17(2): 51-53.
- [5] 李建华. 煤矸石灰分与发热量的回归分析及其运用[J]. 洁净煤技术, 2004, 10(2): 58-60.
- [6] 王晓红, 吴德会. 一种燃煤发热量的综合预测方法[J]. 煤炭科学技术, 2006, 34(6): 16-18.

(上接第62页)

压缩后脱水,设备投资高,TEG损耗少,综合运行费用低,但会增加压缩工段负荷和投资;压缩前脱水,设备投资低,TEG损耗高,综合运行费用高,不会增加压缩工段负荷和投资。

5 结 语

TEG脱水工艺技术成熟、运行稳定,必将在煤制天然气安全、顺利的管输过程中发挥重要的作用。但是,目前TEG脱水装置在国内煤制天然气行业尚无运行业绩,压缩后脱水工艺和压缩前脱水工艺哪种更为经济、合理,目前还有待于今后工业装

置实际运行验证。

参考文献:

- [1] 朱瑞春, 公维恒, 范少锋. 煤制天然气工艺技术研究[J]. 洁净煤技术, 2011, 17(6): 81-85.
- [2] 何策, 张晓东. 国内外天然气脱水设备技术现状及发展趋势[J]. 石油机械, 2008, 36(1): 69-73.
- [3] 郝蕴. 三甘醇脱水工艺探讨[J]. 中国海上油气(工程), 2001, 6(3): 22-29.
- [4] 李明国, 徐立, 张艳玲, 等. 天然气脱水生产中三甘醇的使用情况[J]. 钻采工艺, 2005, 28(3): 107-108.