

# 基于主成分分析及距离判别法的 煤矿安全状态判别模型的研究

刘中原, 张法兴, 程艳晨, 孙均雨

青岛理工大学理学院, 山东 青岛

收稿日期: 2023年2月27日; 录用日期: 2023年3月28日; 发布日期: 2023年4月10日

## 摘要

本文致力于煤矿竖井安全状态判别模型的研究, 采用主成分分析方法解析影响煤矿竖井安全的多特征因素, 简化多元数据结构, 然后结合距离判别方法将处理过的样本数据作为判别因子, 得到相应的判别函数进行预测, 建立了一套基于主成分分析和距离判别法的煤矿竖井安全状态判别模型, 实现了煤矿竖井安全状态有效判别, 准确率达到100%, 保证了煤矿安全运营, 这对煤矿竖井井筒的安全性和稳定性的评价具有借鉴意义。

## 关键词

煤矿, 竖井, 安全状态判别

# Research on the Discrimination Model of Coal Mine Safety State Based on Principal Component and Distance Discrimination Method

Zhongyuan Liu, Faxing Zhang, Yanchen Cheng, Junyu Sun

School of Science, Qingdao University of Technology, Qingdao Shandong

Received: Feb. 27<sup>th</sup>, 2023; accepted: Mar. 28<sup>th</sup>, 2023; published: Apr. 10<sup>th</sup>, 2023

## Abstract

This paper draws lessons from the principal component analysis theory of multiple information of coal mine information refining, the multiple related, the information of index variables, combining into unrelated comprehensive index instead of the original index, simplifying multiple data struc-

tures, then combining the processed sample data as the discrimination method, get the corresponding discriminant function to predict, established a set of coal mine shaft safety discrimination based on principal component analysis and distance discrimination model, realize the coal mine shaft safety state effective discrimination, accuracy of 100%, ensure the safe operation of coal mine. This has reference significance for evaluating the safety and stability of coal mine shaft.

## Keywords

Coal Mine, Shaft, Security Status Discrimination

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

二十世纪八十年代以来,徐淮、兖州等地不同矿区的井壁破裂、水喷砂、安全问题多种多样,给提升运输带来困难,导致矿山停产。据统计,仅在兖州和徐淮矿区发生的煤矿竖井断裂造成的损失就不计其数。因此,准确预测和判别井筒破裂的地质灾害的发生,对矿井的安全运营和经济效益具有重要意义[1]。目前,大多数学者认为井筒周围的低部含水层会因失水而下沉,并产生对井筒的负阻力(附加力),从而导致井筒断裂。通过进一步的煤矿竖井破裂室内模拟试验的分析和推测,井筒破裂主要由井筒周围深厚土体、水体和井筒的相互作用导致的[2] [3] [4]。

判别和预测井筒破裂的方法主要有两种[5]:一是通过观测地表沉降、井筒变形、应力和地下水来研究煤矿竖井变形规律,对煤矿竖井破裂预测判别;其次,通过开发新的智能分析方法、智能技术、煤矿开采观测方法的多信息综合分析来对煤矿竖井安全状态进行预测,这些方法能更好地预测一些不确定性问题,特别是在工程地质灾害预测领域,如应用神经网络、灰色理论或深度学习分析方法来预测工程的可靠性。但这些方法大多没有考虑到井筒不稳定性指标数据重叠造成误判率高的问题,最终导致决策失误,造成安全隐患和经济损失。黎锦贤等[6]采用主成分分析法,建立了煤矿安全评价的综合数学函数评价模型,实现了煤矿安全程度定量化目标评价。金洪伟等[7]为解决煤矿瓦斯涌出量指标繁杂致使预测精度低的问题,运用主成分分析法对瓦斯涌出量进行降维预测,取得了较好的预测效果。丁坤等[8]采用主成分分析对原始多维变量进行预处理,并利用距离来表征光伏系统的健康状态,研究结果表明该方法能够更加灵敏、准确地反映光伏系统的性能状态。

综上,本文借鉴主成分分析理论解析煤矿多特征因素,把煤矿安全状态评价过程中多个彼此相关、数据重叠的指标变量进行重组,组合成互不相关的综合指标代替初始指标,从中优选较少的综合指标来反映初始指标的信息,克服煤矿安全状态评价中信息重叠和指标繁杂导致预测不准的问题,有效的描述不同煤矿竖井现有特征[9]。主成分分析本质是通过数学降维来简化多元数据结构,然后结合距离判别方法将处理过的样本数据作为判别因子,得到相应的判别函数进行预测,以此建立一套基于主成分分析和距离判别法的井筒安全状态判别模型,对煤矿竖井安全状态进行判别。

## 2. 主成分分析法与距离判别法的概念

### 2.1. 主成分分析法概念与基本原理

在多元指标处理过程中经常遇到高维知识集,由于信息具有较高的空间性和大量的变量,这些变量

之间通常存在一些相关性,因此这些样本知识很难复制整体的大部分信息,太多的变量会对计算量产生影响,增加分析的复杂性。主成分分析法作为一种空间约简和提取技术,将知识引入低维区域,尽可能减少信息损失,初始指标的线性组合被用来表示主成分,减少了信息的维度,简化了信息结构,指标的重要性完全取决于线性组合中原指数系数绝对值的大小[10]。

假设煤矿竖井安全状态涉及  $P$  个评价指标,分别用  $X_1, X_2, \dots, X_p$  来表示,这  $P$  个指标构成  $P$  维随机向量,  $X = (X_1, X_2, \dots, X_p)$ , 转换矩阵为  $A$ , 对随机变量  $X$  进行线性变换,形成新的综合变量,用  $Y$  表示。新的综合变量可以由原始变量线性表示  $Y = AX$ , 即[11]:

$$\begin{aligned} Y_1 &= a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ Y_2 &= a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ &\vdots \\ Y_p &= a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p \end{aligned} \quad (1)$$

式中,满足  $a_{i1} + a_{i2} + \dots + a_{ip} = 1$ ;  $Y_i$  与  $Y_j$  之间互不相关 ( $i \neq j$ ;  $i, j = 1, 2, \dots, p$ );  $Y_1$  是  $X_1, \dots, X_p$  的一切线性组合中方差最大的;  $Y_1, Y_2, \dots, Y_p$  的方差之和等于  $X_1, \dots, X_p$  方差之和。主成分的求解过程也就是求转换矩阵  $A$  的过程。求解主成分的一般步骤如下:

(1) 由于选取的影响竖井井筒破裂的特征因素具有不同的量纲,为了消除多元信息的量纲差异,需要对样本数据进行归一化处理,归一化后,数据被转化为无量纲的纯量,便于不同单位和不同量级的指标解析和比较。考虑到样本数据的离散性和数据处理的方便性,本文采用离差标准化方法对数据进行归一化处理。离差标准化方法的表达式为:

$$v_k = \frac{V_k - \min}{\max - \min} \quad (2)$$

式中,  $k$  为样本数据的序号数,  $v_k$  是样本数据的归一函数值。

(2) 计算  $P$  个变量之间的协方差矩阵  $\Sigma$ , 得到特征向量  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ , 单位向量为  $T_1, T_2, \dots, T_p$ 。其中,令转换矩阵  $A = T'$ , 即  $A$  的第  $i$  行就是  $\Sigma$  的第  $i$  个特征根对应的单位特征向量  $T_i$ , 且第  $i$  个主成分  $Y_i$  的方差就等于  $\Sigma$  的第  $i$  个特征根  $\lambda_i$  [11]。

(3) 第  $k$  个主成分  $Y_k$  的方差贡献率为  $\eta_k = \lambda_k / \sum_{k=1}^p \lambda_k$ 。若取  $m$  ( $m < p$ ) 个主成分,主成分  $Y_1, Y_2, \dots, Y_m$  的累积贡献率为  $\xi_m = \sum_{k=1}^m \lambda_k / \sum_{k=1}^p \lambda_k$  [12]。

(4) 主成分个数取决于累积方差贡献率,通常取  $m$  个主成分使得其方差的累积贡献率达到 80% 以上,则对应的前  $m$  个主成分的样本信息量包含  $p$  个原始变量所能提供的绝大部分信息[13]。

## 2.2. 距离判别法概念与基本原理

距离判别法是先计算样本数据到各个分类的距离,根据距离的大小建立相应的判别规则,得出判别函数,进而对未知样本进行分类[14]。

设有  $k$  个  $m$  维总体:  $G_1, G_2, \dots, G_k$  ( $k > 2$ ), 从中任意取 2 个总体  $G_p, G_q$  相同,新样本  $X$  到总体  $G_p$  和  $G_q$  的马氏距离平方差为[14]:

$$d^2(X, G_p) - d^2(X, G_q) = -2(W_q(X) - W(X)) \quad (3)$$

式中:

$$\begin{aligned} W_p(X) &= (S^{-1}u_p)^T X - 0.5u_p^T S^{-1}u_p \\ W_q(X) &= (S^{-1}u_q)^T X - 0.5u_q^T S^{-1}u_q \end{aligned} \quad (4)$$

由此有:

$$\begin{aligned} d^2(X, G_q) \geq d^2(X, G_p) &\Leftrightarrow W_p(X) \geq W_q(X) \\ d^2(X, G_q) < d^2(X, G_p) &\Leftrightarrow W_p(X) < W_q(X) \end{aligned} \quad (5)$$

一般情况下, 总体的均值向量  $u = \{u_1, u_2, \dots, u_m\}^T$  和公共协方差矩阵  $\Sigma$  是未知的, 可以用各总体的学习样本作估计。

假设  $X_{(t)}^{(q)} = \{x_{t1}^{(q)}, x_{t2}^{(q)}, \dots, x_{tm}^{(q)}\}$  ( $q = 1, 2, \dots, k; t = 1, 2, \dots, n_q$ ), 为来自总体  $G_q$  的学习样本, 其中  $n_q$  为总体  $G_q$  的学习样本个数, 则  $u_q$  的无偏估计为[14]:

$$\bar{X}^{(q)} = \left\{ \frac{1}{n_q} \sum_{t=1}^{n_q} x_{t1}^{(q)}, \dots, \frac{1}{n_q} \sum_{t=1}^{n_q} x_{tm}^{(q)} \right\}^T \quad (6)$$

则学习样本的协方差矩阵  $\Sigma_q$  的无偏估计为组内协方差矩阵  $S_q$ :

$$S_q = \frac{1}{n_q - 1} \sum_{t=1}^{n_q} (X_{(t)}^{(q)} - \bar{X}^{(q)}) (X_{(t)}^{(q)} - \bar{X}^{(q)})^T \quad (7)$$

当各总体的协方差矩阵相等时, 总体协方差矩阵  $\Sigma$  的无偏估计为  $W_q(X)$ 。

$$S = \frac{1}{\sum_{q=1}^k n_q - k} \sum_{q=1}^k (n_q - 1) S_q \quad (8)$$

以  $\bar{X}^{(q)}$  和  $S$  分别代替  $u_q$  和  $\Sigma$ , 得到  $W_q(X)$  ( $q = 1, 2, \dots, k$ ) 的估计为:

$$\bar{W}_q(X) = (S^{-1} \bar{X}^{(q)})^T X - 0.5 (\bar{X}^{(q)})^T S^{-1} \bar{X}^{(q)} \quad (9)$$

因此, 多总体情况下的距离判别准则为, 若总体  $G_q$  满足式(10):

$$W_{q0}(X) = \max_{1 \leq q \leq k} W_q(X) \quad (10)$$

则判别  $X \in G_q$ 。

### 3. 煤矿竖井井筒安全状态判别模型及应用

根据兖州矿区立井井筒破裂的实际资料[15], 提取表土层厚度、底含厚度、底含水位速降、井筒外径、井壁厚度和井筒投入使用时间 6 个特征因素, 如下表 1 所示。采用其中 16 组数据作为主成分分析和距离判别法的样本集, 剩余 4 组作为井筒安全状态判别模型精度可靠性的检验样本。

**Table 1.** System resulting data of standard experiment [15]

**表 1.** 各矿井特征因素集[15]

井筒名称	表土层厚度/m	底含厚度/m	底含水位速降/m	井筒外径/m	井壁厚度/m	井筒投入使用时间/月	实际状态
兴隆庄主井	189.31	34.10	60.225	8.92	1.21	192	破裂
兴隆庄副井	190.41	30.00	60.225	10.10	1.30	225	破裂
兴隆庄东风井	190.41	32.85	60.000	6.40	0.70	241	破裂
兴隆庄西风井	189.50	29.90	70.000	7.40	0.95	230	破裂
鲍店主井	148.69	56.29	84.000	8.50	1.00	194	破裂

## Continued

鲍店副井	148.60	55.00	82.000	10.00	1.00	187	破裂
鲍店北风井	202.56	59.00	80.000	6.60	0.80	190	破裂
杨村主井	185.50	57.72	87.500	6.40	0.70	146	破裂
杨村副井	184.50	56.00	87.000	7.00	1.00	139	破裂
杨村北风井	173.40	65.30	86.200	5.90	0.70	136	破裂
兴隆庄主井	189.31	34.10	0	8.92	1.21	0	完整
兴隆庄副井	190.41	30.00	0	10.10	1.30	0	完整
兴隆庄东风井	190.41	32.85	0	6.40	0.70	0	完整
兴隆庄西风井	189.50	29.90	0	7.40	0.95	0	完整
鲍店主井	148.69	56.29	0	8.50	1.00	0	完整
鲍店副井	148.60	55.00	0	10.00	1.00	0	完整
鲍店北风井	202.56	59.00	0	6.60	0.80	0	完整
杨村主井	185.50	57.72	0	6.40	0.70	0	完整
杨村副井	184.50	56.00	0	7.00	1.00	0	完整
杨村北风井	173.40	65.30	0	5.90	0.70	0	完整

## 3.1. 主成分分析

为了后续模型建立方便性, 将影响井筒破裂的表土层厚度、底含厚度、底含水位速降、井筒外径、井壁厚度和井筒投入使用时间 6 个特征因素分别用  $V_1$ 、 $V_2$ 、 $V_3$ 、 $V_4$ 、 $V_5$ 、 $V_6$  表示。对选取的 6 个评价指标进行归一化处理(式 2), 得到对应的归一化后的变量  $v_1$ 、 $v_2$ 、 $v_3$ 、 $v_4$ 、 $v_5$ 、 $v_6$ , 其数值为表 2 所示。

Table 2. Normalized shaft sample data

表 2. 归一化处理后的井筒样本数据

井筒名称	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$
兴隆庄东风井	0.58	0.11	0.69	0.12	0	1
兴隆庄西风井	0.55	0	0.8	0.37	0.83	0.95
鲍店主井	-0.85	1	0.96	0.63	1	0.8
鲍店副井	-0.85	0.95	0.94	1	1	0.78
鲍店北风井	1	1.1	0.91	0.17	0.33	0.79
杨村主井	0.41	1.05	1	0.12	0	0.61
杨村副井	0.38	0.99	0.99	0.27	1	0.58
杨村北风井	0	1.34	0.99	0	0	0.56
兴隆庄东风井	0.58	0.11	0	0.12	0	0
兴隆庄西风井	0.55	0	0	0.37	0.83	0
鲍店主井	-0.85	1	0	0.63	1	0
鲍店副井	-0.85	0.95	0	1	1	0
鲍店北风井	1	1.1	0	0.17	0.33	0
杨村主井	0.41	1.05	0	0.12	0	0
杨村副井	0.38	0.99	0	0.27	1	0
杨村北风井	0	1.34	0	0	0	0

**Table 3.** Correlation coefficient matrix of each feature factor  
**表 3.** 各特征因素相关系数矩阵

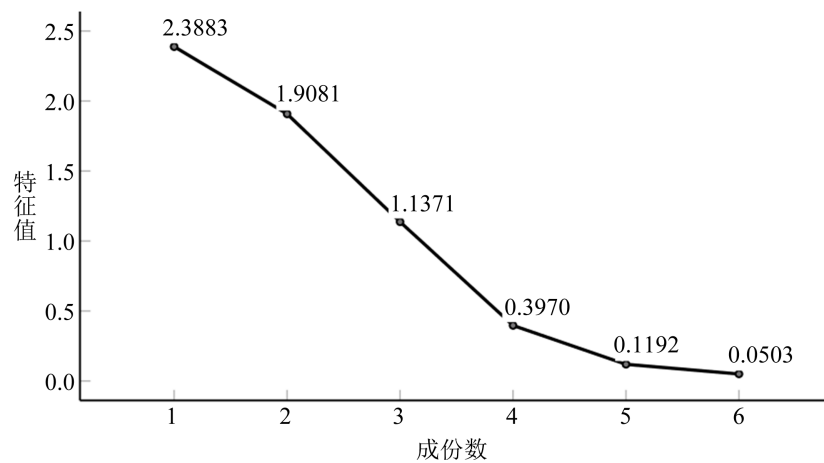
指标	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$
$v_1$	1.000					
$v_2$	-0.296	1.000				
$v_3$	-0.039	0.100	1.000			
$v_4$	-0.767	-0.014	0.012	1.000		
$v_5$	-0.509	-0.063	0.022	0.774	1.000	
$v_6$	0.023	-0.167	0.908	0.043	0.019	1.000

**Table 4.** Characteristic values, variance contribution and cumulative contribution  
**表 4.** 各成分特征值、方差贡献率及累积贡献率

成分	初始特征值			提取载荷平方和		
	特征值/ $\lambda_i$	贡献率%	累积贡献率%	特征值/ $\lambda_i$	贡献率%	累积贡献率%
$Y_1$	2.388	39.805	39.805	2.388	39.805	39.805
$Y_2$	1.908	31.801	71.606	1.908	31.801	71.606
$Y_3$	1.137	18.952	90.558	1.137	18.952	90.558
$Y_4$	0.397	6.616	97.174			
$Y_5$	0.119	1.987	99.161			
$Y_6$	0.050	0.839	100.000			

**Table 5.** Principal component coefficient matrix  
**表 5.** 主成分系数矩阵

成分	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$
$Y_1$	-0.5585	0.0874	0.0705	0.6135	-0.5442	-0.0531
$Y_2$	0.0944	-0.1039	1.0096	-0.0588	-0.0504	1.0274
$Y_3$	-0.2541	0.9012	0.1726	-0.1425	-0.2588	-0.0769



**Figure 1.** Principal component analysis gravel map  
**图 1.** 主成分分析碎石图

由表 3 各特征因素之间的相关系数矩阵可以看出, 输入的几个特征因素彼此之间具有明确的相关性, 如:  $v_3$  和  $v_6$  相关系数达 0.908、 $v_4$  和  $v_5$  相关系数达 0.774、 $v_4$  和  $v_1$  相关系数达 0.767 等, 这必然会对井筒安全状态判别的精度造成影响, 容易导致井筒安全状态的误判。因此, 有必要对样本数据进行主成分分析。通过编写 R 语言程序, 对表 2 中归一化处理的样本数据进行主成分分析, 得到各成分特征值、方差贡献率及累积贡献率如表 4 所示, 把输出结果的每个元素除以特征值的平方根  $\sqrt{\lambda}$ , 得到主成分分析的系数矩阵如表 5 所示。

主成分的提取顺序是根据特征值大小从大往小依次选取的, 特征向量也根据相应的特征值大小顺序依次选取, 特征值越大, 相对应的主成分因素越重要。主成分数据的选取一般要求大于 80%, 即累积贡献率不小于 80% 的原则。由表 4 可知, 前 3 个主成分特征值  $\lambda_1$ 、 $\lambda_2$  和  $\lambda_3$  的累积贡献率为 90.558%, 即包含了原始数据的 90.558% 的信息, 可以概括原始变量的主要信息, 同时与图 1 中的主成分分析碎石图相吻合。由表 5 可知,  $Y_1$  包含原始变量  $v_1$ 、 $v_4$ 、 $v_5$  的信息较多, 它反映了原始变量 39.805% 的信息;  $Y_2$  主要是对原始变量  $v_3$ 、 $v_6$  表征, 它反映了原始变量 31.801% 的信息;  $Y_3$  主要包含原始变量  $v_2$  的信息, 它反应了原始变量 18.952% 的信息, 这与图 1 中主成分分析碎石图相吻合。依据表 5 中各主成分系数矩阵, 可得到  $Y_1$ 、 $Y_2$ 、 $Y_3$  与原始变量  $v_1$ 、 $v_2$ 、 $v_3$ 、 $v_4$ 、 $v_5$ 、 $v_6$  之间的函数表达式为:

$$Y_1 = -0.5585v_1 + 0.0874v_2 + 0.0705v_3 + 0.6135v_4 + 0.5442v_5 + 0.0531v_6 \quad (11)$$

$$Y_2 = 0.0944v_1 - 0.1039v_2 + 1.0096v_3 - 0.0588v_4 - 0.0504v_5 + 1.0274v_6 \quad (12)$$

$$Y_3 = -0.2542v_1 + 0.90125v_2 + 0.1726v_3 - 0.1426v_4 - 0.2588v_5 - 0.0769v_6 \quad (13)$$

因此, 根据式 11、式 12 和式 13 计算得到处理后新的样本数据, 如下表 6 所示。

**Table 6.** Data obtained after the principal component analysis  
**表 6.** 主成分分析后的数据

指标	$Y_1$	$Y_2$	$Y_3$
兴隆庄东风井	-0.14	1.76	-0.02
兴隆庄西风井	0.48	1.78	-0.34
鲍店主井	1.6	1.52	0.87
鲍店副井	1.82	1.46	0.77
鲍店北风井	-0.07	1.69	0.73
杨村主井	0.04	1.55	0.95
杨村副井	0.68	1.46	0.62
杨村北风井	0.22	1.44	1.34
兴隆庄东风井	-0.24	0.04	-0.06
兴隆庄西风井	0.37	-0.01	-0.41
鲍店主井	1.49	-0.27	0.77
鲍店副井	1.72	-0.29	0.67
鲍店北风井	-0.18	-0.05	0.63
杨村主井	-0.06	-0.08	0.83
杨村副井	0.58	-0.13	0.5
杨村北风井	0.12	-0.14	1.21

### 3.2. 距离判别模型的构建与检验

对表 1 中前 16 组样本数据进行训练, 后 4 组样本作为验证。以破裂和完整作为 2 个不同的总体, 并假设 2 个总体的协方差矩阵相等。将上述通过主成分分析方法得到的三个主成分指标  $Y_1$ 、 $Y_2$ 、 $Y_3$  作为距离判别模型的判别因子, 按照上文提出的距离判别分析计算理论进行计算、建模, 程序中将井筒破裂和完整状态分别用数字“1”和数字“0”表征。然后, 对训练样本进行计算、学习后可求得相应的判别系数, 如表 7 所示, 进而得到煤矿竖井井筒的安全状态判别函数:

$$\psi(Y) = 1.448Y_1 + 2.506Y_2 + 1.557Y_3 \quad (14)$$

**Table 7.** Distance discriminant function coefficient

**表 7.** 距离判别函数系数

指标	$Y_1$	$Y_2$	$Y_3$
判别函数系数	1.448	2.506	1.557

经过学习后的模型, 利用留一交叉验证法对学习样本进行判别, 判别结果全部正确, 并对 100.0% 个进行了交叉验证的已分组数据进行了分类, 正确率达到了 100%, 如下表 8 所示。

**Table 8.** Training results of the distance discrimination method

**表 8.** 距离判别法训练结果

井筒名称	判别因子			实际状态	分类	训练结果
	$Y_1$	$Y_2$	$Y_3$			
兴隆庄东风井	-0.14	1.76	-0.02	破裂	1	破裂
兴隆庄西风井	0.48	1.78	-0.34	破裂	1	破裂
鲍店主井	1.6	1.52	0.87	破裂	1	破裂
鲍店副井	1.82	1.46	0.77	破裂	1	破裂
鲍店北风井	-0.07	1.69	0.73	破裂	1	破裂
杨村主井	0.04	1.55	0.95	破裂	1	破裂
杨村副井	0.68	1.46	0.62	破裂	1	破裂
杨村北风井	0.22	1.44	1.34	破裂	1	破裂
兴隆庄东风井	-0.24	0.04	-0.06	完整	0	完整
兴隆庄西风井	0.37	-0.01	-0.41	完整	0	完整
鲍店主井	1.49	-0.27	0.77	完整	0	完整
鲍店副井	1.72	-0.29	0.67	完整	0	完整
鲍店北风井	-0.18	-0.05	0.63	完整	0	完整
杨村主井	-0.06	-0.08	0.83	完整	0	完整
杨村副井	0.58	-0.13	0.5	完整	0	完整
杨村北风井	0.12	-0.14	1.21	完整	0	完整

采用上面已经完成学习的距离判别分析模型, 对剩余 4 个验证样本进行判别, 判别结果与实际情况



全部相符, 判别正确率达 100%, 如下表 9 所示。

**Table 9.** Verify the sample discrimination results  
**表 9.** 验证样本判别结果

井筒名称	表土层厚度/m	底含厚度/m	底含水位速降/m	井筒外径/m	井壁厚度/m	井筒投入使用时间/月	实际状态	预测结果
兴隆庄主井	189.31	34.10	60.225	8.92	1.21	192	破裂	1
兴隆庄副井	190.41	30.00	60.225	10.10	1.30	225	破裂	1
兴隆庄主井	189.31	34.1	0	8.92	1.21	0	完整	0
兴隆庄副井	190.41	30.00	0	10.10	1.30	0	完整	0

综上, 基于主成分分析及距离判别法的煤矿安全状态判别模型的交叉验证结果和测试样本的判别结果均与实际情况完全相符。由此认为, 本文提出的判别模型对于煤矿竖井井筒的安全状态预测预报是完全可行且有效的。

#### 4. 结论

(1) 本文从煤矿竖井井筒非采动破裂特点以及灾害的成因出发, 提取了影响煤矿竖井井筒非采动破裂的几种特征因素, 采用主成分分析方法解析煤矿竖井破裂的多特征因素, 通过降维处理克服预测指标繁杂和信息重叠的影响, 遴选出反映煤矿竖井破裂的主要判别因子, 提高预测的精度。

(2) 依据距离判别分析基本原理, 对上述通过主成分分析方法得到的判别因子进行建模、计算, 建立了反映煤矿竖井井筒的安全状态判别函数。

利用留一交叉验证法对建立的安全状态判别函数进行训练验证, 前 16 组训练样本数据判别结果全部正确, 并对剩余的 4 组检验样本进行判别验证, 判别结果与各煤矿竖井实际安全状态吻合, 说明本文建立的煤矿竖井安全状态判别模型预测效果较佳。

#### 参考文献

- [1] 刘环宇, 李晓, 曾钱帮, 黄福昌, 陈祥军. 兖州矿区立井井筒非采动破裂的非线性预测与判别方法[J]. 工程地质学报, 2005, 13(2): 231-235.
- [2] 刘环宇. 厚冲积层立井井筒破坏的发生机理及防治技术研究[D]: [博士学位论文]. 南京: 河海大学, 2005.
- [3] 骆念海, 杨维好. 井壁垂直附加力的影响因素分析[J]. 煤炭科学技术, 2000, 28(12): 41-43.
- [4] 经来盛. 表土沉降对井壁破裂的影响及防破裂措施的研究[J]. 煤炭学报, 2001, 38(1): 49-53.
- [5] 李文平, 于双忠. 深厚表土中煤矿立井非采动破裂的研究[J]. 工程地质学报, 1995, 3(1): 45-55.
- [6] 黎锦贤, 胡千庭. 主成分分析法在煤矿安全评价中的应用[J]. 矿业安全与环保, 2007, 36(5): 71-73+76+90.
- [7] 金洪伟, 周捷, 杨守国, 高巧红. 基于主成分分析法的煤矿瓦斯涌出量预测[J]. 煤炭技术, 2019, 38(4): 125-128.
- [8] 丁坤, 刘振飞, 高列, 等. 基于主成分分析和马氏距离的光伏系统健康状态研究[J]. 可再生能源, 2017, 35(1): 1-7.
- [9] 刘庆军, 陈坤, 刘晓光. 煤与瓦斯突出预测 PCA——距离判别法研究[J]. 中国煤炭, 2016, 42(10): 97-101.
- [10] 张国英, 王娜娜, 张润生, 马兵胜. 基于主成分分析的 BP 神经网络在岩性识别中的应用[J]. 北京石油化工学院学报, 2008, 16(3): 43-46.
- [11] 琚棋定, 胡友彪, 张淑莹. 基于主成分分析与贝叶斯判别法的矿井突水水源识别方法研究[J]. 煤炭工程, 2018, 50(12): 90-94.
- [12] 胡友彪, 邢世平, 张淑莹. 基于可拓模型判别矿井突水水源[J]. 安徽理工大学学报(自然科学版), 2017, 37(6): 34-40.

- 
- [13] 马济国, 姜春露, 朱赛君, 谢毫, 毕波, 郑刘根. 基于主成分分析的潘谢矿区突水水源 Fisher 判别模型[J]. 煤炭技术, 2020, 39(9): 132-134.
- [14] 郭亚奔, 杨金林, 陈忠. 边坡稳定性预测的 DDA 模型及应用[J]. 有色金属, 2011, 63(3): 70-72.
- [15] 邵良杉, 张宇. 煤矿立井井筒非采动破裂预测[J]. 煤炭学报, 2009, 34(2): 184-186.