

基于统计模式识别的压力管道损伤检测

侯小强¹, 朱军伟², 任建民²

(1. 甘肃建筑职业技术学院, 兰州 730050; 2. 兰州交通大学 土木工程学院, 兰州 730070)

摘要: 基于统计模式识别理论, 采用 AMRA 时序分析方法, 通过长自回归模型计算残差法和最小二乘法的结合来估计模型参数, 从而建立了系统模型。运用均值控制图的方法对压力管道的振动信息特征进行提取、选择, 从而有效地判别压力管道的无损与损伤问题。数值模拟结果表明, 基于统计模式识别的均值控制图的结构异常检验方法, 能够准确诊断检测结构的损伤状态, 而且对于损伤程度和损伤位置都有很强的敏感性。

关键词: 统计模型; AMRA 时序分析; 特征值; 均值控制图; 损伤检测

中图分类号: TV 732.4 **文献标志码:** A **文章编号:** 1672-1683(2015)04-0813-04

Damage detection of pressure pipe based on statistical pattern recognition theory

HOU Xiaoliang¹, ZHU Junwei², REN Jianmin²

(1. Gansu Construction Vocational Technical College, Lanzhou 730050, China;

2. School of Civil Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: Based on the statistical pattern recognition theory, the AMRA timing analysis method combined with the long autoregressive model residual method and the least squares method were used to estimate the model parameters and then to develop the system model. The mean control chart method was used to extract and select the vibration information and feature of the pressure pipe, therefore the damage of the pressure pipe can be identified effectively. The simulation results showed that structural abnormality test method of the mean control chart, which is based on the statistical pattern recognition theory, can diagnose the structural damage accurately, and is sensitive to the damage degree and location.

Key words: statistical model; AMRA timing analysis; eigen value; mean control chart; damage detection

统计模式进行损伤识别^[1-2]具有剔除偶然环境因素造成的组内差异, 将组内差异最小化, 组间差异最大化, 从而能够识别结构轻微损伤^[3-4], 具有较高的灵敏度。

其关键步骤就是要找到能够明显区别正常统计模式和异常统计模式的指标。文章采用 AMRA 时序分析^[5-6]方法建立系统模型, 采用主成分分析方法提取特征参数, 并运用均值控制图的方法对泵站压力管道的健康状态进行判断。均值控制图方法^[7-8]将监测结果自图中表现出来, 可视化性较强, 准确度较高, 而且对于损伤程度和损伤位置都有很强的敏感性

1 建立统计模型

文中基于统计模式来识别压力管道的损伤问题的研究中, 采用 AMRA 时序分析方法建立系统模型, 包括 ARMA 时间序列模型的定阶、模型参数提取、特征参数的减缩和差

异指标的构造等。采用主成分分析方法提取特征参数, 并提出了均值控制图的识别分类判别方法。时间序列是指一组按照时间、空间或是其他方式排列成的有序随机数据。ARMA 时间序列是指采用 ARMA 参数观测输出到的振动响应数据进行分析和处理的一种统计学的数学方法。所以 ARMA 时序模型建模的结构损伤诊断是基于统计模式识别^[9-10]的损伤诊断方法。

1.1 AMRA 模型

假设通过监测模拟, 在一定采样频率下所测得一系列按时间先后顺序排列的监测数据, 这些离散有序的随机数据就组成一经典的观测时间序列, 这个时序服从正态分布, 表示成时间的序列为

$$\{x_t\} = \{x_1, x_2, x_3, \dots, x_n\} \quad t = 1, 2, \dots, n \quad (1)$$

其中: 对于平稳、正态、零均值为零的序列 $\{x_t\}$, 若 x_t 的取值不仅与前步的各个取值 $x_{t-1}, x_{t-2}, \dots, x_{t-n}$ 有关, 还与干

收稿日期: 2014-07-22 修回日期: 2015-05-08 网络出版时间: 2015-07-24

网络出版地址: <http://www.cnki.net/kcms/detail/13.1334.TV.20150724.0919.004.html>

基金项目: 甘肃省科技支撑计划研究项目(1304FKA055); 甘肃省高等学校科研项目(2013B124); 国家电网甘肃电力科技项目(5227011600D9)

作者简介: 侯小强(1977-), 男, 甘肃正宁人, 高级工程师, 主要从事滑坡治理研究和教学方面工作。E-mail: houxiq@foxmail.com

扰值 α_i , 以及前 m 步的 $\alpha_{i-1}, \alpha_{i-2}, \dots, \alpha_{i-m}$ 各个干扰值有关, 则引入多远线性回归思想, 可以把经典 AMRA 模型^[1]表示为下式:

$$x_i = \varphi_1 x_{i-1} + \varphi_2 x_{i-2} + \dots + \varphi_n x_{i-n} + \alpha_i - \theta_1 \alpha_{i-1} - \theta_2 \alpha_{i-2} - \dots - \theta_m \alpha_{i-m} \quad (2)$$

移项整理为

$$x_i - \sum_{i=1}^n \varphi_i x_{i-i} = \alpha - \sum_{j=1}^m \theta_j \alpha_{i-j}, \alpha_i \sim NID(0, \sigma_\alpha^2) \quad (3)$$

式中: $\varphi_i (i=1, 2, 3, \dots, n)$ 是自回归参数; $\theta_j (j=1, 2, 3, \dots, m)$ 为滑动平均参数; α_i 是残差, 若式(3)能够正确反映结构的时序规律时, 则 $\{\alpha_i\}$ 是白噪声, 其均值为零, 方差为 σ_α^2 。式(3)等式左边称为 n 阶自回归部分, 右边称为 m 阶滑动平均部分, 所以方程(3)即是 n 阶自回归 m 阶滑动平均模型, 简记为 $ARMA(n, m)$ 时序模型。

1.2 AMRA 模型定阶

ARMA 模型定阶就是确定模型的阶次 n, m , 也称模型的适用性检验。ARMA 时序模型的传统定阶方法一般有三种: 白噪声检验准则、残差平方和(残差方差)检验准则以及 Akaike 信息准则^[11]。Akaike 信息准则以残平方和检验准则为基础, 在此之上填补了前两种方法模型阶次高、参数增加所带来的误差影响。同时 Akaike 信息准则还具有计算简单, 便于操作和实现, 效果显著的优点而在实际中得到较为广泛应用。Akaike 信息准则主要用到 AIC 准则。

AIC 准则如下:

$$AIC(p) = N \ln \sigma_\alpha^2 + 2p \quad (4)$$

式中: N 表示样本的数据长度; p 表示建立的统计模型个数,

$$p = n + m. \sigma_\alpha^2 \text{ 为模型残差的方差, 表达式为: } \sigma_\alpha^2 = \frac{1}{N - n - 2} \cdot$$

$\sum_{i=n+1}^N \alpha_i^2 = \frac{s}{N - n - 2}, s = \sum_{i=n+1}^N \alpha_i^2$ 。对于一个新建立的 $ARMA(n, m)$ 时序模型, 增加模型的拟合阶数 p 的值, 那么残差的方差 σ_α^2 就会降低。因此 $AIC(p)$ 的求解值一定会有极小值, 使得 $n = p$, 此时的 n 就是 $ARMA(n, m)$ 模型的最佳阶数。

1.3 AMRA 模型参数估计

本文首先采用长自回归模型^[12]计算时序残差, 再采用最小二乘法估计模型参数。长自回归计算残差法避免了繁杂的计算过程, 建模速度大大提高, 再采用最小二乘法估计参数, 提高了精度。两种方法的结合使用既能够简化复杂的计算工程, 又能保证参数的估计精度。

通过长自回归模型计算残差序列 $\{\alpha_i\}$, 采用线性回归中的最小二乘法估计模型参数, 将 $\{\alpha_i\}$ 与 $\{x_i\}$ 代入 $ARMA(n, m)$ 模型来估计其模型参数, 可得如下方程:

$$y = X\beta + \alpha \quad (5)$$

$$\text{其中: } y = \begin{bmatrix} x_{p+n} & x_{p+n+1} & \dots & x_n \end{bmatrix}^T, \alpha = \begin{bmatrix} \alpha_{p+n} & \alpha_{p+n+1} & \dots & \alpha_n \end{bmatrix}^T$$

$$\beta = \begin{bmatrix} \varphi_1 & \varphi_2 & \dots & \varphi_n - \theta_1 - \theta_2 & \dots & -\theta_m \end{bmatrix}^T$$

$$X = \begin{bmatrix} x_{p+n-1} & x_{p+n-2} & \dots & x_{p+n} & \alpha_{p+n-1} & \alpha_{p+n-2} & \dots & \alpha_{p+n-m} \\ x_{p+n} & x_{p+n-1} & \dots & x_{p+n+1} & \alpha_{p+n} & \alpha_{p+n-1} & \dots & \alpha_{p+n-1-m} \\ \dots & \dots \\ x_{N-1} & x_{N-2} & \dots & x_{N-n} & \alpha_{N-1} & \alpha_{N-2} & \dots & \alpha_{N-m} \end{bmatrix}$$

则, 估计最小二乘 β 为

$$\beta = (X^T X)^{-1} X^T y \quad (6)$$

由上可见, 长自回归计算残差法将 ARMA 模型参数估计的非线性过程转化为 AR 模型参数的线性估计过程。在概念上简单明了, 避开了 ARMA 建模中大量的繁杂计算, 建模速度大为提高。而且, 两次建模(先建 $AR(p)$ 模型, 再建 $ARMA(n, m)$ 模型)可使用同一种参数估计方法(最小二乘法)。

2 特征提取和选择

特征提取的实质是从已有的模式向量(初步特征)中提取出具有代表性的有效成分构成新的特征向量(精细特征), 从而使得识别工作量大为减小, 提高效率和精度。文章采用主成分分析法进行特征选择, 重新构造统计模型识别的物理特征参数。

2.1 主成分分析法

主成分分析法^[13]是构造原变量适当的线性组合, 以产生一组互不相关的新变量, 从中选出少数几个新变量并使它们含有尽可能多的原变量带有的信息, 从而采用这组新变量代替原变量分析解决问题成为可能。当所研究的问题确定之后, 变量中所含“信息”的大小通常用该变最的方差或样本方差来度量。

设结构的某一状态由 n 维模式向量 X 描述, $X = [x_1 \ x_2 \ \dots \ x_n]^T (n \geq 2)$, 其中各 x_i 之间相关, 通过一系列数学方法, 对向量 X 进行降维变换, 变换矩阵为 A , 并得到低维向量 $Y = [y_1 \ y_2 \ \dots \ y_m]^T (m < n)$, 即: $Y_{m \times 1} = A_{m \times n} X_{n \times 1}$

这样, 对模式向量 X 进行了降维处理, 得到了低维、各元素相互独立的特征向量 Y , 而且 Y 中包含有原模式向量中大部分信息, 因此也同时实现了原数据的信息凝聚。

2.2 基于 ARMA 模型的特征提取

在结构损伤检测中, 首先将所采集的数据分成等长的数据样本 $\{x_i\}_j (j=1, 2, \dots, k)$, 每一个样本作为一个时序建立 $ARMA(n, m)$ 模型。时序模型的参数 $\varphi_i (i=1, 2, 3, \dots, n)$, $\theta_j (j=1, 2, 3, \dots, m)$ 、模型残差方差 σ_α^2 均可组成表征结构状态的模式向量。结构的检测积累了大量数据, 建立时序模型将大量数据所蕴含的信息凝聚成为少数几个模型参数, 然后对模型参数进行主成分分析进一步得到特征向量, 从而一步步实现对原始监测数据的降维处理和信息凝聚。

设自回归参数^[14] $\varphi_i (i=1, 2, 3, \dots, n)$ 构成了结构某一参考状态所对应的模式向量 $X = [x_1 \ x_2 \ \dots \ x_n]^T$, 并且得到 X 的 k 个训练样本 $x_j (j=1, 2, 3, \dots, k)$, 采用正交变换, 将原来的 x_1, x_2, \dots, x_n 变成相互独立的 y_1, y_2, \dots, y_n , 取前 m 个主成分 ($m < n$) 以实现数据的降维处理, 见表 1。

3 模型检验

均值控制图^[7]是一种统计学上的假设检验方法, 均值控制图的理论依据是正态总体均值的置信区间, 根据样本构造置信上限与置信下限, 从而作出均值控制图。其需要对标准样本的整体确定出三个参数, 即上控制线(UCL)、中心线(CL)、下控制线(LCL)。

采用均值控制图识别结构是否损伤的依据是根据控制

图中管道的振动信息特征所描点的出界个数来判别, 每个样本平均值都画在图上, 如果均值落在控制线之间, 则表明压力管道正常没有损伤, 否则有损伤。文章采用 3σ 控制均值

图。均值控制图的三个控制参数可以表示为

$$CL_{\bar{x}} = \bar{\bar{x}}, UCL_{\bar{x}} = \bar{\bar{x}} + \frac{3\bar{S}}{c\sqrt{n}}, LCL_{\bar{x}} = \bar{\bar{x}} - \frac{3\bar{S}}{c\sqrt{n}} \quad (8)$$

表 1 数据的降维处理

Tab. 1 Data dimension reduction

序号	原始数据	主成分	提取前 m 个主成分
1	$x_{11}, x_{21}, \dots, x_{m1}, \dots, x_{n1}$	$y_{11}, y_{21}, \dots, y_{m1}, \dots, y_{n1}$	$y_{11}, y_{21}, \dots, y_{m1}$
2	$x_{12}, x_{22}, \dots, x_{m2}, \dots, x_{n2}$	$y_{12}, y_{22}, \dots, y_{m2}, \dots, y_{n2}$	$y_{12}, y_{22}, \dots, y_{m2}$
...
k	$x_{1k}, x_{2k}, \dots, x_{mk}, \dots, x_{nk}$	$y_{1k}, y_{2k}, \dots, y_{mk}, \dots, y_{nk}$	$y_{1k}, y_{2k}, \dots, y_{mk}$

式中: n 为样本总体量; c 为关于样本容量的常数; $\bar{\bar{x}}$ 为样本均值的平均值, \bar{S} 为样本方差的平均值。均值控制图所表示的正常状态与异常状态的关系见图 1。

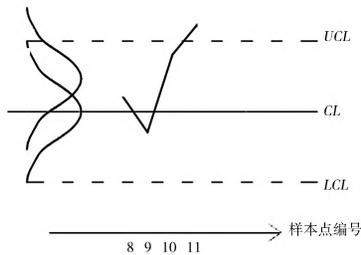


图 1 均值控制图识别方式

Fig. 1 The recognition method of mean control chart

4 数值模拟验证

数值实验模型以某泵站 1 号压力管道为例。1 号压力管道分为左进水口支管、右进水口支管和出水口总管, 其管道平面尺寸图形见图 2。图中管道 1 部分: 内径 1.40 m, 壁厚 0.12 m; \circ 部分: 左端内径 1.40 m, 壁厚 0.12 m, 右端内径 1.00 m, 壁厚 0.12 m; \gg 部分: 转弯处外径 2.50 m, 内径 1.50 m, 轴线半径 2.00 m, 壁厚 0.12 m; $\frac{1}{4}$ 部分: 内径 0.80 m, 壁厚 0.12 m。钢管可简化为均质材料, 密度为 7.85 g/cm^3 , 弹性模量 $2.06 \times 10^5 \text{ MPa}$, 泊松比 0.25。

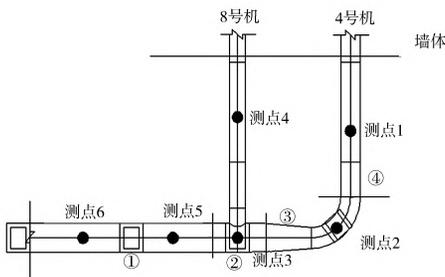


图 2 数值模拟样本测点号

Fig. 2 Measuring points of samples in numerical simulation

4.1 建立识别模型

依据统计模式识别的 $ARMA(n, m)$ 模型定阶方法, 对 1 至 6 测点分别进行模型定阶。采用 $AIC(p) = N \ln \Phi_a^2 + 2p$ 准则, 求得函数 $AIP(p)$ 的极小值 $p = 8$, 即 $n = p = 8$ 。定阶结果表明, 模型自回归部分阶数能够稳定在 8 阶, 滑动平均部分阶数有一定的离散性, 不能稳定。所定阶数为 $ARMA(8, 7)$, $ARMA(n, m)$ 模型定阶结果见表 2。

征参数。对建立起来的 $ARMA(8, 7)$ 模型, 经过主成分分析, 第一阶主成分方差即占总方差的 85% 以上, 所以可选取第一阶主成分指标代替原来多参数指标。

最后绘出测点 1 至 6 控制图的上控制线 UCL 、中心线 CL 、下控制线 LCL 三个控制参数。

表 2 $ARMA(n, m)$ 模型定阶结果

Tab. 2 $ARMA(n, m)$ model order results

编号	测点 1		测点 2		测点 3		测点 4		测点 5		测点 6	
	n	m										
1	8	7	8	7	8	5	8	6	8	7	8	7
2	8	7	8	6	8	7	8	7	8	5	8	6

4.2 识别结果

在有限元模拟计算中, 如果有单元损伤, 那么该单元抗弯刚度就会必然降低。结构中特定部分的质量和刚度损失而引起的模态参数的变化, 都将在模态模拟中有所体现, 当系统的模态模拟结果与完好结构系统的模态值之间出现差异时, 就表明结构出现了一定的损伤, 进而可以确定损伤的位置及程度。作者先用有限元仿真分析模拟的方法分别计算得到结构正常状态和待检测状态下响应, 进行识别。模拟结果表明: 损伤位于测点 3 处, 抗弯刚度折减 30%。

以测点 1 和测点 3 为例, 验证基于统计模式损伤识别的均值控制图方法, 图 3、图 4 分别是测点 1、测点 3 的正常状态控制图和待检测状态控制图。

从图 3 均值控制图中可见, 测点 1 的均值控制图无论是正常状态还是待检测状态的控制点走在上控制线和下控制线之内, 没有控制点越界。在图 4 的测点 3 的均值控制图中, 正常状态控制点均在上下控制线之间, 没有点出界, 测点 3 的待检测状态的控制点从图 4(b) 中可以得知有 7 个样本点出界, 表明此处有损伤。其结果也符合模拟假设中的 3 测点损伤位置这一假设。经过模型的验证, 从中可以说明这种均值控制图方法在统计模式损伤识别中的应用是有较高实际应用价值的, 可以表现出基于统计模式损伤识别的均值控制图方法对检验结构损伤位置及其程度具有较强的敏感性。

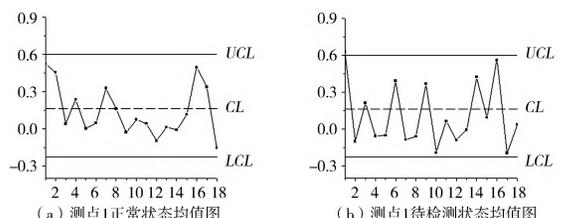


图 3 测点 1 均值控制图

Fig. 3 The mean control chart of measuring point 1

基于前面章节提到的主成分分析的方法提取和缩减特

根据正态分布与数理统计的知识,进行建筑物结构诊断是概率统计应用的一个重要方面,特别是控制均值图,不仅用于土木工程结构质量诊断,还广泛用于企业全面质量管理,金融风险分析与管理等各个方面。

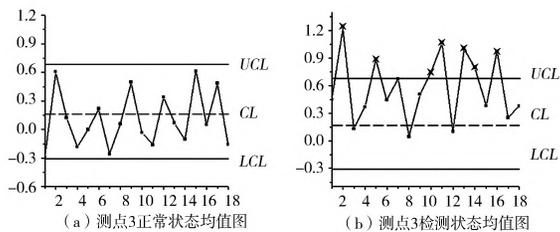


图 4 测点 3 均值控制图

Fig. 4 The mean control chart of measuring point 3

5 小结

(1) 检测输出的数据通过 AMRA 时序模型的转换计算确定 ARMA(n, m) 模型的最佳阶数;通过长自回归模型计算残差法和最小二乘法的结合来估计模型参数;采用主成分分析法缩减参数,降低模型计算的复杂程度,又能够同时保证参数的原有信息含量。

(2) 均值控制图作为判别指标的一种构造方法,相比其他方法有着很大的优势。建立正常状态和待检测状态的两种统计识别模式,进行描点比对,能直观准确地识别结构的健康状况。通过识别结果分析,1 号压力管道测点 1 处和测点 3 处存在结构损伤。

(3) 基于统计模式识别的结构监测在土木工程结构监测领域有很广的应用前景,本文提供的检测方法同时可以为桥梁和其他建筑提供参考。

参考文献 (References):

[1] 陈志为. 基于统计模式识别技术的结构异常检验[D]. 福建: 福州大学, 2005. (CHEN Zhi wei. The test of structural abnormalities based on statistical pattern recognition techniques [D]. Fuzhou: Fuzhou University, 2005. (in Chinese))

[2] 刘毅, 李爱群. 基于结构响应的损伤诊断方法及其应用[J]. 东南大学学报: 自然科学版, 2010, 40(4): 810-815. (LIU Yi, LI Ai qiong. structural response based damage detection method and its application [J]. Southeast University: Natural Science Edition, 2010, 40(4): 810-815. (in Chinese))

[3] 黄斌, 史文海. 基于统计模型的结构损伤识别[J]. 工程力学, 2006, 23(12): 47-54. (HU ANG Bin, SHI Wen hai. Damage identification of structures using stochastic models. Engineering me-

chanics, 2006, 23(12): 47-54. (in Chinese))

[4] Birkenheuer G, Brinkmann A, Hogqvist M, et al. Infrastructure federation through virtualized delegation of resources and services[J]. J Grid Computing, 2011(9): 355-377.

[5] 何书元. 应用时间序列分析[M]. 北京: 北京大学出版社, 2003. (HE Shuyuan. Applied Time Series Analysis [M]. Beijing: Peking University Press, 2003. (in Chinese))

[6] 朱军华. 基于时间序列分析与高阶统计矩的结构损伤检测[J]. 东南大学学报: 自然科学版, 2012, 42(1): 7-13. (ZHU Jun hua. Damage detection based on time series analysis and higher statistical moments [J]. Southeast University: Natural Science Edition, 2012, 42(1): 7-13. (in Chinese))

[7] 周丙常, 师义民, 于蕾. 有偏总体的均值控制图[J]. 昆明理工大学学报: 理工版, 2005, 30(3): 123-126. (ZHOU Bing chang, SHI Yi min, YU Lei. Biased population mean control chart [J]. Kunming University of Technology: Science & Engineering Edition, 2005, 30(3): 123-126. (in Chinese))

[8] Sergios Theodoridis, Konstantinos Koutroumbas. 模式识别(第二版)[M]. 李晶皎译. 北京: 电子工业出版社, 2004. (Sergios Theodoridis, Konstantinos Koutroumbas. Pattern Recognition (second edition) [M]. Jing Jiao Beijing: Electronic Industry Press, 2004 (in Chinese))

[9] Abraham Wald. Contributions to the theory of statistical estimation and testing of hypotheses[J]. Annals of mathematical Statistics, 10: 299-326, 1939.

[10] Wilbur H Highleyman. Linear decision function, with application to pattern recognition[J]. Proceedings of the IRE, 50: 1501-1504, 1962.

[11] 接标, 刘冠晓, 冯乔生. 统计模式识别的研究[J]. 云南师范大学学报, 2005, 25(6): 19-21. (JIE Biao, LIU Guan xiao, FENG Qiao sheng. Statistical pattern recognition research [J]. Yunnan Normal University, 2005, 25(6): 19-21. (in Chinese))

[12] 童佳威. 基于主成分的特征提取方法研究[D]. 杭州: 浙江树人大学, 2011. (TONG Jia wei. Study of Structural Abnormalities Based on Statistical Pattern Recognition Techniques [D]. Hangzhou: Zhejiang Shuren University, 2011. (in Chinese))

[13] 王真, 杜利珍, 程远胜. AR 参数灵敏度及结构损伤控制图识别[J]. 振动、测试与诊断, 2010, 30(4): 394-399. (WANG Zhen, DU Li zhen, CHENG Yuan sheng. AR parameter sensitivity and structural damage control pattern recognition [J]. Vibration, Measurement & Diagnosis, 2010, 30(4): 394-399. (in Chinese))

[14] 王真, 程远胜. 基于时间序列模型自回归系数灵敏度分析的结构损伤识别方法[J]. 工程力学, 2008, 25(10): 38-43. (WANG Zhen, CHENG Yuan sheng. Structural damage identification based on sensitivity analysis of autoregressive coefficients of time series models [J]. Engineering mechanics, 2008, 25(10): 38-43. (in Chinese))

(上接第 807 页)

[16] 於孝春, 贾朋美, 张兴. 基于模糊 Bow-tie 模型的城镇燃气管道泄漏定量风险评价[J]. 天然气工业, 2013, 33(7): 134-139. (YU Xiao chun, JIA Peng mei, ZHANG Xing. Quantitative risk assessment of urban gas pipelines based on the fuzzy Bow-tie model [J]. Natural Gas Industry, 2013, 33(7): 134-139. (in Chinese))

[17] Wang Y M, Yang J B, Xu D L, et al. On the centroids of fuzzy numbers[J]. Fuzzy sets and systems, 2006, 157(7): 919-926.

[18] 吕智颖, 黄天民, 靳凤霞. 模糊多属性序决策的冗余指标的

消除策略[J]. 数学的实践与认识, 2013, 43(10): 173-181. (LV Zhi ying, HU ANG Tian ming, JIN Feng xia. Fuzzy multiple attribute lattice decision making method based on the elimination of redundant similarity index [J]. Mathematics in Practice and Theory, 2013, 43(10): 173-181. (in Chinese))

[19] Cher-Tung C, Ching-Torng L, Sue-Fn H. A fuzzy approach for supplier evaluation and selection in supply chain management [J]. International Journal of Production Economics, 2006, 102(2): 289-301.