

# 依据主成分和协整性的大坝变形奇异诊断\*

杨光<sup>1,2</sup>, 李姝昱<sup>3</sup>, 孙锦<sup>4</sup>

(1. 华北水利水电大学水利学院 郑州, 450046)

(2. 河海大学水文水资源与水利工程科学国家重点实验室 南京, 210098)

(3. 黄河水利科学研究院 郑州, 450003)

(4. 华北水利水电大学测绘与地理信息学院 郑州, 450046)

**摘要** 针对常规方法对大坝变形原位监测数据中奇异成分的诊断效率较低问题, 综合应用主成分分析(principal component analysis, 简称 PCA)和协整分析(co-integration analysis, 简称 CA), 提出一种新方法。首先, 基于 PCA, 构建平方预测误差(squared prediction error, 简称 SPE)统计量, 结合假设检验, 提出奇异成分辨识准则; 其次, 依据 CA, 运用拓展的迪基-福勒(augmented Dickey-Fuller, 简称 ADF)检验和逐步回归法, 建立奇异成分似然估计模型; 最后, 通过工程实例分析, 检验方法的有效性。结果表明: PCA、拉依达、狄克松和  $t$  准则分别可辨识出相对误差为 3.81%, 7.61%, 7.61% 和 5.08% 的孤立型奇异; CA 模型对斑点型奇异的估计精度最高, 其次是统计模型, 自回归模型最差, 复相关系数分别为 0.994 5, 0.871 5 和 0.743 2。与常规方法相比, PCA-CA 方法性能有较大提升, 可为大坝变形奇异诊断提供有效的途径。

**关键词** 大坝; 变形原位监测; 奇异诊断; 主成分; 协整性  
**中图分类号** TV698.1

## 引言

变形是大坝服役性态变化的综合反映, 是衡量结构安全与否的重要标志。科学地分析变形原位监测信息, 不仅是现行规范的要求, 而且是监控大坝安全的有效手段<sup>[1-2]</sup>。受多种因素的干扰, 原位监测资料中常包含奇异成分, 表现为测值在某时刻或某时段的异常突跳, 即孤立型或斑点型奇异成分。若未有效诊断出监测资料中的奇异成分, 极可能影响大坝安全监控结论的客观性, 造成虚假报警或将危险状况遗漏。尤其对于运行多年的老坝, 监测仪器的性能和稳定性较差, 此问题尤为突出。因此, 提出科学的大坝变形原位监测数据奇异成分诊断方法, 具有重要的理论意义和应用价值。

大坝变形原位监测数据奇异成分诊断包括辨识和估计 2 个环节。常规的辨识方法有<sup>[3-4]</sup>: ①过程线法, 通过绘制变形过程线, 直观辨识奇异成分; ②假设检验法, 如拉依达检验、狄克松检验。传统的估计

方法包括<sup>[3-4]</sup>: ①忽略法, 若无法还原或难度较大, 可直接忽略; ②似然估计法, 如临近插值、三次 Hermite 插值、统计模型及自回归模型。总体来看, 常规诊断方法大多以变形时间序列为分析对象, 无法有效地捕捉多测点间隐含的变形关系, 且主观性较强, 易误诊。

相较于 1 维时间序列, 多维数据序列具有信息量丰富、自由度高及稳定性强的优势, 包含了隐匿、新颖及有潜在价值的信息<sup>[5]</sup>。PCA<sup>[6-9]</sup>和 CA<sup>[10-12]</sup>是两种多维数据特征挖掘方法。PCA 可将多维数据序列投影到不同方向上, 得到主要规律和无法解释成分, 其中无法解释成分包含了奇异数据、噪声等。对于大坝上临近的监测点, 变形监测序列存在关联性, 若未出现超标准洪水、极端天气等非常规状况, 某监测点变形数据异常, 其他临近点的数据未见异常, 则可判定该监测点的变形数据奇异。CA 刻画了多维数据序列的长期关系, 若未发生非常规状况, 多测点变形监测序列的协整关系一般不会明显改

\* 国家自然科学基金资助项目(52109155); 水文水资源与水利工程科学国家重点实验室“一带一路”水与可持续发展科技基金资助项目(2019492211); 华北水利水电大学高层次人才启动资助项目(202005002, 202005017); 黄河水利科学研究院基本科研业务费专项资助项目(HKY-JBYW-2021-10)

收稿日期: 2020-11-29; 修回日期: 2021-09-08

变,据此可实现对奇异成分的估计。

综上所述,笔者利用多维数据的优势,基于 PCA,研究奇异成分辨识准则;依据 CA,探究奇异成分似然估计模型;在此基础上,结合实际工程,检验 PCA-CA 方法的有效性。

## 1 诊断方法

### 1.1 奇异成分辨识

依据 2 维数据格式,建立大坝变形原位监测数据集,即

$$X = [x_1 \ x_2 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_m] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1i} & \dots & x_{1j} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2i} & \dots & x_{2j} & \dots & x_{2m} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ x_{k1} & x_{k2} & \dots & x_{ki} & \dots & x_{kj} & \dots & x_{km} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{ni} & \dots & x_{nj} & \dots & x_{nm} \end{bmatrix} \quad (1)$$

其中: $m$ 为监测点总数; $n$ 为监测时长; $x_i, x_j$ 分别为测点  $i$  和  $j$  的监测序列; $x_{ki}, x_{kj}$ 分别为测点  $i$  和  $j$  在  $k$  时刻的监测值。

PCA 辨识步骤如下。

1) 分析环境量监测资料,若未发生非常规状况,则执行后续步骤。

2) 对矩阵  $X$  进行标准化,得到矩阵  $\bar{X} = [\bar{x}_1 \ \bar{x}_2 \ \dots \ \bar{x}_i \ \dots \ \bar{x}_j \ \dots \ \bar{x}_m]$ ,  $\bar{x}_i$  为标准化的  $x_i$ ,  $\bar{x}_j$  为标准化的  $x_j$ 。相关系数矩阵  $R$  为

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1j} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2j} & \dots & r_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{i1} & r_{i2} & \dots & r_{ij} & \dots & r_{im} \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mj} & \dots & r_{mm} \end{bmatrix} \quad (2)$$

其中: $r_{ij}$ 为  $\bar{x}_i$  和  $\bar{x}_j$  的相关系数,计算公式为  $r_{ij} = \frac{1}{n-1} \sum_{k=1}^n \bar{x}_{ki} \bar{x}_{kj} \quad (i, j = 1, 2, \dots, m)$ ;  $\bar{x}_{ki}$  和  $\bar{x}_{kj}$  为标准化的  $x_{ki}$  和  $x_{kj}$ 。

3) 计算相关系数矩阵  $R$  的特征向量矩阵  $P$  和特征值矩阵  $\lambda$ , 计算公式为

$$(R - \lambda I)P = 0 \quad (3)$$

其中: $0$ 为  $m$  阶零矩阵; $I$ 为  $m$  阶单位矩阵。

式(3)中的  $P$  和  $\lambda$  分别表示为

$$P = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_i \\ \vdots \\ p_m \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1j} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2j} & \dots & p_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ p_{i1} & p_{i2} & \dots & p_{ij} & \dots & p_{im} \\ \vdots & \vdots & & \vdots & & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mj} & \dots & p_{mm} \end{bmatrix} \quad (4)$$

$$\lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_j & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & \lambda_m \end{bmatrix} \quad (5)$$

其中: $\lambda_i$ 为矩阵  $R$  的第  $i$  个特征值,满足  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ ;  $p_i$  为  $\lambda_i$  对应的特征向量。

4) 建立  $n \times m$  矩阵  $T$ , 记为  $(t_1, t_2, \dots, t_j, \dots, t_m)$

$$T = \bar{X}P^T \quad (6)$$

5) 计算第  $j$  项  $t_j (j = 1, 2, \dots, m)$  对原始序列  $X$  的解释能力  $e_j$

$$e_j = \left( \lambda_j / \sum_{i=1}^m \lambda_i \right) \times 100\% \quad (7)$$

由式(7)可知:①前  $j$  项对  $X$  的累积解释能力为  $\sum_{i=1}^j e_i$ , 所有  $m$  项的累积解释能力为 1; ②  $e_1 \geq e_2 \geq \dots \geq e_m$ , 即  $t_1 \rightarrow t_m$  对  $X$  的解释能力逐渐变弱; ③  $e_1$  越大,  $X$  的主要规律越明显。因此,为保证 PCA 辨识精度,要求  $e_1$  尽可能大。

6) 建立 SPE 统计量。由于  $P$  为正交矩阵,满足

$$P^T = P^{-1} \quad (8)$$

则式(6)可变换为

$$\bar{X} = TP \quad (9)$$

矩阵  $\bar{X}$  的主要规律用  $n \times m$  矩阵  $\hat{D}$  表示, 记为  $(\hat{d}_1, \hat{d}_2, \dots, \hat{d}_j, \dots, \hat{d}_m)$ 。  $\hat{D}$  计算公式为

$$\hat{D} = t_1 p_1 \quad (10)$$

$k$  时刻 SPE 统计量记为  $SPE_k$ , 计算公式为

$$SPE_k = \sum_{j=1}^m (\bar{x}_{kj} - \hat{d}_{kj})^2 \quad (11)$$

其中: $\bar{x}_{kj}$  和  $\hat{d}_{kj}$  分别为矩阵  $\bar{X}$  和  $\hat{D}$  第  $k$  行第  $j$  列的元素。

7) 依据假设检验,在显著性水平  $\alpha$  下,建立 SPE 控制限<sup>[13-14]</sup>, 记为  $SPE_\alpha$ , 即

$$SPE_\alpha = \theta_1 \left[ \frac{C_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1} \right]^{\frac{1}{h_0}} \quad (12)$$

其中： $\theta_l = \sum_{j=2}^m \lambda_j^l$  ( $l=1, 2, 3$ )； $h_0 = 1 - 2\theta_1\theta_3/3\theta_2$ ；

$C_\alpha$  为  $\alpha$  所对应的正态分布函数值。

假设检验中  $\alpha$  常取为 0.05 或 0.01，即接受原假设时，正确的概率为 95% 或 99%。

8) 图 1 为 SPE 统计与控制限的关系。若  $SPE_k$  未超过  $SPE_\alpha$ ，如图 1(a) 所示，则不存在奇异成分；若  $SPE_k$  在某时段超过  $SPE_\alpha$ ，如图 1(b) 所示，则为斑点型奇异；若  $SPE_k$  在某时刻超过  $SPE_\alpha$ ，如图 1(c) 所示，则为孤立型奇异。若奇异，则执行步骤 9。

9) 测点  $j$  对  $k$  时刻  $SPE_k$  超过  $SPE_\alpha$  的贡献度用  $CSPE_{kj}$  统计量表征，即

$$SPE_k = \sum_{j=1}^m (\bar{x}_{kj} - \hat{d}_{kj})^2 = \sum_{j=1}^m CSPE_{kj} \quad (13)$$

$CSPE_{kj}$  越大，则测点  $j$  对  $k$  时刻  $SPE_k$  超限的贡献度越大，而  $CSPE_{kj}$  最大处，即为存在奇异成分的测点。

### 1.2 奇异成分似然估计

假设  $m$  个变形监测序列中有  $f$  个存在奇异成

分，记为  $x'_1 \cdots x'_h \cdots x'_f$ ，剩余  $p = m - f$  个不存在奇异成分，记为  $x''_1 \cdots x''_q \cdots x''_p$ 。CA 似然估计步骤如下。

1) 建立多测点变形原位监测序列的协整模型，即

$$x'_h = \sum_{g=1}^{a_1} c_{1g} x''_1{}^g + \cdots + \sum_{g=1}^{a_q} c_{qg} x''_q{}^g + \cdots + \sum_{g=1}^{a_p} c_{pg} x''_p{}^g + \epsilon \quad (14)$$

其中： $c_{qg}$  为  $x''_q$  的系数； $a_q$  为  $x''_q$  的最高次数，依据变形散点关系确定； $\epsilon$  为余量序列。

2) 应用逐步回归算法，确定式 (14) 的系数  $c_{qg}$ ，得到余量序列  $\epsilon$ 。

3) 利用 ADF 检验<sup>[15-16]</sup>，测试  $\epsilon$  的平稳性。若  $\epsilon$  平稳，则协整，执行步骤 4；若非平稳，则不协整，此时算法不适用。

4) 利用式 (14)，估计  $x'_h$  的奇异成分，修正矩阵  $X$ 。

5) 利用 PCA 准则，辨识矩阵  $X$  中是否仍然存在奇异成分，若存在，重复执行步骤 1~4，直至无奇异成分。

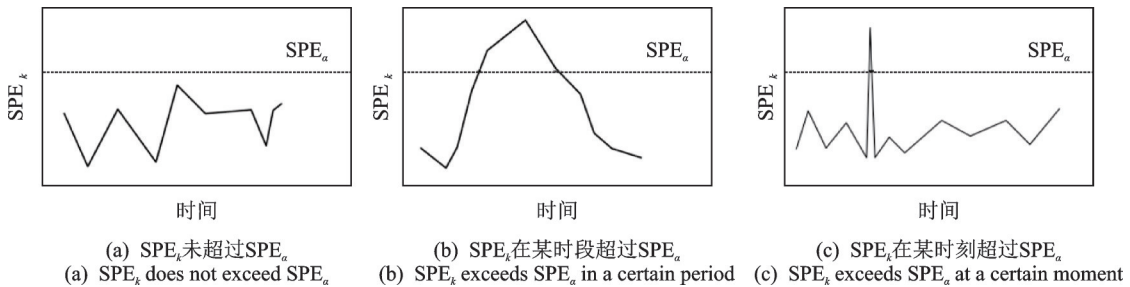


图 1 SPE 统计量与控制限的关系

Fig.1 Relationships between SPE and its control limit

## 2 工程实例分析

### 2.1 工程实例 1

陈村拱坝最大坝高为 76.3 m，设计汛限水位为 117.0 m，汛后正常高水位为 119.0 m，500 年校核水位为 123.8 m，5 000 年校核水位为 126.1 m，保坝水位为 127.7 m。18 正下 (1 号测点)，26 正下 (2 号测点) 和 29 倒 (3 号测点) 测点的位置如图 2 所示，径向变形过程线和置信区间法计算结果如图 3 所示，环境量过程线如图 4 所示。

依据水压-温变-时效 (hydraulic-seasonal-time, 简称 HST) 模式<sup>[17]</sup>，建立变形监控模型，HST 模型精度如表 1 所示。由表可知，3 个模型的复相关系数

$R$  均大于 0.95，且标准差  $S$  较小，拟合精度较高。应用置信区间法<sup>[17]</sup>，评判变形变化是否正常， $2S$  为正常与基本正常的临界值， $3S$  为基本正常与异常的临界值。由图 3(b)~(d) 可知：① 1 号测点未超出  $3S$ ，在 42 天超出了  $2S$ ，未报警；② 2 号测点未超出  $3S$ ，在 23 天超出了  $2S$ ，未报警；③ 3 号测点在 2 天超出了  $3S$ ，且在 148 天超出了  $2S$ ，发生报警。

由图 4 可知，该时段未出现非常规水位和温度状况，因此受奇异成分影响，监控模型虚假报警。利用本研究方法，诊断 3 个测点的奇异监测数据。经计算， $e_1$  为 95.23%，且余量序列协整，诊断出的奇异监测数据有斑点型和孤立型，如：1 号测点的变形监测值在 2012-07-11 为孤立型奇异；3 号测点的测值在 2011-05-25 至 2011-07-19 期间为斑点型奇异。

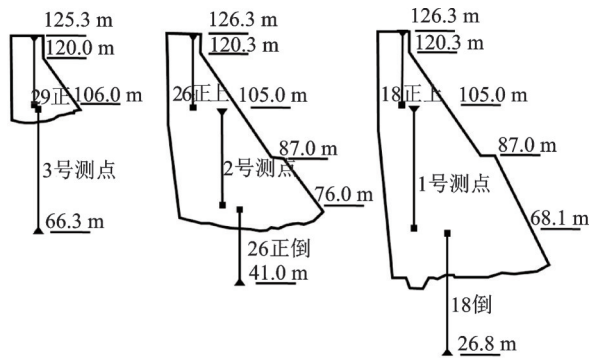


图 2 29#,26#和18#坝段垂线监测点  
Fig.2 Vertical monitoring points of the 29#, 26# and 18# dam sections

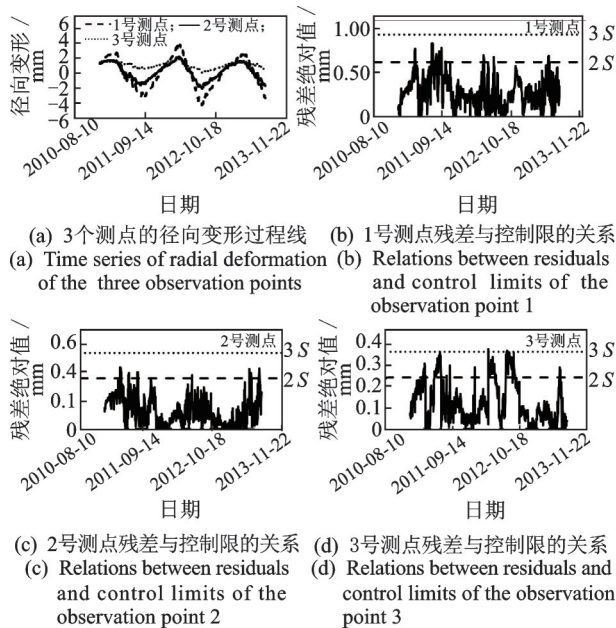


图 3 径向变形过程线和置信区间法计算结果  
Fig.3 Time series of radial deformation and calculation results of confidence interval approach

表 1 HST 模型精度  
Tab.1 Accuracy of the HST models

评价参数	1号测点	2号测点	3号测点
复相关系数 $R$	0.990 5	0.988 1	0.963 2
标准差 $S$	0.312 7	0.177 9	0.120 7

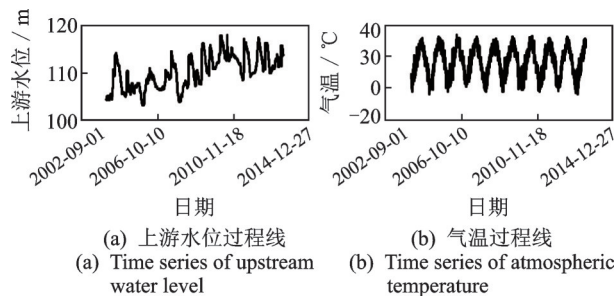


图 4 环境量过程线  
Fig.4 Time series of environment factors

重新建立监控模型,结果显示:①1号测点在 3 天超出了 2S,未超过 3S;②2号测点在 5 天超出了 2S,未超过 3S;③3号测点在 14 天超出了 2S,未超过 3S。

## 2.2 工程实例 2

斑点型奇异可视作孤立型奇异的集合,因此孤立型奇异辨识和斑点型奇异估计具有分析的代表性。结合锦屏一级拱坝 3 个垂线点 PL11-3, PL11-4 和 PL13-3 的径向变形监测资料,通过人为构造奇异成分<sup>[18]</sup>,检验 PCA-CA 方法的性能。测点布置如图 5 所示,图 6 为径向变形原位监测过程线。

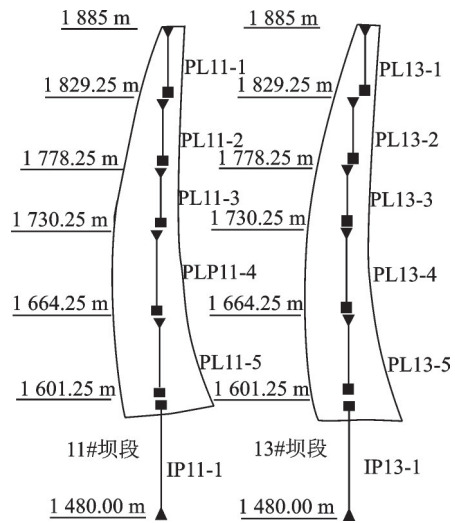


图 5 11#,13#坝段垂线监测点布置  
Fig.5 Vertical monitoring points of the 11# and 13# dam sections

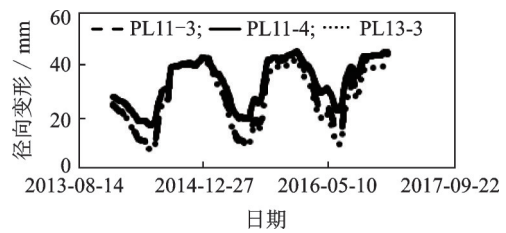


图 6 PL11-3,PL11-4 和 PL13-3 的径向变形监测序列  
Fig.6 Time series of radial deformation of PL11-3, PL11-4 and PL13-3

### 2.2.1 PCA 辨识性能分析

图 7 为 PCA 方法的诊断结果。依据 PCA 辨识步骤 1~5,得到了  $t_1 \sim t_3$  效应量,如图 7(a)~(c)所示,测点变形时间序列的相关系数矩阵如表 2 所示,  $t_1 \sim t_3$  的特征值、解释能力以及累积解释能力如表 3 所示。依据步骤 6,可得到  $SPE_k$  过程线,在执行步

骤 7 时,  $\alpha$  取为 0.01, 计算得到控制限  $SPE_{0.01} = 0.11$ , 由图 7(d) 可知,  $SPE_k < SPE_{0.01}$ , 因此 3 个测点的变形原位监测数据不包含奇异成分。

表 2 相关系数矩阵

**Tab.2 Correlation coefficient matrix**

测点编号	PL11-3	PL11-4	PL13-3
PL11-3	1.000 0	0.994 1	0.993 5
PL11-4	0.994 1	1.000 0	0.975 9
PL13-3	0.993 5	0.975 9	1.000 0

表 3 特征值、解释能力及累积解释能力

**Tab.3 Eigenvalues, explanatory capabilities and accumulated explanatory capabilities**

效应量	第 $j$ 项的特征值 $\lambda_j$	第 $j$ 项的解释能力 $e_j$	前 $j$ 项的累积解释能力 $\sum_{i=1}^j e_i$
$t_1(j=1)$	2.974 6	0.991 5	0.991 5
$t_2(j=2)$	0.024 8	0.008 3	0.999 8
$t_3(j=3)$	0.000 6	0.000 2	1.000 0

PL11-4 测点在 2014-10-03 的径向变形监测值为 39.4 mm, 将其构造为不同程度的孤立型奇异成分, 记作 0#~6# 情况, 分别为 39.6, 39.9, 40.4, 40.9, 41.4, 42.4 和 43.4 mm。图 8 为 1#~5# 情况下  $SPE_k$  与  $SPE_{0.01}$  的关系图, 图 9 为 1#~6# 情况的 CSPE 统计量。由图 8 可知, 在 2#~5# 情况下,  $SPE_k$  超过了  $SPE_{0.01}$ , 且从图 9 可以看出, 在 2#~6# 情况下, PL11-4 测点的 CSPE 统计量明显高于 PL11-3 和 PL13-3, 因此 2#~6# 可判定为孤立型奇异成分。表 4 比较了 PCA 准则、拉依达准则、狄克松准则和  $t$  准则的性能, 可以看出, 4 种方法分别可辨识出相对误差为 3.81%, 7.61%, 7.61% 和 5.08% 的孤立型奇异成分。

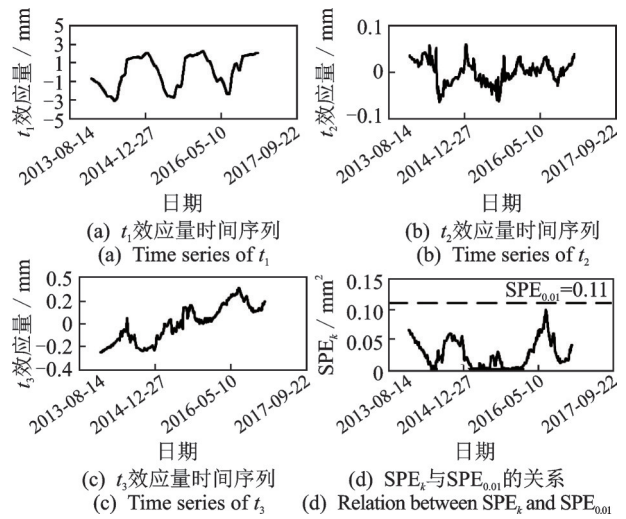


图 7 PCA 方法的诊断结果

Fig.7 Diagnosis results of the PCA method

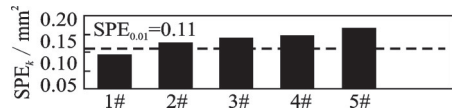


图 8 1#~5# 情况下  $SPE_k$  与  $SPE_{0.01}$  的关系

Fig.8 Relations between  $SPE_k$  and  $SPE_{0.01}$  under 1# case~5# case

2.2.2 CA 估计性能分析

假设 PL11-4 测点在 2016-07-15 至 2016-09-15 期间的变形监测数据为斑点型奇异, 利用 2014-01-01 至 2016-07-14 的原位监测资料建立协整模型, 对奇异成分进行估计。图 10 为 3 个测点径向变形散点关系。由图可知, 式(16)的最高次数  $a_q$  均取 1, 利用逐步回归法, 可得

$$x'_{PL11-4} = 1.328\ 581x''_{PL11-3} - 0.626\ 32x''_{PL13-3} + \epsilon \quad (17)$$

其中:  $x''_{PL11-3}$ ,  $x'_{PL11-4}$  和  $x''_{PL13-3}$  分别为 PL11-3, PL11-4 和 PL13-3 测点的径向变形原位监测序列。

图 11 为 CA 方法的估计结果。PL11-4 的监测值、拟合值和余量序列如图 11(a) 所示, 复相关系数

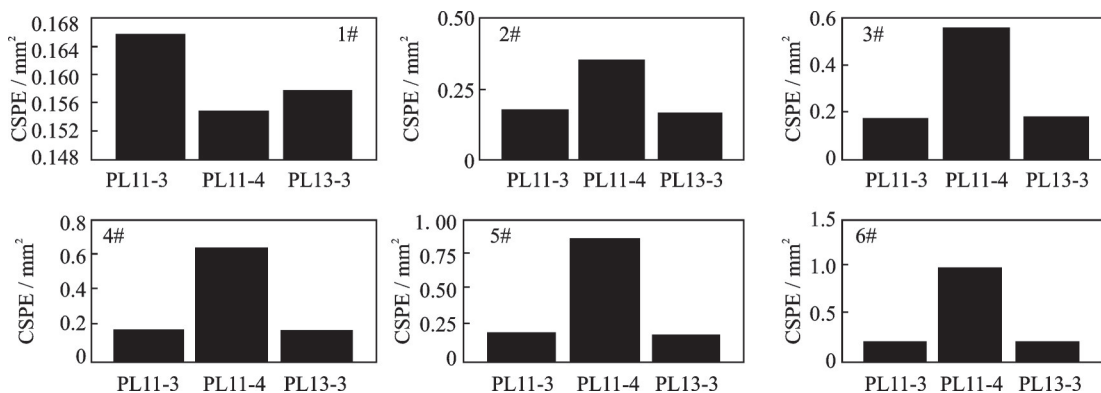


图 9 1#~6# 情况的 CSPE 统计量

Fig.9 CSPE statistics of 1# case~6# case

表4 4种方法性能对比  
Tab.4 Performance comparison of four methods

情况	监测值/mm	奇异值/mm	相对误差/%	PCA 准则	常规方法		
					拉依达准则	狄克松准则	$t$ 准则
0	39.4	39.6	0.51	未识别	未识别	未识别	未识别
1	39.4	39.9	1.27	未识别	未识别	未识别	未识别
2	39.4	40.4	2.53	未识别	未识别	未识别	未识别
3	39.4	40.9	3.81	识别	未识别	未识别	未识别
4	39.4	41.4	5.08	识别	未识别	未识别	识别
5	39.4	42.4	7.61	识别	识别	识别	识别
6	39.4	43.4	10.15	识别	识别	识别	识别

为 0.997 1。ADF 检验结果显示： $t$  统计量的值为 -2.724 4，1% 和 5% 两个显著性水平的临界值分别为 -2.400 和 -2.150，因此  $t$  统计量小于临界值，即原假设均被拒绝，余量序列  $\epsilon$  平稳。图 11(b) 为自回归模型、统计模型和 CA 模型对奇异成分的估计结果，3 种模型的复相关系数分别为 0.743 2, 0.871 5 和 0.994 5。

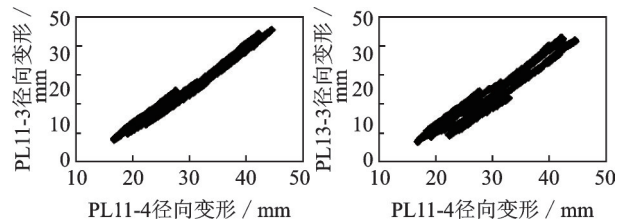
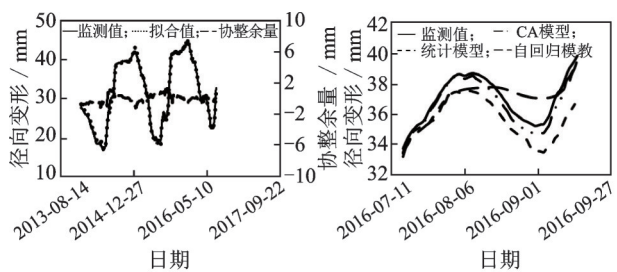


图 10 PL11-3, PL11-4 和 PL13-3 测点径向变形散点关系  
Fig.10 Scatter relations among radial deformation of PL11-3, PL11-4 and PL13-3



(a) PL11-4 监测值、拟合值和余量序列 (b) 奇异成分监测值和估计值  
(a) Monitoring values, fitting values and residual values of PL11-4 (b) Monitoring values and estimation values of singular components

图 11 CA 方法的估计结果

Fig.11 Estimation results of the CA method

### 3 结论

1) 为消除奇异成分对大坝变形安全监控结论客观性的影响，运用 PCA 理论，建立了 SPE 和 CSPE，借助假设检验，提出了 PCA 辨识准则，在此基础上，基于 CA 原理，应用 ADF 检验和逐步回归法，提出了 CA 似然估计模型。利用陈村拱坝和锦

屏一级拱坝径向变形原位监测资料，检验了 PCA-CA 方法的有效性。

2) 拉依达准则、狄克松准则和  $t$  准则以 1 维时间序列为分析对象，辨识性能对样本分布型式的依赖性较大，仅体现了概率意义。PCA 准则针对多维数据序列进行分析，不仅包含概率含义，而且考虑了多测点变形的主成分关系，辨识性能最佳。

3) 自回归模型和统计模型着重分析 1 维时间序列，其中：自回归模型仅考察了前期监测数据的时间波动规律，外延性较差；统计模型刻画了水压、温变和时效因素，外延性优于自回归模型，但属于半经验模型，估计精度仍不足；CA 模型以多维数据序列为分析对象，既刻画了变形随时间的波动特征，亦表征了多测点变形间的协整关系，同时也反映出外部作用的综合效应，外延性较好，取得了最佳的估计精度。

4) PCA-CA 方法要求：①  $t_1$  效应量对原始变形序列  $X$  的解释程度  $e_1$  尽可能大，通过工程实例 1 可知，当  $e_1 = 95.23\%$  时，PCA 准则可有效辨识出奇异数据；② 余量序列平稳。对于我国已修筑的部分老坝，布置的变形监测仪器较少，不易满足上述条件，此时，仍需采用常规方法。

### 参 考 文 献

[1] WANG S W, XU Y L, GU C S, et al. Hysteretic effect considered monitoring model for interpreting abnormal deformation behavior of arch dams: a case study[J]. Structural Control and Health Monitoring, 2019, 26(10): e2417.  
[2] LI B, YANG J, HU D X. Dam monitoring data analysis methods: a literature review[J]. Structural Control and Health Monitoring, 2020, 27(3): e2501.  
[3] 黄潇霏. 混凝土坝健康状态的实测资料时空诊断方法[D]. 南京: 河海大学, 2018.

- [4] 吴中如, 顾冲时. 重大水工混凝土结构病害检测与健康诊断[M]. 北京: 高等教育出版社, 2005: 25-47.
- [5] 李德仁, 王树良, 李德毅. 空间数据挖掘理论与应用[M]. 3版. 北京: 科学出版社, 2019: 6-20.
- [6] 付云骁, 贾利民, 秦勇, 等. 基于LMD-CM-PCA的滚动轴承故障诊断方法[J]. 振动、测试与诊断, 2017, 37(2): 249-255.  
FU Yunxiao, JIA Limin, QIN Yong, et al. Roller bearing fault diagnosis method based on LMD-CM-PCA [J]. Journal of Vibration, Measurement & Diagnosis, 2017, 37(2): 249-255. (in Chinese)
- [7] 李振, 李伟光, 赵学智, 等. 基于主成分分析的特征频率提取算法及应用[J]. 振动、测试与诊断, 2018, 38(4): 834-842.  
LI Zhen, LI Weiguang, ZHAO Xuezhi, et al. Feature frequency extraction algorithm based on principal component analysis and its application[J]. Journal of Vibration, Measurement & Diagnosis, 2018, 38(4): 834-842. (in Chinese)
- [8] PIROG D, FIDELUS-ORZECZOWSKA J, WIEJ-ACZKA L, et al. Hierarchy of factors affecting the social perception of dam reservoirs[J]. Environmental Impact Assessment Review, 2019, 79: 106301.
- [9] 李伟光, 郭明军, 杨期江, 等. 改进PCA算法及其在转子特征提取中的应用[J]. 振动、测试与诊断, 2020, 40(3): 555-562.  
LI Weiguang, GUO Mingjun, YANG Qijiang, et al. Improved PCA algorithm and its application in rotor feature extraction[J]. Journal of Vibration, Measurement & Diagnosis, 2020, 40(3): 555-562. (in Chinese)
- [10] LI F Q, WANG Z Y, LIU G H. Towards an error correction model for dam monitoring data analysis based on co-integration theory [J]. Structural Safety, 2013, 43: 12-20.
- [11] SHEN Y, SIU T K. Optimal investment and consumption in a continuous-time co-integration model [J]. IMA Journal of Management Mathematics, 2017, 28(4): 501-530.
- [12] 李耀龙, 李洪儒, 王冰, 等. 基于协整理论的滚动轴承退化特征提取[J]. 振动、测试与诊断, 2021, 41(2): 385-391.  
LI Yaolong, LI Hongru, WANG Bing, et al. Extrac-
- tion of degradation feature for rolling bearings based on co-integration theory[J]. Journal of Vibration, Measurement & Diagnosis, 2021, 41(2): 385-391. (in Chinese)
- [13] 侯彦东, 闫治宇, 金勇. 小样本下基于特征子空间估计的故障诊断算法[J]. 上海交通大学学报, 2015, 49(6): 825-829.  
HOU Yandong, YAN Zhiyu, JIN Yong. Fault diagnosis algorithm of based feature subspace estimation in small sample circumstance[J]. Journal of Shanghai Jiaotong University, 2015, 49(6): 825-829. (in Chinese)
- [14] ZENG L, LONG W, LI Y Y. A novel method for gas turbine condition monitoring based on KPCA and analysis of statistics T-2 and SPE[J]. Processes, 2019, 7(3): 124.
- [15] KUO C Y. Sustainability of current accounts: evidence from the quantile unit-root test[J]. Applied Economics Letters, 23(10): 747-755.
- [16] ZHANG L X. Performance of unit-root tests for non linear unit-root and partial unit-root processes[J]. Communications in Statistics-Theory and Methods, 2016, 45(15): 4528-4536.
- [17] YANG G, GU H, CHEN X D, et al. Hybrid hydraulic-seasonal-time model for predicting the deformation behaviour of high concrete dams during the operational period [J]. Structural Control and Health Monitoring, 2021, 28(3): e2685.
- [18] YANG J, QU X D, HU D X, et al. Research on singular value detection method of concrete dam deformation monitoring[J]. Measurement, 2021, 179: 109457.



**第一作者简介:** 杨光, 男, 1989年9月生, 博士、讲师。主要研究方向为水工结构安全监控与变异诊断。曾发表《Hybrid hydraulic-seasonal-time model for predicting the deformation behaviour of high concrete dams during the operational period》(《Structural Control and Health Monitoring》2021, Vol.28, No.3)等论文。  
E-mail: yangguanghhu@outlook.com