

基于原型监测资料的特高拱坝变形实时风险率模型

牛景太¹, 姜灵¹, 邓志平¹, 王志强¹, 刘知远², 张阳¹

(1. 南昌工程学院 水利与生态工程学院, 江西 南昌 330099; 2. 南昌工程学院 信息工程学院, 江西 南昌 330099)

摘要: 混凝土坝风险率分析模型通常基于结构极限状态功能函数对单个监测点一维时间序列建模, 未考虑变形监测点之间的相关性及多重共线性问题。基于原型监测资料, 考虑各分区所有测点的相关性及不同分区变形之间的协同性, 引入面板数据理论对特高拱坝监测点进行聚类分区, 在拟定单测点风险率函数的基础上, 提出计算特高拱坝变形分区单测点实时风险率的方法, 基于 Copula 函数进一步构建基于原型监测资料的特高拱坝整体实时风险率分析模型。实例分析表明, 所构建的模型确定了依据大坝长序列立体监测数据建立变形实测效应量与风险率的函数关系, 可有效分析特高拱坝各分区变形风险率及整体变形风险率, 能够客观刻画特高拱坝整体风险率变化的基本规律。

关键词: 特高拱坝; 变形风险率; 聚类分区; Copula 函数; 原型监测

中图分类号: TV642.4⁺2; TV213.4

文献标识码: A

文章编号: 1672-643X(2021)05-0166-09

Real-time risk rate model for deformation of ultra-high arch dams based on prototype monitoring data

NIU Jingtai¹, JIANG Ling¹, DENG Zhiping¹, WANG Zhiqiang¹, LIU Zhiyuan², ZHANG Yang¹

(1. School of Hydraulic and Ecological Engineering, Nanchang Institute of Technology, Nanchang 330099, China;

2. School of Information Engineering, Nanchang Institute of Technology, Nanchang 330099, China)

Abstract: The risk rate analysis model of concrete dams is usually used for one-dimensional time series of individual monitoring points based on the structure limit state function, without considering the correlation and multicollinearity between deformation monitoring points. Based on the prototype monitoring data, considering the correlation of all monitoring points in each partition and the synergy of the deformation in different partitions, the panel data theory is introduced for the cluster partition of the monitoring points of ultra-high arch dams, and the calculation method of real-time single monitoring point risk rate of ultra-high arch dam deformation partition is proposed based on the estimated single monitoring point risk rate function. Then the overall real-time risk rate analysis model of ultra-high arch dams based on prototype monitoring data is further constructed combined with Copula function. The case analysis shows that the constructed model establishes the functional relationship between the actual deformation effect and the risk rate based on the dam long time series monitoring data. It can effectively analyze the deformation risk rate of each partition and the whole dam, and can objectively describe the basic law of the overall risk rate change of ultra-high arch dams.

Key words: ultra-high arch dam; deformation risk rate; cluster partition; Copula function; prototype monitoring

1 研究背景

特高拱坝失事一般都要经历从渐变到突变的过

程, 在此发展演变过程中, 若能及时准确地对特高拱坝的运行状态进行实时监控, 根据监测信息在最短时间内对特高拱坝运行风险率做出客观评价, 并及

收稿日期: 2021-05-23; 修回日期: 2021-07-04

基金项目: 江西省教育厅科技项目 (GJJ161096、GJJ180951、GJJ151100); 国家自然科学基金项目 (51769017、51969018、52009054)

作者简介: 牛景太 (1977-), 男, 山东东明人, 博士, 教授, 硕士生导师, 主要从事水工建筑物安全监控研究。

通讯作者: 邓志平 (1990-), 男, 江西南昌人, 博士, 副教授, 主要从事水工岩土可靠度分析及风险控制研究。

时预警,就有可能避免重大事故的发生^[1-2],因此,基于变形原型监测数据分析特高拱坝的实时风险率,对保障特高拱坝长效健康服役具有十分重要的理论意义和实用价值。

目前,关于混凝土坝风险率的研究多集中在坝体滑动方面^[3-5],其风险率分析模型由常规模型发展到人工智能模型,计算精度、稳健性及外延性也逐渐得到提高。而有关特高拱坝实时风险率的研究还不多,邓建华等^[6]对特高拱坝坝体和坝基两重要部位各取一个测点,基于 Copula 函数计算特高拱坝多变量变形风险率。

针对运维期内特高拱坝结构安全风险的动态性、复杂性以及失事后果严重性,王子成等^[7]基于层次分析法、模糊理论和定量风险评估分析法,研发了一套特高拱坝动态安全风险分析系统。从现有的研究成果看,虽有学者已从时空角度进行了多测点、多测次的特高拱坝风险率分析,但忽略了变形监测点之间的相似性及影响因子之间的共线性问题,以及海量监测数据中时间和空间两个维度上的关联性和大坝动态空间结构的协同性,因而分析模型的拟合精度和稳定性较低,且难以全面掌握特高拱坝整体结构及完整时段的变化性态。

鉴于上述问题,基于面板数据理论在时空维度上表征特高拱坝的原型监测信息,依据相似性准则对多个测点进行聚类分析,实现特高拱坝变形动态分区。建立特高拱坝面板数据随机系数模型,拟定基于原型监测数据的单测点实时风险率函数。在此基础上,考虑分区内所有测点的相关性及不同分区变形之间的协同性,结合 Copula 函数,提出基于变形原型监测资料的特高拱坝实时风险率分析方法。最后,通过实例模拟计算验证所构建模型的合理性及有效性。

2 基于面板数据理论的特高拱坝变形分析模型

特高拱坝受荷载作用、外界环境、材料特性等多因素协同影响,影响因子之间存在多重共线性程度,导致特高拱坝整体变形性态呈现区域特征,不同区域部位之间存在差异性和协同性,且特高拱坝时间和空间维度上的变形监测信息存在一定的关联性。因此,在统计模型的基础上,引入面板数据理论在时空维度上表征特高拱坝的原型监测信息,依据相似性准则对多个测点进行聚类动态分区,以降低多重共线性的影响,构建特高拱坝变形面板数据随机系

数模型。

2.1 变形统计分析模型

特高拱坝的变形统计分析模型^[8]中主要荷载因子包括水压分量、温度分量和时效分量,其表达式为:

$$\delta = \delta_H + \delta_T + \delta_\theta = \sum_{i=1}^n a_i H^i + \sum_{i=1}^2 \left(b_{i1} \sin \frac{2\pi i t}{365} + b_{i2} \cos \frac{2\pi i t}{365} \right) + c_1 \theta + c_2 \ln \theta \quad (1)$$

式中: δ 为在某时刻 t 时特高拱坝某监测点的变形值; $\delta_H, \delta_T, \delta_\theta$ 分别为变形水压、温度及时效分量; a_i, b_i, c_1, c_2 为回归系数; H 为上游水深, m; 拱坝中 n 取 4; t 为起始日至监测日的累计天数; θ 为累计天数除以 100(每天增加 0.01)。

2.2 变形随机系数面板模型

2.2.1 基于面板数据的特高拱坝变形分区 特高拱坝是整体空间结构,各测点之间内在联系紧密,依据欧氏距离^[9]作为相似性指标,判别各测点之间的亲疏程度。在此基础上,基于离差平方和法(Ward 法)^[10-11]对特高拱坝进行聚类分析,从而实现变形动态分区。设特高拱坝 N 个监测点在 T 个不同时段内的位移实测值为 δ_{it} ,用面板数据表征特高拱坝位移实测值的具体形式如下:

$$\delta_{it} = \begin{bmatrix} \delta_{i1} & \delta_{i2} & \cdots & \delta_{it} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2t} \\ \vdots & \vdots & & \vdots \\ \delta_{N1} & \delta_{N2} & \cdots & \delta_{Nt} \end{bmatrix} \quad (2)$$

($i = 1, 2, \dots, N; t = 1, 2, \dots, T$)

式中: i 为监测点个数; t 为监测时间; δ_{it} 包含了横截面 N 和时间 T 两个维度。

欧氏距离 d_{ij} 反映了实测值之间的亲疏程度,其计算公式为:

$$d_{ij} = \left[\sum_{t=1}^T (\delta_{it} - \delta_{jt})^2 \right]^{1/2} \quad (3)$$

式中: δ_{it} 为测点 i 在 t 时段内的变形监测值, mm; δ_{jt} 为测点 j 在 t 时段内的变形监测值, mm。

将 N 个测点分为 k 个区域,记为 $G_1, G_2, \dots, G_k$,用 X_{it} 表示 G_i 类中的第 i 个测点的变形值,则 G_i 区域中不同测点序列的离差平方和为:

$$W_i = \sum_{i=1}^{N_i} \sum_{t=1}^T [(X_{it} - \bar{X}_i)' (X_{it} - \bar{X}_i)] \quad (4)$$

$$\bar{X}_i = \frac{1}{N_i} \sum_{t=1}^{N_i} X_{it} \quad (5)$$

式中: W_i 为总离差平方和; X_{it} 为 t 时段内测点 i 的变形值; $\bar{X}_i$ 为 G_i 类的测值重心。

离差平方和反映了分区间的相似性,按照同类测点的离差平方和较小进行分区。

因此,大坝聚类分区步骤如下:

(1) 设最初有 N 类,即每个变形测点各自成一类 $\{G_1, G_2, \dots, G_n\}$, 根据公式(3) 计算 N 个变形测点两两间的欧氏距离 d_{ij} , 记为矩阵 D_0 ;

(2) 依据离差平方和 W_i 最小准则,将欧氏距离最小的两个区域合并为一个新的区域;

(3) 计算新区域和其他区域的欧氏距离 d_{ij} , 得到新的距离矩阵 D_i , 重复合并区域步骤(2) ~ (3), 直到所有测点归于一个区域为止;

(4) 绘制聚类树状图;

(5) 根据聚类树状图分类并结合实际情况,确定大坝变形分区数 M 。

2.2.2 构建随机系数面板模型 上节基于离差平方和对特高拱坝进行了聚类分析,实现了动态分区,下面重点研究构建能反映特高拱坝实际动态变形规律的变系数面板模型方法。

面板数据模型^[12]的一般形式可表示为:

$$y_{it} = \sum_{k=1}^K \beta_{ki} x_{kit} + \varepsilon_{it} \quad (6)$$

$$(i = 1, 2, \dots, N; t = 1, 2, \dots, T)$$

式中: y_{it} 为特高拱坝变形监测的面板数据; t 为面板数据中的时间序数; i 为面板数据的截面序列; β_{ki} 为第 i 个监测点的随机系数; x_{kit} 为解释变量的面板数据组, k 为解释变量的序数; ε_{it} 为随机误差。

基于特高拱坝 N 个变形监测数据,建立如下面板模型:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} X_1 & & 0 \\ & X_2 & \\ & & \ddots \\ 0 & & & X_N \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix} \quad (7)$$

式中: y_i 为监测时间序列, $y_i = (y_{i1}, y_{i2}, \dots, y_{iT})'$; X_i 为监测点 $t \times k$ 的解释变量矩阵, $X_i =$

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ik} \end{bmatrix}; x_{ik} = (1, H_t^1, H_t^2, H_t^3, H_t^4,$$

$T_{1,t}, \dots, T_{m,t}, \theta_t, \ln \theta_t)'$; $H_t^1, H_t^2, H_t^3, H_t^4$ 为变形监测值的水压分量影响因素; $T_{1,t}, T_{2,t}, \dots, T_{m,t}$ 为温度分量影响因素; $\theta_t, \ln \theta_t$ 为时效分量影响因素; $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$ 为测点的待估参数; $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{ik})'$ 为监测序列的随机误差。

由于各监测点的变形实测值变化规律不同, β_i 亦非为固定值,若考虑特高拱坝每个分区中不同测点的相似性,其面板模型系数可用均值系数 $\bar{\beta}$ 来表示,分区内不同监测点之间的差异性记为 α_i ,也反映特高拱坝不同区域之间的多重共线性,将 $\beta_i = \bar{\beta} + \alpha_i$ 视为随机变量,则基于 NT 个变形监测数据建立的变系数面板数据矩阵方程式为:

$$y = X\bar{\beta} + \tilde{X}\alpha + \varepsilon \quad (8)$$

式中: $y_{NT \times 1} = (y'_1, y'_2, \dots, y'_N)'$; $\bar{\beta} = (\bar{\beta}_0, \bar{\beta}_1, \dots, \bar{\beta}_k)'$ 为共同均值系数; $\varepsilon = (\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_N)'$;

$$X_{NT \times K} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}; \tilde{X}_{NT \times NK} = \begin{bmatrix} X_1 & & 0 \\ & X_2 & \\ & & \ddots \\ 0 & & & X_N \end{bmatrix};$$

$\alpha = (\alpha'_1, \alpha'_2, \dots, \alpha'_N)$; $\alpha_i = (\alpha_{0i}, \alpha_{1i}, \alpha_{ki})'$ 为同一分区内不同测点均值的偏差, α_i 为均值为 0、方差和协方差为常数的随机变量。即假定:

$$E\alpha_i = 0, E\alpha_i \alpha'_j = \begin{cases} \Delta & (i = j) \\ 0 & (i \neq j) \end{cases}, E\varepsilon_i \varepsilon'_j = 0,$$

$$E\alpha_i \varepsilon'_j = 0, E\varepsilon_i \varepsilon'_j = \begin{cases} \sigma_i^2 I_T & (i = j) \\ 0 & (i \neq j) \end{cases}。$$

复合误差项 $\tilde{X}\alpha + \varepsilon$ 的协方差矩阵是对角矩阵,第 i 个对角块为:

$$\Psi_i = X_i \Delta X'_i + \sigma_i^2 I_T \quad (9)$$

式中: Ψ_i 为协方差矩阵第 i 个对角块; $\sigma_i^2 I_T$ 为第 i 个监测序列随机误差的协方差; I_T 为第 i 个监测序列随机误差的协方差系数。

对变形随机系数面板模型,传统的 y 对 x 线性回归得出的 β 估计量不是有效估计,且常用来计算估计量的最小二乘公式也不适用。因此,下面研究随机系数中共同均值系数 $\bar{\beta}$ 的估计方法,进一步得到变形系数面板模型随机系数 β 的预测方法。

2.2.3 变系数面板模型随机系数的估计 为了检验建立的变形随机系数面板模型的有效性,需对随机系数 β 进行估计,下面先探究共同均值系数 $\bar{\beta}$ 的最佳估计量。由文献[13]、[14]可知, $\bar{\beta}$ 的最优线性无偏估计量为:

$$\hat{\bar{\beta}}_{GLS} = \left(\sum_{i=1}^N X'_i \Psi_i^{-1} X_i \right)^{-1} \left(\sum_{i=1}^N X'_i \Psi_i^{-1} y_i \right) = \sum_{i=1}^N W_i \hat{\beta}_i \quad (10)$$

$$\text{Var}(\hat{\beta}_{GLS}) = \left(\sum_{i=1}^N X_i' \Psi_i^{-1} X_i \right)^{-1} = \left\{ \sum_{i=1}^N [\Delta + \sigma_i^2 (X_i' X_i)^{-1}]^{-1} \right\}^{-1} \quad (11)$$

$$W_i = \left\{ \sum_{i=1}^N [\Delta + \sigma_i^2 (X_i' X_i)^{-1}]^{-1} \right\}^{-1} [\Delta + \sigma_i^2 (X_i' X_i)^{-1}]^{-1} \quad (12)$$

$$\hat{\beta}_i = (X_i' X_i)^{-1} X_i' y_i \quad (13)$$

式中: $\hat{\beta}_{GLS}$ 为广义最小二乘 (GLS) 估计量; Ψ_i^{-1} 为协方差矩阵第 i 个对角块矩阵的转置矩阵; X_i 为第 i 个变形监测点 $t \times k$ 的解释变量矩阵; X_i' 为第 i 个变形监测点 $t \times k$ 的解释变量矩阵的转置矩阵。

$\hat{\beta}_{GLS}$ 服从渐进正态分布, 为共同均值系数 $\bar{\beta}$ 的有效估计量。在随机系数面板模型中, 每一个随机变量 β_i 与解释变量 X_i 无关, 且有固定的均值 $\bar{\beta}$, 因此在得出共同均值系数 $\bar{\beta}$ 的有效估计量的基础上可进一步得到随机系数 β 的预测值 $\hat{\beta}$, 即:

$$\hat{\beta} = \hat{\beta}_{GLS} + \Delta X_i' (X_i \Delta X_i' + \sigma_i^2 I_T)^{-1} (y_i - X_i \hat{\beta}_{GLS}) \quad (14)$$

由于 $E(\hat{\beta} - \beta) = 0$, 因此随机系数估计量 $\hat{\beta}$ 为最佳线性无偏估计量, 其期望是无条件期望。面板数据随机系数模型中每一个解释变量的随机系数估计量均为服从渐进正态分布的随机变量, 可进一步得到每一区监测点的随机系数估计量, 建立不同监测点位移实测值之间的关联性和差异性, 从而更好地分析控制由于实测资料波动而引起的估计量较大变动, 以增加模型的稳定性和有效性。

3 基于变形监测资料的特高拱坝实时风险率分析模型

变形是大坝服役状态最直观反映, 依据原型监测数据更能精确分析特高拱坝的风险率。因此, 在前文构建特高拱坝变形面板数据随机系数模型的基础上拟定监控指标, 依据测点变形预测值和 Copula 函数构建特高拱坝分区实时风险率分析模型, 提出特高拱坝整体风险率实时分析方法。

3.1 监控指标的拟定

选取置信区间法^[15] 拟定监控指标, 位移监测效应量的监控指标 $\hat{\delta}_m$ 为:

$$\hat{\delta}_m = \delta_m \pm \beta \sigma \quad (15)$$

式中: δ_m 为监测量的极值; β 为可靠性指标, 与显著性水平 α (一般取值 1% ~ 5%) 有关; σ 为剩余标准差。

3.2 构建特高拱坝实时风险率分析模型

基于特高拱坝面板数据随机系数模型, 在任意时刻变形实测值序列中的测值均服从正态分布, 根据拟定的监控指标, 构建分区测点风险率函数, 确定特高拱坝变形分区某一测点实时风险率。

风险率函数的表达式为:

$$P_i = K \cdot \int_{\mu_i - |\mu_i - \delta_i|}^{\mu_i + |\mu_i - \delta_i|} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\delta_i - \mu_i)^2}{2\sigma^2}\right) d\delta_i \quad (16)$$

式中: P_i 为测点效应量第 i 天的风险率; K 为风险率系数; μ_i 为第 i 天的预测值; δ_i 为第 i 天的实测值; σ 为剩余标准差。

由于拟定监控指标时, $P\{R > R_{\text{设计}}\} = 1 - \alpha$, 其中 R 为大坝安全准则中的设计允许值, 则称在 $(1 - \alpha) \times 100\%$ 置信度下保证了要求的可靠度 R 设计, 变形实测值服从正态分布, $P\{\delta_i < \delta_{m1}\} = \frac{\alpha}{2}$,

$P\{\delta_i > \delta_{m2}\} = 1 - \frac{\alpha}{2}$ (δ_{m1} 、 δ_{m2} 为监控指标的上、下限), 因此, 风险率系数 K 取 $\frac{\alpha}{2}$ (α 为显著性水平, 一般取值范围为 1% ~ 5%)。

则基于原型监测数据的特高拱坝变形分区某一测点风险率函数可表示为:

$$P_i = \frac{\alpha}{2} \cdot \int_{\mu_i - |\mu_i - \delta_i|}^{\mu_i + |\mu_i - \delta_i|} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\delta_i - \mu_i)^2}{2\sigma^2}\right) d\delta_i \quad (17)$$

3.3 基于 Copula 函数的特高拱坝分区监测点风险率及整体风险率计算

对 d 维随机变量 $(x_1, x_2, \dots, x_d)$ 的联合累积概率分布函数为 $F(x_1, x_2, \dots, x_d)$, 边缘分布函数为 $F_1, F_2, \dots, F_d$, 联合累积概率分布函数用 Copula 函数 C 表示为:

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d)) \quad (18)$$

3.3.1 最佳 Copula 函数的选择 将变形监测点的风险率函数作为边缘概率密度函数, 基于均方根误差 (root mean square error, RMSE) 准则统计检验方法^[16] 确定最佳 Copula 函数。

均方根误差 RMSE 反映了所选联合分布函数的理论概率与样本数据的经验概率之间的差异。RMSE 的方程式为:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [P_c(i) - P_0(i)]^2} \quad (19)$$

式中: n 为样本容量; P_c 为 Copula 函数的计算值; P_0 为样本数据的经验值。

按公式(19)计算出最常见的 5 种 Copula 函数^[19]的 RMSE 值,选取最小值的 Copula 函数为最佳 Copula 函数。

3.3.2 Copula 函数的参数估计 确定最佳 Copula 函数后,采用极大似然估计法^[21]对其进行参数估计,似然函数表示为:

$$L(\theta) = \prod_{i=1}^n c(u_1, u_2, \dots, u_n; \theta) \\ = \prod_{i=1}^n \frac{\partial^n C(u_1, u_2, \dots, u_n; \theta)}{\partial u_1 \partial u_2 \dots \partial u_n} \quad (20)$$

为简便运算,将公式(20)转换为似然函数 $\ln L(\theta)$ 的对数形式进行求解:

$$\ln L(\theta) = \sum_{i=1}^n \ln c(u_1, u_2, \dots, u_n) \quad (21)$$

通过求解方程(21),可以得到极大似然估计量 $\hat{\theta}$,满足 $\ln L(\hat{\theta}) \geq \ln L(\theta)$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0 \quad (22)$$

采用极大似然估计方法对最佳 Copula 函数进行参数估计,确定其函数表达式,最后基于特高拱坝测点谱系聚类树状图并结合最佳 Copula 函数对各分区风险率进行计算,进一步推算特高拱坝整体风险率。因此,特高拱坝整体风险率计算步骤如下:①将特高拱坝监测点作为随机变量,基于公式(18)拟定的监测点变形风险函数作为随机变量的边缘概率密度函数;②基于均方根误差检验方法和待选 Copula 函数计算结果对比选择最佳 Copula 函数;③对最佳 Copula 函数进行参数估计,确定其表达式;④构建特高拱坝分区风险率模型并计算各分区风险率;⑤基于 Copula 函数再次联合各分区风险率,构建特高拱坝整体风险率模型。

4 算例分析

4.1 实例工程简介

云南省某双曲拱坝坝顶高程为 1 885.0 m,坝基最低建基面高程为 1 580.0 m,最大坝高 305.0 m,水库正常蓄水位 1 880.0 m,2012 年 11 月 30 日,导流洞下闸开始首次蓄水。拱坝由 26 个坝段构成,在第 5、9、11、13、16、19 坝段垂线共布置 24 个位移监测点,各监测点位置分布如图 1 所示。

选取 2013 年 6 月 16 日至 2015 年 9 月 28 日的径向位移监测数据(下游为正,上游为负)进行分析研究,因此,面板数据模型中横截面序列个数为 24,时间序列为 274,如图 2 所示,该时段大坝库水位过

程线如图 3 所示。

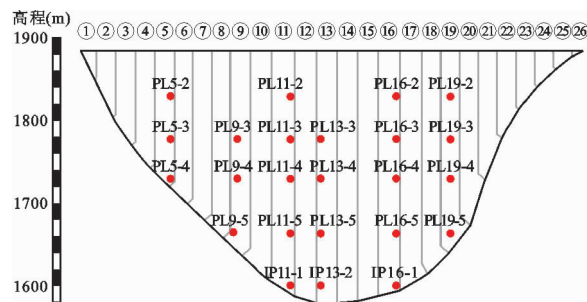


图1 大坝垂线位移测点分布图

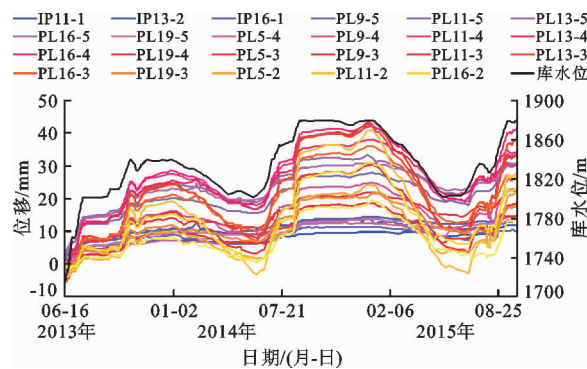


图2 面板数据模型中截面及时间序列数据

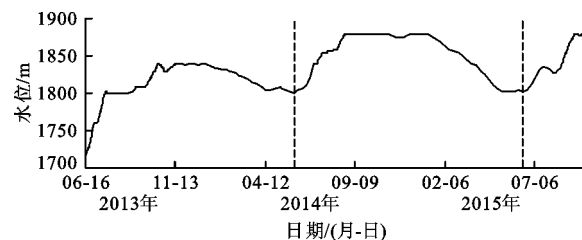


图3 实例工程研究时段大坝库水位过程线

4.2 特高拱坝变形分区

按照特高拱坝的大坝变形分区流程,对特高拱坝的 24 个变形监测点进行大坝变形的分区,得到聚类树状图如图 4 所示。根据实际情况及内部相似性指标评估确定欧氏距离范围,以此来确定特高拱坝变形的区域数,坝体变形区域分布如图 5 所示。

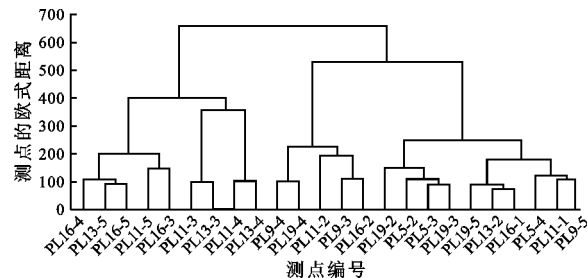


图4 特高拱坝测点谱系聚类树状图

由图 5 可以看出,全坝段 24 个位移变形监测点基本沿拱冠梁两侧对称分布,且各测点变形规律较相似。

4.3 分区随机系数面板模型

基于聚类结果和大坝变形分区,分别对各区建立面板数据随机系数模型进行建模分析。建模前,需对因子进行归一化处理,通过对 4 个区域内测点进行面板随机系数模型建模分析,一区~四区各测点位移实测值及拟合值过程线如图 6~8 所示,随机系数模型的相关系数 R 和标准差 S 见表 1。并得到各区域变形面板模型影响因素系数及检验统计量,以测点数目最多且变化规律较为复杂的四区为例,列举其模型影响因素系数及检验统计量结果见表 2。

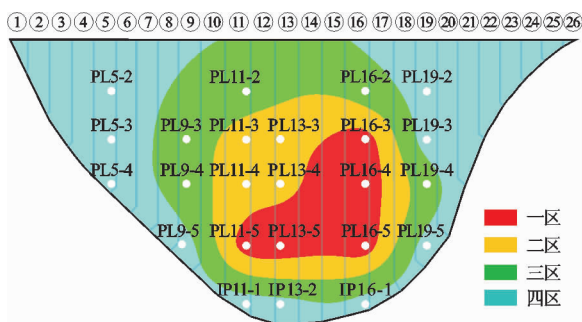


图 5 坝体变形区域分布图

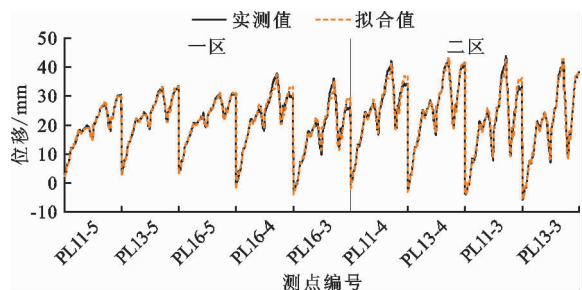


图 6 坝体一区和二区面板随机系数模型
实测值及拟合值过程线

由图 6~8 及表 2 可以看出:特高拱坝变形序列随机系数面板模型整体拟合效果很好,能捕捉有效信息,可较好地描述特高拱坝坝体的变形规律,增强了模型的解释能力。

模型整体检验值 F 值 $Wald\ chi^2(10) = 918.34$, 且 P 值 $Prob > \chi^2 = 0.0000$, 说明该区域变形面板序列采用随机系数面板模型是合理的。由 t 值检验结果来分析模型影响因子,各影响因子 ($P > |t|$) 值均小于 5%, 说明各影响因子在 5% 水平上显著, 所建模型样本及参数选择合理,模型中各因子能较好地描述坝体变形规律。随机系数模型建模过程中考虑了各测点解释变量之间的相互关系,剔除了相关性较强的解释变量,消除了模型中变量间共线性的影响,使最终得到的解释变量分离较为客观。

4.4 分区测点变形风险率计算

取 2015 年 9 月 22 日~2015 年 9 月 28 日分别对建立随机系数面板模型,其中一区各测点位移实测值和预测值统计结果见表 3(限于篇幅,二、三、四区未列出),基于风险函数(本文中显著性水平取值为 5%)求得特高拱坝变形分区各测点在预测日期内的风险率,4 个区各测点风险率计算结果分别见表 4~7。

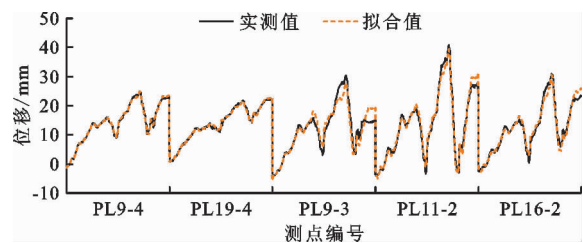


图 7 坝体三区面板随机系数模型实测值及拟合值过程线

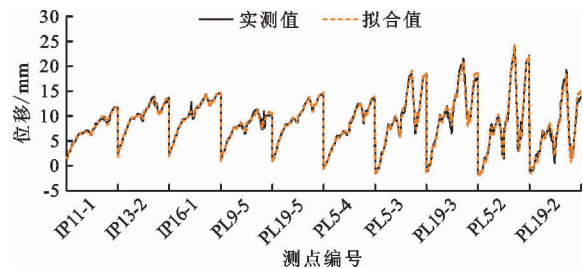


图 8 坝体四区面板随机系数模型实测值及拟合值过程线

表 1 随机系数模型的相关系数 R 和标准差 S

区域	测点编号	R	S	区域	测点编号	R	S	区域	测点编号	R	S
一区	PL16-4	0.987	1.399	二区	PL13-4	0.999	0.515	四区	PL5-3	0.998	0.368
	PL13-5	0.999	0.382		PL9-4	0.998	0.456		PL19-3	0.998	0.868
	PL16-5	0.999	0.373		PL19-4	0.998	0.393		PL19-5	0.999	0.151
	PL11-5	0.999	0.342	三区	PL11-2	0.984	0.486		IP13-2	0.994	0.297
	PL16-3	0.999	0.369		PL9-3	0.957	2.344		IP16-1	0.998	0.223
	PL11-3	0.992	1.470		PL16-2	0.987	1.302		PL5-4	0.997	0.285
二区	PL13-3	0.999	0.611	四区	PL19-2	0.983	0.952		IP11-1	0.997	0.206
	PL11-4	0.992	1.316		PL5-2	0.997	0.523		PL9-5	0.984	0.442

表 2 坝体四区面板随机系数模型结果

四区随机系数面板模型						
Random – coefficients regression		Number of obs: 2740				
Group variable: _j		Number of groups: 10				
		Obs per group: min = 274 avg = 274.0 max = 274				
		Wald chi2(10): 918.34				
		Prob > chi2: 0.0000				
var	Coef.	Std. Err.	z	P > z	[95% Conf. Interval]	
x1	– 355.6792	79.26087	– 4.49	0.000	– 511.0276	– 200.3307
x2	726.7902	151.1626	4.81	0.000	430.517	1023.063
x3	– 641.6031	129.7297	– 4.95	0.000	– 895.8687	– 387.3375
x4	218.9433	43.41306	5.04	0.000	133.8552	304.0313
x5	2.302854	0.768735	3.00	0.003	0.7961616	3.809547
x6	2.686276	1.811475	1.48	0.138	– 0.86415	6.236702
x7	– 0.8040001	0.2384235	– 3.37	0.001	– 1.271302	– 0.336985
x8	– 0.306704	0.1847532	– 1.66	0.097	– 0.6688136	0.0554056
x9	– 0.1351805	0.0612589	– 2.21	0.027	– 0.2552458	– 0.0151153
x10	0.1122696	0.0371336	3.02	0.002	0.0394891	0.1850502
_cons	62.74983	15.05091	4.17	0.000	33.25059	92.24906
Test of parameter constancy: chi2(33) = 89818.17 Prob > chi2 = 0.0000						

表 3 一区各测点位移实测值和计算值

日期	PL16 – 4		PL13 – 5		PL16 – 5		PL11 – 5		PL16 – 3	
	mm									
	实测值	预测值	实测值	预测值	实测值	预测值	实测值	预测值	实测值	预测值
2015-09-22	30.18	30.01	33.14	32.89	30.55	31.04	29.88	32.74	25.80	28.77
2015-09-23	30.17	30.06	33.21	32.95	30.65	31.09	30.03	32.81	26.02	28.87
2015-09-24	30.20	30.12	33.24	33.03	30.68	31.16	30.09	32.89	26.07	29.00
2015-09-25	30.21	30.14	33.25	33.04	30.69	31.17	30.10	32.89	26.13	29.01
2015-09-26	30.33	30.24	33.40	33.17	30.78	31.27	30.27	33.04	26.38	29.24
2015-09-27	30.38	30.30	33.46	33.25	30.84	31.34	30.37	33.14	26.51	29.40
2015-09-28	30.43	30.33	33.47	33.28	30.82	31.37	30.33	33.16	26.45	29.44

表 4 一区各监测点变形风险率

日期	PL16 – 4	PL13 – 5	PL16 – 5	PL11 – 5	PL16 – 3
2015-09-22	0.0031	0.0019	0.0026	0.0005	0.0028
2015-09-23	0.0031	0.0019	0.0028	0.0007	0.0028
2015-09-24	0.0032	0.0021	0.0027	0.0007	0.0027
2015-09-25	0.0031	0.0021	0.0027	0.0006	0.0028
2015-09-26	0.0031	0.0019	0.0026	0.0006	0.0027
2015-09-27	0.0031	0.0018	0.0027	0.0005	0.0027
2015-09-28	0.0032	0.0019	0.0026	0.0006	0.0028

表 5 二区各监测点变形风险率

日期	PL11 – 3	PL13 – 3	PL11 – 4	PL13 – 4
2015-09-22	0.0031	0.0006	0.0032	0.0006
2015-09-23	0.0031	0.0005	0.0033	0.0008
2015-09-25	0.0031	0.0006	0.0033	0.0009
2015-09-26	0.0032	0.0007	0.0032	0.0007
2015-09-27	0.0031	0.0008	0.0032	0.0007
2015-09-28	0.0032	0.0007	0.0033	0.0007

表 6 三区各监测点变形风险率

日期	PL9 - 4	PL19 - 4	PL11 - 2	PL9 - 3	PL16 - 2
2015-09-22	0.0027	0.0029	0.0031	0.0032	0.0032
2015-09-23	0.0028	0.0031	0.0031	0.0032	0.0032
2015-09-24	0.0028	0.0031	0.0032	0.0032	0.0031
2015-09-25	0.0029	0.0029	0.0032	0.0031	0.0031
2015-09-26	0.0028	0.0029	0.0031	0.0032	0.0031
2015-09-27	0.0029	0.0029	0.0032	0.0031	0.0031
2015-09-28	0.0028	0.0029	0.0032	0.0032	0.0032

4.5 分区及整体变形风险率计算

由于监测点较多,以一区中 PL13 - 5 和 PL16 - 5

两测点为例,根据计算出的 2013 年 6 月 16 日 - 2015 年 9 月 28 日两测点各 274 组变形风险率作为样本总数,以样本总数中全部小于某一日期对应的两测点风险率数值的组数与样本总数之比来近似表示两测点的真实综合风险率。得出的真实综合风险率与各 Copula 联合分布函数计算得出的风险率见表 8,各 Copula 联合分布函数 RMSE 准则统计检验结果见表 9。由表 8、9 可以看出,以 Frank 分布函数计算出的两测点联合风险率与真实概率最接近,同时 RMSE 准则统计检验结果中的 Frank Copula 数值最小,因而确定最佳二维联合概率分布函数为 Frank 分布函数。

表 7 四区各监测点变形风险率

日期	PL19 - 2	PL5 - 2	PL5 - 3	PL19 - 3	PL19 - 5	IP13 - 2	IP16 - 1	PL5 - 4	IP11 - 1	PL9 - 5
2015-09-22	0.0024	0.0003	0.0007	0.0028	0.0026	0.0009	0.0008	0.0017	0.0003	0.0028
2015-09-23	0.0025	0.0003	0.0005	0.0028	0.0027	0.0009	0.0008	0.0017	0.0004	0.0028
2015-09-24	0.0024	0.0003	0.0005	0.0028	0.0028	0.0011	0.0009	0.0018	0.0005	0.0027
2015-09-25	0.0026	0.0003	0.0006	0.0028	0.0028	0.0011	0.0009	0.0019	0.0004	0.0028
2015-09-26	0.0025	0.0002	0.0007	0.0029	0.0029	0.0011	0.0011	0.0018	0.0003	0.0029
2015-09-27	0.0025	0.0002	0.0006	0.0029	0.0029	0.0011	0.0011	0.0021	0.0002	0.0028
2015-09-28	0.0025	0.0002	0.0006	0.0029	0.0028	0.0011	0.0012	0.0021	0.0002	0.0029

表 8 一区 PL13 - 5 和 PL16 - 5 测点各 Copula 联合分布函数风险率计算结果对比

PL13 - 5	PL16 - 5	真实概率	Frank	Gumbel	Clayton	t	Gaussian
0.0019	0.0026	0.4964	0.4982	0.5148	0.5218	0.5151	0.5189
0.0019	0.0028	0.4891	0.5064	0.5237	0.5303	0.5238	0.5275
0.0021	0.0027	0.4161	0.4264	0.4360	0.4440	0.4376	0.4413
0.0021	0.0027	0.4197	0.4298	0.4397	0.4476	0.4412	0.4449
0.0019	0.0026	0.4781	0.4686	0.4819	0.4889	0.4827	0.4863
0.0018	0.0027	0.4124	0.4114	0.4197	0.4281	0.4217	0.4254
0.0019	0.0026	0.3650	0.3698	0.3741	0.3803	0.3763	0.3791

表 9 Copula 联合分布函数 RMSE 检验指标结果

Frank Copula	Gumbel Copula	Clayton Copula	t - Copula	Gaussian Copula
0.0402288	0.0424739	0.0425395	0.0424585	0.0426847

由于监测点数目较多,计算出的 Frank Copula 函数参数较多,仍以一区中 PL13 - 5 和 PL16 - 5 两

测点为例,代入极大似然估计求出的参数 $\hat{\theta}$,其二维 Frank Copula 表达式为:

$$C(u_1, v_1, \theta) = -\frac{1}{2.8568} \ln \left[1 - \frac{(1 - e^{-2.8568u_1})(1 - e^{-2.8568v_1})}{1 - e^{-2.8568}} \right]$$

(23)

对于特高拱坝变形风险分析,PL13 - 5 或 PL16 - 5 变形风险率异常均为大坝变形异常。因

此,考虑 PL13 - 5 和 PL16 - 5 变形风险的特高拱坝多变形变量风险率计算公式为:

$$P_i = P(X_1 \leq u_1 \cup X_2 \leq v_1) = u_1 + v_1 - \frac{1}{2.8568} \ln \left[1 - \frac{(1 - e^{-2.8568u_1})(1 - e^{-2.8568v_1})}{1 - e^{-2.8568}} \right]$$

(24)

式中: u_i 、 v_i 分别为 PL13-5 和 PL16-5 于监测日期当天的风险率。

通过上述算法根据实例工程特高拱坝测点谱系聚类树状图(图 4)依次按照聚类流程算出各分区及整体变形风险率见表 10。

表 10 各区及整体变形风险率

日期	一区	二区	三区	四区	整体
2015-09-22	9.03×10^{-13}	3.28×10^{-10}	2.72×10^{-10}	0	0
2015-09-23	1.36×10^{-12}	3.76×10^{-10}	3.02×10^{-10}	0	0
2015-09-24	1.44×10^{-12}	3.88×10^{-10}	3.02×10^{-10}	0	0
2015-09-25	1.24×10^{-12}	5.07×10^{-10}	2.83×10^{-10}	0	0
2015-09-26	1.05×10^{-12}	4.61×10^{-10}	2.74×10^{-10}	0	0
2015-09-27	8.57×10^{-13}	5.10×10^{-10}	2.83×10^{-10}	0	0
2015-09-28	1.12×10^{-12}	4.75×10^{-10}	2.91×10^{-10}	0	0

5 结 论

依据原型监测资料,基于面板数据理论和 Copula 函数,构建了特高拱坝整体实时风险率分析模型,主要工作总结如下:

(1) 依据构建的同时考虑时间效应和横截面序列的面板数据模型,根据特高拱坝监测点变形监测值之间的相关性和差异性对其聚类分区,建立面板数据随机系数模型对每一区测点进行建模,确定依据大坝长序列立体监测数据建立变形实测效应量与风险率的函数关系。

(2) 基于 Copula 函数,考虑变形分区监测点之间的相关性及其不同分区间的协同性,分析特高拱坝各分区变形风险率和整体变形风险率,建立特高拱坝实时风险率评估模型。实例分析表明,所构建的模型能有效分析特高拱坝各分区变形风险率及整体变形风险率,能客观刻画特高拱坝整体风险率变化的基本规律。

参考文献:

[1] 赵二峰,顾冲时. 混凝土坝长效服役性态健康诊断研究述评[J]. 水力发电学报,2021,40(5):22-34.
[2] 牛景太,杨光,雷鹏. 基于权重熵的碾压混凝土坝温度场预警方法[J]. 水资源与水工程学报,2020,31(6):126-131.

[3] 胡国华,夏军. 风险分析的灰色-随机风险率方法研究[J]. 水利学报,2001,32(4):1-6.
[4] 陈在铁,任青文. 计算机模拟技术在大坝失效概率研究中的应用[J]. 水力发电,2008,34(1):92-94.
[5] 徐强,陈健云,李静. 基于拉格朗日乘数法计算大坝结构可靠度[J]. 工程力学,2009,26(11):108-113.
[6] 邓建华,胡雅婷,孟珍珠,等. 基于 Copula 函数的混凝土特高拱坝变形风险量化模型[J]. 水电能源科学,2020,38(6):59-62.
[7] 王子成,许后磊,赵志勇,等. 特高拱坝动态安全风险分析系统研发及应用[J]. 水利水运工程学报,2020,(1):112-118.
[8] 顾冲时,吴中如. 大坝与坝基安全监控理论和方法及其应用(第1版)[M]. 南京:河海大学出版社,2006.
[9] 何晓群. 多元统计分析(第2版)[M]. 北京:中国人民大学出版社,2008.
[10] 李因果,何晓群. 面板数据聚类方法及应用[J]. 统计研究,2010,27(9):73-79.
[11] 张立军,彭浩. 面板数据加权聚类分析方法研究[J]. 统计与信息论坛,2017,32(4):21-26.
[12] BALTAGI B H. Econometric analysis of panel data[M]. New Jersey: John Wiley & Sons, 2009.
[13] 程慧燕,赵艳伟,朱道元. 半变系数模型改进的轮廓最小二乘估计[J]. 数学的实践与认识,2017,47(15):249-253.
[14] 张颖,曹连英. 时空变系数半参数模型的轮廓最小二乘估计[J]. 哈尔滨师范大学自然科学学报,2018,34(5):1-4.
[15] 徐颖,傅志敏. 复合土工膜防渗土石坝渗流安全监控指标拟定[J]. 水电能源科学,2021,39(1):83-86.
[16] 韩云霞,马义中,欧阳林寒,等. 数据污染下的稳健设计及最优参数的置信区间估计[J]. 工业工程与管理,2019,24(1):64-71+86.
[17] 左伯庆. 拱坝安全监测与预警方法研究[D]. 济南:山东大学,2018.
[18] 张流洋,马义中,任鸣鸣,等. 基于均方根误差建模的多响应稳健参数设计[J]. 统计与决策,2020,36(6):20-25.
[19] 代剑锋. 基于 Copula 的整体风险度量模型构建[D]. 武汉:华中科技大学,2019.
[20] 宋松柏. Copula 函数在水文多变量分析计算中的问题[J]. 人民黄河,2019,41(10):40-47+57.
[21] 陈望学,杨瑞,谢民育. 完美和非完美中位数排序集抽样下刻度参数的极大似然估计[J]. 应用数学学报,2020,43(3):572-583.