DOI:10.3880/j.issn.1006-7647.2023.02.004

基于分位数回归的升船机变形监控模型构建方法

杨晨昊[1,2],郑东健[1,2]

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

摘要:为深入分析升船机变形影响因素,提出了一种基于分位数回归的升船机变形监控模型构建方法。该方法根据升船机的结构特点,将温度、前期上游水位均值等因素引入候选影响因子集,采用自适应弹性网络分位数回归对影响因子进行筛选,建立各分位数下的回归模型,并根据拟合的良好性和检验的有效性原则选出最优的升船机变形监控模型。实例验证结果表明:相对于常规的逐步回归模型,本文方法构建的最优模型的预测精度波动性小,具有较强的稳定性,同时具有良好的长期预测能力。

关键词:升船机变形;变形预测;分位数回归;自适应弹性网络;模型优选方法

中图分类号:TV698 **文献标志码**:A **文章编号**:1006-7647(2023)02-0027-06

Construction method for ship lift deformation monitoring model based on quantile regression//YANG Chenhao[1,2], ZHENG Dongjian[1,2] (1. *College of Water Conservancy and Hydropower Engineering*, *Hohai University*, *Nanjing* 210098, *China*; 2. *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering*, *Hohai University*, *Nanjing* 210098, *China*)

Abstract: To deeply analyze the influence factors of ship lift deformation, a method for building ship lift deformation monitoring model based on quantile regression is proposed. Based on the structural characteristics of the ship lift, the influence factors such as air temperature and mean value of upstream water level in early stages are introduced into the candidate influence factor set. Adaptive elastic net penalized quantile regression is used for choosing the influence factors of ship lift deformation and building the quantile regression model. The optimal model can be obtained based on the principle of good fitting and validity check. Example verification shows that compared with traditional stepwise regression model, the proposed optimized model has low volatility for prediction accuracy with strong stability, indication a good ability for long term prediction.

Key words: ship lift deformation; deformation prediction; quantile regression; adaptive elastic net; model optimization method

对水工建筑物进行安全监控是保障水利工程安全运行的主要手段[1]。升船机在水利工程中发挥着通航功能,它的安全不仅关乎水利工程的整体安全,也保障了江河的通航能力,具有积极的经济作用[2]。随着我国升船机数量的增加[3],探究升船机的安全监控方法具有重要的理论价值和现实意义。

变形是掌握和分析升船机工作性态的重要因素[1],常用的水工建筑物变形预测模型可分为最小二乘回归类和智能算法类[4-5],但这两类方法都存在一定的局限性:最小二乘回归类假设随机误差项服从正态分布,对异常值敏感[6];智能算法类计算复杂度高,而且不利于工程师对监测量物理成因进行把握[7]。

自适应弹性网络(adaptive elastic net)可以较好地对变量进行筛选,从而可以避免因候选影响因子过多而影响模型的可解释性及计算复杂度。分位数回归数学形式简单,无须假设随机误差的分布形式,适用于异方差情况,稳健性强,且更关注因变量局部信息,因此更适合挖掘数据信息[6]。目前,分位数回归已在经济学、水文学等[8-9]专业得到了较好的应用,而将分位数回归应用于升船机变形监控模型尚未见报道。

基金项目:国家自然科学基金面上项目(52179128)

作者简介:杨晨昊(1998—),男,硕士研究生,主要从事水工结构安全监测研究。E-mail:1418132626@qq.com

通信作者:郑东健(1965—),男,教授,博士,主要从事水工结构安全监测研究。E-mail:zhengdj@hhu.edu.cn

为深入分析升船机变形的影响因素,本文提出了一种基于分位数回归的升船机变形监控模型构建方法。该方法从升船机的变形特点出发,用自适应弹性网络分位数回归方法构建回归模型,而后从中选出最优模型,并在工程实例应用中取得了较好的预测结果。

1　基本理论

1.1　分位数回归

分位数回归由 Koenker 等[10]于 1978 年首次提出,其与传统的均值回归仅考察自变量和因变量均值之间的关系不同,分位数回归更关心因变量在不同分位数下与自变量的回归关系,更有利于挖掘数据蕴含的信息。同时,建模时分位数回归无须考虑随机误差的分布形式,稳健性强,适用范围广。

对于简单多元线性模型,分位数回归表达式与均值回归相同:

$$\boldsymbol{Y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon} \tag{1}$$

其中
$$\boldsymbol{Y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$$
$$\boldsymbol{X}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)^{\mathrm{T}}$$
$$\boldsymbol{X}_i=(1,x_{i1},x_{i2},\cdots,x_{ip})^{\mathrm{T}}$$
$$\boldsymbol{\beta}=(\beta_0,\beta_1,\beta_2,\cdots,\beta_p)^{\mathrm{T}}$$

式中:$\boldsymbol{Y}$ 为观测样本中的被解释变量;$\boldsymbol{X}$ 为观测样本中的解释变量,是 $n\times(p+1)$ 维矩阵;$\boldsymbol{\beta}$ 为解释变量的系数向量,其中 β_0 为截距系数;$\boldsymbol{\varepsilon}$ 为随机误差。

对于线性模型,分位数回归可被视作条件分位数函数,可表示为

$$Q_{\boldsymbol{Y}}(\tau\mid\boldsymbol{X})=\boldsymbol{X\beta}(\tau) \tag{2}$$

其中 $\boldsymbol{\beta}(\tau)=(\beta_0(\tau),\beta_1(\tau),\beta_2(\tau),\cdots,\beta_p(\tau))^{\mathrm{T}}$

式中:τ 为分位数,$\tau\in(0,1)$;$Q_{\boldsymbol{Y}}(\tau\mid\boldsymbol{X})$ 为 $\boldsymbol{X}$ 条件时 $\boldsymbol{Y}$ 的 τ 分位数下的条件分位数函数;$\boldsymbol{\beta}(\tau)$ 为 τ 分位数下的系数向量,其中 $\beta_0(\tau)$ 为 τ 分位数下的截距系数。

与最小二乘回归的损失函数不同,分位数回归将加权误差绝对值之和定义为损失函数,并通过最小化损失函数估计回归系数,表达式为

$$\hat{\boldsymbol{\beta}}(\tau)=\arg\min_{\boldsymbol{\beta}(\tau)}\frac{1}{n}\sum_{i=1}^{n}\rho_\tau(y_i-\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}(\tau)) \tag{3}$$

式中$\rho_\tau(\cdot)$为检查函数,表达式为

$$\rho_\tau(u)=(\tau-I(u<0))u \tag{4}$$

式中:$I(t)$ 为示性函数,t 为真时,$I(t)=1$;t 为假时,$I(t)=0$。

1.2　自适应弹性网络分位数回归参数估计

在弱正则条件下,自适应弹性网络具备 Oracle 性质,在仿真试验中比其他惩罚项(如自适应 Lasso)表现更优[11]。同时,自适应弹性网络包含 L2 正则项,可以一定程度上消除变量间的多重线性相关性,从而增强模型的可解释性。因此,本文选择自适应弹性网络与分位数回归结合来处理升船机变形影响因子的选择问题。

自适应弹性网络在弹性网络的 L1 正则项中加入权重,其与分位数回归结合后的参数估计表达式为

$$\hat{\boldsymbol{\beta}}(\tau)_{\mathrm{AEN}}=\arg\min_{\boldsymbol{\beta}(\tau)}\frac{1}{n}\sum_{i=1}^{n}\rho_\tau(y_i-\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}(\tau))+\lambda_1\sum_{j=1}^{p}w_j\left|\beta_j(\tau)\right|+\lambda_2\sum_{j=1}^{p}\beta_j(\tau)^2 \tag{5}$$

其中 $$w_j=\left(\left|\hat{\beta}_j(\tau)_{\mathrm{EN}}+\frac{1}{n}\right|\right)^{\gamma}$$

式中:λ_1 为 L1 正则项的系数;λ_2 为 L2 正则项的系数;w_j 为权重;γ 为正常数,根据文献[11-12]取 1;$\hat{\beta}_j(\tau)_{\mathrm{EN}}$ 为用弹性网络分位数回归得到的初始值[12]。

分位数回归的损失函数具有非光滑性,给数学计算带来了巨大的挑战,故考虑到计算效率等因素,选择 Gu 等[12]于 2018 年提出的 scdADMM 算法对自适应弹性网络分位数回归模型进行参数估计。

2　监控模型构建方法

2.1　升船机变形候选影响因子集

研究水工建筑物的长期变形时,主要考虑水荷载、温度荷载和时效作用的影响[1],故本文从这 3 类影响因子出发去探讨影响升船机变形的因素,从而为升船机变形监控模型构建候选影响因子集。

2.1.1　水压分量

大坝变形监控模型通常以上下游水位的前 q 次方项为水荷载因子。对于升船机,其上闸首为拦河坝的一部分,上游水荷载会通过坝体传递给升船机,所以可将上游水位的前 3 次方项(重力坝)或前 4 次方项(拱坝)以及下游水位前 3 次方项引入候选因子集合。

升船机与混凝土坝体结构不同,升船机顺河向长度远大于坝体,升船机塔楼这一主体承重建筑通常在上工作门段之后,故坝体传递给升船机塔楼的上游水压荷载可能存在滞后效应,故可将前期上游水位平均值引入候选因子集合,时间范围保守选择在 30 d 以内。

综上,引入候选因子集合的水压分量因子如下:

$$\delta_{\mathrm{w}}=\{H_{\mathrm{u}},H_{\mathrm{u}}^2,\cdots,H_{\mathrm{u}}^q,H_{\mathrm{d}},H_{\mathrm{d}}^2,H_{\mathrm{d}}^3,\overline{H}_{1-3},\cdots,\overline{H}_{m-n}\} \tag{6}$$

式中:H_{u} 为上游水位,重力坝时 $q=3$,拱坝时 $q=4$;H_{d} 为下游水位;$\overline{H}_{m-n}$为观测日前第 m 天至第 n 天上游水位均值,m 和 n 可根据实际工程情况调整。

2.1.2 温度分量

温度荷载是影响升船机塔楼变形的主要荷载之一,其变形主要与温度有关,具有明显的年周期性[13-14]。水工建筑物变形监控模型中,一般以两组谐波因子表示具有明显周期变化规律的温度因子,它们有明确的物理意义,故对于升船机塔楼,可用两组谐波因子作为周期性温度因子。

气温并非是一个严格的周期函数,它会受到许多随机因素影响而表现出随机性[15]。升船机塔楼温度变形的有限元仿真[16]表明,短期气温变化会对升船机变形造成影响,如寒潮等。因此,考虑到升船机为高耸薄壁结构,温度场易受气温影响,故本文引入“随机气温因子”以考察其是否会对升船机变形带来影响。对于“随机气温因子”,王建等[15]已给出了一个合理的定义与计算方法,并用实例验证了可行性。

一天内气温不可避免地会出现上升下降现象,这种现象也是一个潜在的温度影响因子。根据自然规律,一天中通常6:00气温最低,14:00气温最高,故定义日变温作为一个温度因子,即当监测时刻6:00<t'≤14:00时,日变温为$T_{t'}-T_6$;当14:00<t'≤6:00时,日变温为$T_{t'}-T_{14}$,注意此时的6:00指第2天6:00。

综上,引入候选因子集合的温度分量因子如下:

$$\delta=\left\{\sin\frac{2\pi it}{365}-\sin\frac{2\pi it_0}{365},\cos\frac{2\pi it}{365}-\cos\frac{2\pi it_0}{365},\right.$$
$$\left.T_{r,t},T_{r,t-1},\cdots,T_{r,t-k},T_{\mathrm{d}}\right\} \tag{7}$$

式中:t为监测日距始测日天数,因升船机建成时间一般要晚于大坝等建筑物,故本文将始测日定为升船机验收当日;t_0为建模系列第一个监测日距始测日天数;$i=1$、2;$T_{r,t}$、$T_{r,t-1}$、$\cdots$、$T_{r,t-k}$为随机气温因子,具体计算公式见文献[15];k为滞后天数;T_{d}为日变温,每日6:00的气温减去前一天14:00的气温。

2.1.3 时效分量

同大坝等水工建筑物一样,升船机时效分量的组成成分复杂,是一个可以综合反映升船机钢筋混凝土徐变、基岩变形等因素的综合性影响因子。故升船机变形的时效因子同样可采用多项式形式和对数函数形式表达。引入候选因子集合的时效分量因子如下:

$$\delta_\theta=\{\theta-\theta_0,\ln\theta-\ln\theta_0\} \tag{8}$$

式中:θ为t除以100;θ_0为t_0除以100。

2.2 模型优选方法

模型优选指从一系列关注局部信息的分位数回归模型中,选出一个最能代表整体信息的模型作为预测模型。黄耀英等[7]给出了最佳统计模型优选原则:拟合的良好性、模型的简单性和检验的有效性。本文每个回归模型的影响因子均是经过自适应弹性网络筛选得到,已具备稀疏性,并且具有一定的物理意义,所以本文优选模型时只需服从拟合的良好性和检验的有效性这2个原则即可。

模型在训练集上的拟合良好性或在验证集上的有效性均可通过决定系数(R^2)、均方根误差(RMSE)和平均绝对误差(MAE)进行综合评价。对于分位数回归,还可用伪决定系数(pseudo R-squared,PR2)作为评价指标,PR2计算方法可参考文献[17]。

以某一测点为例,测点的评价指标共计8个,需对它们进行归一化处理使其在同一量级。归一化前,需剔除R^2、PR2指标异常(即值小于0.5)的模型。值越大越优的指标用式(9)归一化,值越小越优的指标用式(10)归一化。

$$\check{f}_j=\frac{f_j-f_{\min}}{f_{\max}-f_{\min}} \tag{9}$$

$$\check{f}_j=\frac{f_{\max}-f_j}{f_{\max}-f_{\min}} \tag{10}$$

式中:f_j为各分位数下的某一评价指标;$f_{\min}$、$f_{\max}$分别为f_j的最小和最大值。

归一化后的指标无须再进行标准化处理,因为上述归一化处理已相当于经公式$y=x$进行了标准化,避免了文献[7]中评价指标的高值或低值在标准化后区分度不明显的问题。

回归模型的最终评价标准为

$$M_{\mathrm{final}}=\sum_{c=1}^{4}\lambda_cF_c+\sum_{d=1}^{4}\mu_dV_d \tag{11}$$

式中:F_1、F_2、F_3、F_4分别为训练集上某分位数模型归一化后的R^2、PR2、RMSE和MAE;V_1、V_2、V_3、V_4、分别为验证集上某分位数模型归一化后的R^2、PR2、RMSE和MAE;λ_c、μ_d分别为相应权重,本文均取1/8。

计算得到各回归模型的M_{final}后,选择M_{final}最大的模型为该测点最优模型。

2.3 模型构建流程

以单个测点为例,基于分位数回归的升船机变形预测模型建模流程可总结为如下5个步骤:

步骤1 将变形监测数据与候选影响因子监测数据划分为训练集、验证集和测试集,并进行归一化处理。

步骤2 确定所需的分位数,以某一分位数为例,将训练集输入弹性网络分位数回归模型,并在固定L2正则项系数λ_2后[18],用k折交叉验证法搜索最优的L1正则化系数λ_1,得到该分位数下自适应弹性网络分位数回归模型的权重。

步骤 3 同样以某一分位数为例,将训练集输入自适应弹性网络分位数回归模型,令 L2 正则项系数 λ_2 与该分位数在步骤 2 中的 λ_2 值一致,并用 k 折交叉验证法搜索最优的 L1 正则化系数 λ_1,以此类推,就能得到各分位数下的自适应弹性网络分位数回归数学模型以及变量筛选结果。

步骤 4 计算各回归模型在训练集和验证集上的评价指标,并将所得评价指标输入模型优选方法,最终确定该测点的最优模型。

步骤 5 在测试集上考察最终模型的预测结果,验证本文所提方法的合理性和可行性。

3 模型验证

3.1 工程概况

某水利工程拦河坝为混凝土重力坝,升船机为全平衡钢丝绳卷扬式垂直升船机。升船机主要建筑物塔楼为高耸薄壁结构,由 8 座钢筋混凝土承重塔柱和 2 座交通楼构成,并在塔楼的中心线左右两侧呈对称布置。为监测升船机水平位移变化情况,共计布置了 10 条正垂线,分别布置在升船机塔楼的下平衡段、交通段中部、下提升段 E 形结构等部位,共计 20 个测点,可同时监测升船机塔楼上下游方向(顺河向)和左右岸方向(横河向)的水平位移。

对数据的可靠性和完整性进行分析后,选择升船机塔楼的左右岸方向绝对水平位移的 13 个测点进行建模。建模序列为 2018 年 1 月 4 日至 2020 年 12 月 31 日,共计 1 093 个测值,监测时间为每天 6:00,其中训练集为前 800 个测值,验证集为序列中的第 801 ~900 个测值,测试集则为后 193 个测值。

3.2 候选影响因子组成

候选影响因子集合共包含 20 个影响因子,并按如上顺序将因子分别编号为 1 ~20。

a. 水荷载因子集为

$$\delta_{1-10}=\{H_u,H_u^2,H_u^3,H_d,H_d^2,H_d^3,\overline{H}_{1-4},\overline{H}_{5-9},\overline{H}_{10-15},\overline{H}_{16-30}\} \tag{12}$$

b. 温度因子集为

$$\delta_{11-18}=\left\{\sin\frac{2\pi it}{365}-\sin\frac{2\pi it_0}{365},\cos\frac{2\pi it}{365}-\cos\frac{2\pi it_0}{365},T_{r,t},T_{r,t-1},\cdots,T_{r,t-2},T_d\right\} \tag{13}$$

c. 时效因子集为

$$\delta_{19-20}=\{\theta-\theta_0,\ln\theta-\ln\theta_0\} \tag{14}$$

为初步验证选择的合理性,在测点 PL10-2 与式(15)常规影响因子集做对比(图 1)。由图 1 可知,本文候选影响因子集极大改善了常规影响因子集的拟合效果,较好地拟合了实测值全局与局部变化趋势,常规影响因子集则拟合不出实测值局部的波动情况。

$$\left\{H_u,H_u^2,H_u^3,H_d,H_d^2,H_d^3,\sin\frac{2\pi it}{365}-\sin\frac{2\pi it_0}{365},\cos\frac{2\pi it}{365}-\cos\frac{2\pi it_0}{365},\theta-\theta_0,\ln\theta-\ln\theta_0\right\} \tag{15}$$

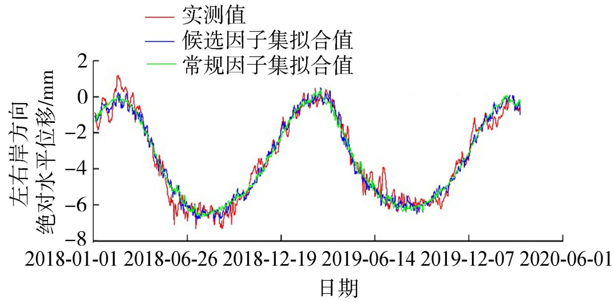
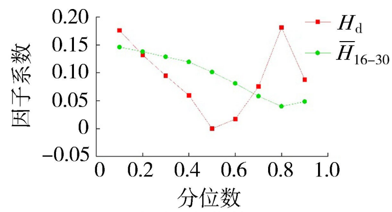


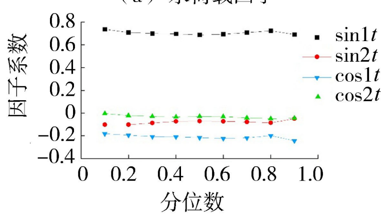
图 1 测点 PL10-2 测值拟合结果对比

3.3 影响因子筛选结果与分析

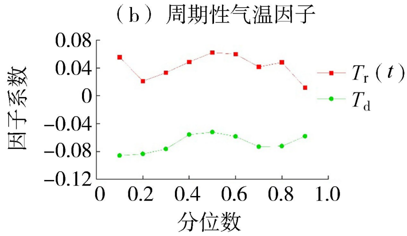
以 0.1 为间隔,取 0.1 ~0.9 共计 9 个分位数分析不同分位数下测点影响因子系数的变化规律。以测点 PL10-2 为例,筛选后的主要影响因子系数变化如图 2 所示。



(a) 水荷载因子



(b) 周期性气温因子



(c) 随机气温因子和日变温

图 2 测点 PL10-2 影响因子系数变化情况

由图 2 可知,下游水位(H_d)的系数随分位数增大呈先减小后增大的趋势,即对变形影响先减小后增大;前期上游水位均值($\overline{H}_{16-30}$)的系数随分位数增大呈先减小趋势,说明位移越大其影响越小;4 个周期温度因子(两组谐波因子)的系数随分位数变化不断调整,说明对于不同变形大小,周期性规律略有差异;随机气温因子($T_{r,t}$)的系数变化有一定波动性;日变温(T_d)的系数为负并随分位数增大呈先增

大后减小趋势，说明其在变形值较大或较小时影响增大。除图 2 中的影响因子外，在个别分位数下，上游水位前期平均值（$\overline{H}_{1-4}$，$\overline{H}_{5-9}$）、时效因子（$\theta-\theta_0$）也被选中，可称其为次要因子。这可能是因为它们仅在某些变形分位数下才会对变形产生一些影响。可以看出，分位数回归可以分析升船机变形在不同分位数时的因子作用情况，与均值回归相比能得到更多信息。

以各测点不同分位数下共 117 个回归模型的影响因子被选情况为依据，采用被选百分比为指标，可得到以下规律：

a. 候选影响因子除水位次方项（H_u^2，H_u^3，H_d^2，H_d^3）以外的影响因子均会不同程度影响变形。对于全部测点，这些因子可以被分为必要因子（大于 70%）、次要因子（30% ~70%）和非必要因子（小于 30%）。

b. 必要因子定义为各测点各分位数下被选中频率较高的因子，包括周期温度因子（两组谐波因子）、随机气温因子（$T_{r,t}$，$T_{r,t-2}$）、观测日前第 16—30 天上游水位均值（$\overline{H}_{16-30}$），它们对升船机变形中起到了主要影响作用。

c. 次要因子定义为仅在某些测点变形中发挥了较大作用，包括随机气温因子（$T_{r,t-1}$）、前期上游水位均值（$\overline{H}_{1-4}$，$\overline{H}_{5-9}$，$\overline{H}_{10-15}$）和时效因子（$\theta-\theta_0$）。

d. 非必要因子指仅在一个测点或某些测点的某几个分位数下有作用，且作用较小，包括上下游水位因子（H_u，H_d）、日变温（T_d）、时效因子（$\ln\theta-\ln\theta_0$）。

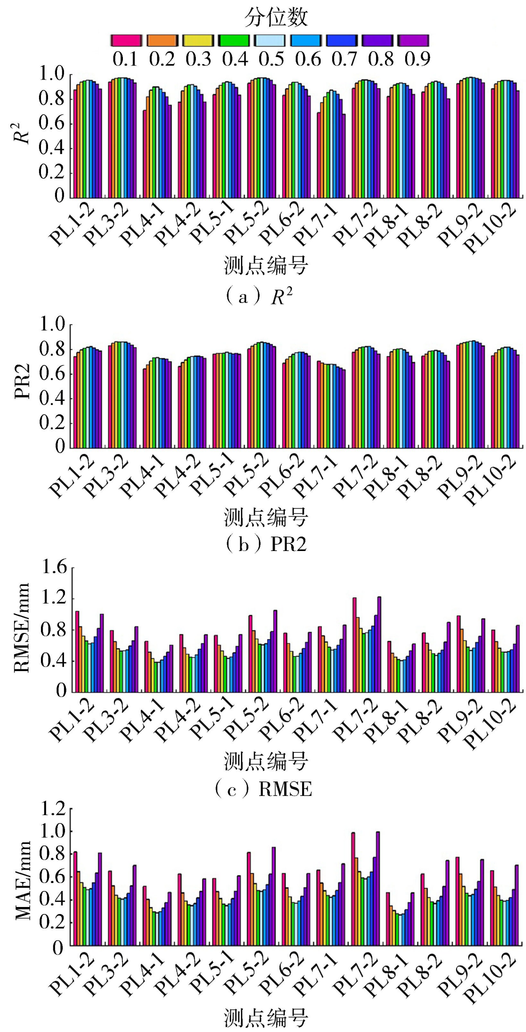
e. 对 13 个测点模型中水荷载因子、温度因子和时效因子对变形影响大小定量分析，可得温度因子为影响升船机变形的主要因素，水荷载因子次之，时效因子影响则最小。

3.4　最优模型及验证

以训练集为例，各测点各分位数下回归方程的 R^2、PR2、RMSE 和 MAE 如图 3 所示，大多数测点的回归方程拟合效果较好，并以 0.5、0.6 分位数时的回归模型拟合效果最佳。所有测点的 R^2 均随分位数增大呈先增大后减小的变化趋势；PR2 随分位数增大，变化趋势与 R^2 类似；所有测点的 RMSE 和 MAE 则随分位数增大呈先减小后增大的变化趋势。

实际水利工程的定检工作中，逐步回归法使用较多，这表明其能经受实际工程考验，有一定的实用性和说服力，故本文选择逐步回归模型与最优模型在测试集上进行对比，结果见表 1，其中最优模型选择结果以分位数表示。

由表 1 可知，对于大部分测点，最优模型的 R^2、RMSE 和 MAE 均要优于逐步回归模型，对于逐步回归模型表现异常的测点，最优模型也具有相对较好的表现；对于少部分测点，最优模型的评价指标比逐步回归模型要差，但差异不大；整体上，最优模型的预测精度波动性小，具有较强的稳定性。同时，本例中测试集与验证集数据长度之和为 293 d，时长已接



（a）R^2

（b）PR2

（c）RMSE

（d）MAE

图 3　训练集上各回归模型评价指标

表 1　模型评价指标对比

测点编号	最优模型评价指标				逐步回归模型评价指标		
	分位数	R^2	RMSE/mm	MAE/mm	R^2	RMSE/mm	MAE/mm
PL1-2	0.5	0.796	0.852	0.751	0.355	1.5164	1.403
PL3-2	0.5	0.958	0.501	0.424	0.960	0.488	0.373
PL4-1	0.2	0.815	0.228	0.187	0.593	0.339	0.270
PL4-2	0.5	0.699	0.409	0.326	−0.676	0.966	0.891
PL5-1	0.8	0.749	0.494	0.393	−2.377	1.814	1.779
PL5-2	0.6	0.952	0.576	0.474	0.515	1.823	1.735
PL6-2	0.5	0.710	0.519	0.428	0.745	0.487	0.377
PL7-1	0.6	0.703	0.506	0.382	0.550	0.623	0.458
PL7-2	0.6	0.936	0.653	0.543	0.924	0.708	0.553
PL8-1	0.6	0.891	0.371	0.268	0.834	0.459	0.365
PL8-2	0.6	0.740	0.725	0.612	0.773	0.677	0.563
PL9-2	0.6	0.904	0.997	0.773	0.920	0.909	0.719
PL10-2	0.6	0.731	0.830	0.743	0.878	0.559	0.449

近1a,这从侧面显示出本文所提的最优模型具有良好的长期预测能力。由表1还可以发现最优模型以0.5、0.6分位数居多,这表明对于升船机的大部分测点,0.5、0.6分位数回归模型更具有代表性。

4 结 论

a. 影响升船机左右岸方向绝对水平位移的主要因素是温度和上游水位前期平均值,上下游水位次方项对其影响很小。

b. 分位数回归比均值回归更适合挖掘监测资料中所蕴含的信息,最优模型与逐步回归模型相比,在具有同样简单的数学模型的同时,对测试集的预测效果要更好,稳定性强。

c. 对于大部分测点,0.5、0.6分位数回归模型可较好地代表变形与影响因子间的关系。

参考文献:

[1] 徐卫亚,徐伟,闫龙,等.基于LMBP和SVR的倾倒变形体变形预测[J].河海大学学报(自然科学版),2021,49(1):64-69.(XU Weiya, XU Wei, YAN Long, et al. Deformation prediction of toppling deformed slope based on LMBP and SVR[J]. Journal of Hohai University (Natural Sciences),2021,49(1):64-69.(in Chinese))

[2] 钮新强.三峡升船机结构关键技术研究[D].武汉:华中科技大学,2005.

[3] 胡亚安,李中华,李云,等.中国大型升船机研究进展[J].水运工程,2016(12):10-19.(HU Yaan, LI Zhonghua, LI Yun, et al. Research developments in the field of major ship lift in China[J]. Port & Waterway Engineering,2016(12):10-19.(in Chinese))

[4] 卢骏,戴吾蛟,章浙涛.大坝变形变系数回归建模[J].武汉大学学报(信息科学版),2015,40(1):139-142.(LU Jun, DAI Wujiao, ZHANG Zhetao. Modeling dam deformation using varing coefficient regressin [J]. Geomatics and Information Science of Wuhan University,2015,40(1):139-142.(in Chinese))

[5] 侯回位,郑东健,刘永涛,等.基于EEMD-SE-LSTM的混凝土坝变形监测模型[J].水利水电科技进展,2022,42(1):61-66.(HOU Huiwei, ZHENG Dongjian, LIU Yongtao, el al. Deformation monitoring model of concrete dams based on EEMD-SE-LSTM[J]. Advances in Science and Technology of Water Resources,2022,42(1):61-66.(in Chinese))

[6] 裴耀.分位数回归及其应用[D].武汉:华中师范大学,2014.

[7] 黄耀英,何一洋,沈振中,等.大坝监测量最佳统计模型优选方法[J].水利学报,2022,53(2):154-164.(HUANG Yaoying, HE Yiyang, SHEN Zhenzhong, et al. Optimization method of optimal statistical model of dam monitoring data [J]. Journal of Hydraulic Engineering,2022,53(2):154-164.(in Chinese))

[8] MELIGKOTSIDOU L, VRONTOS I D, VRONTOS S D. Quantile regression analysis of hedge fund strategies[J]. Journal of Empirical Finance,2009,16(2):264-279.

[9] 李婧,闫磊,屈春来,等.分位数回归和GAMLSS模型在非一致性洪水频率分析中的比较[J].水利水电科技进展,2020,40(5):48-54.(LI Jing, YAN Lei, QU Chunlai, et al. Comparison of quantile regression and GAMLSS model in non-stationary flood frequency analysis[J]. Advances in Science and Technology of Water Resources,2020,40(5):48-54.(in Chinese))

[10] KOENKER R, BASSETT G. Regression quantiles[J]. Econometrica,1978,46(1):33-50.

[11] ZOU H, ZHANG H H. On the adaptive elastic-net with a diverging number of parameters[J]. Annals of Statistics,2009,37(4):1733-1751.

[12] GU Y, FAN J, KONG L, et al. ADMM for high-dimensional sparse penalized quantile regression[J]. Technometrics,2018,60(3):319-331.

[13] 朱艳华.水口水电站升船机塔楼变形分析[J].水利与建筑工程学报,2010,8(3):124-126.(ZHU Yanhua. Analysis of deformation behavior of ship lift tower in Shuikou Hydropower Station [J]. Journal of Water Resources and Architectural Engineering,2010,8(3):124-126.(in Chinese))

[14] 邱章云,罗琛.三峡升船机塔柱施工期变形监测[J].中国工程科学,2013,15(9):102-107.(QIU Zhangyun, LUO Chen. Deformation monitoring on Three Gorges Project ship lift tower during construction period[J]. Strategic Study of CAE,2013,15(9):102-107.(in Chinese))

[15] 王建,郑东健,姜海霞.大坝观测资料分析中谐波温度因子的改进及应用[J].河海大学学报(自然科学版),2002,30(1):68-72.(WANG Jian, ZHENG Dongjian, JIANG Haixia. Improvement of harmonic wave temperature factor and its application to analysis of dam monitoring data[J]. Journal of Hohai University(Natural Sciences),2002,30(1):68-72.(in Chinese))

[16] 王旭虹,朱虹,汪基伟.三峡升船机塔柱结构温度变形计算[J].工程力学,2001(增刊1):931-935.

[17] KOENKER R, MACHADO J. Goodness of fit and related inference processes for quantile regression[J]. Journal of the American Statal Association,1999,94(448):1296-1310.

[18] 徐若南.基于自适应弹性网对强相关数据的群组变量选择的研究[D].合肥:合肥工业大学,2017.

(收稿日期:2022-04-13 编辑:刘晓艳)