

DOI: 10. 20040/j. cnki. 1000-7709. 2023. 20221541

基于管网分区流量数据的漏损检测方法研究

王 彤^{a,b}, 李钟毓^{a,b}, 康炳卿^{a,b}, 朱多林^{a,b}, 王晴怡^{a,b},
赵红斌^{a,b}, 许德伦^{a,b}, 洪 磊^{a,b}

(长安大学 a. 建筑工程学院; b. 住房与城乡建设部给水排水重点实验室, 陕西 西安 710061)

摘要: 为减轻城镇供水管网漏损造成资源浪费,以预测分类为基本思想,将小波降噪处理后的管网流量监测数据用于 PSO 寻优最小二乘支持向量机(LSSVM)算法的预测模型,并对预测模型进行训练和评估,再结合模型预测误差分布规律,推出阈值和漏损量的估算方法用于漏损检测。结果表明,模型预测值与实际水量的平均误差较低,稳定性高,结合预测—阈值分类法能够及时检测出突发性漏损并相对准确地估算出漏损量。

关键词: 供水管网; 流量数据; 漏损检测; 预测分类; 误差分析

中图分类号: TU991; TV672⁺.2

文献标志码: A

文章编号: 1000-7709(2023)07-0127-05

1 引言

目前城市供水管网多采用 DMA 分区管理,在分区入口处安装传感器,对分区用水量进行监测。但由于供水管网系统庞杂,而监测装置有限,难以把握供水管网的真实情况。随着计算机技术的发展,支持向量机(SVM)算法凭借优越的性能在管网异常检测中获得广泛应用。宋杰等^[1]利用 PSO-SVM 模型对给水管网的试验数据进行训练和预测,能够为管网漏损诊断提供帮助;吴以文等^[2]发现 LSSVM 交互预测模型有效地降低了数据丢失、数据错误对预测结果的影响,且对较小的爆管响应更加灵敏;陈磊等^[3]利用 AGA 优化 LSSVM 参数,建立了基于 AGA-LSSVM 的管道初次漏损时间预测模型,能够对相同管材管道的初次漏损时间进行高精度预测。这类方法虽能用于部分管网漏失检测,但缺乏对监测数据的有效利用,且管网漏损的影响因素复杂,现有模型的预测结果与管网实际情况仍存在较大差别。为此,本文以城镇供水管网分区流量监测数据为基础,探究了 PSO 优化 LSSVM 时序预测模型的预测水量与实际监测数据的误差分布规律,并提出阈值和漏损量估计方法,以期快速检测漏损突发及漏损点的精确定位提供参考。

2 管网分区流量预测

根据数据分析方法的不同,基于数据驱动的漏损检测方法可分为基于分类、预测分类和概率统计的漏损检测方法三类^[4]。其中,以预测分类为基础的漏损检测方法避免了基于分类的漏损定位方法中需要大量分类监测数据的局限,仅需少量样本对供水管网运行数据进行实时预测。而最小二乘支持向量机(LSSVM)算法在处理小样本、非线性、高维数数据方面具有独特优势,可尝试与不同的寻优算法相结合,用于管网漏损检测与定位研究。

2.1 预测模型搭建

建立时序预测模型需找到时间变化与数据的内在关系^[5]。以时间序列实测值 $\{x_n\}_{n=1}^M$ 为基础搭建时间序列训练样本集 $S = \{(x_i, y_i)\}_{i=1}^N$, 并利用该样本集训练 LSSVM 算法,找到时序预测的非线性映射函数关系 $T(x)$, 便可对 $\{x_n\}_{n=1}^M$ 之外的时间序列数据进行预测。即:

$$x_n = T(x_{n-1}, x_{n-2}, \dots, x_{n-k}) \\ n = k + 1, k + 2, \dots, M \quad (1)$$

式中, k 为嵌入维数。

为提高 LSSVM 预测回归性能,采用高斯核函数(RBF)数选择和超参数优化:

收稿日期: 2022-07-27, 修回日期: 2022-09-23

基金项目: 水资源高效开发利用重点专项(2018YFC0406200)

作者简介: 王彤(1965-), 男, 教授、硕导, 研究方向为给排水系统优化及人工湿地, E-mail: 894496318@qq.com

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (2)$$

式中, δ 为函数的宽度参数。

超参数包括 LSSVM 的正则化参数 c 和核函数中的参数 δ 。利用 PSO 寻优算法不断迭代更新, 找出 LSSVM 算法的最优超参组合 (c, δ) 作为预测模型参数, 优化目标为:

$$\min E = \min \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - \hat{y}_k)^2} \quad (3)$$

式中, E 为实际值与预测值之间的均方根误差; n 为用于优化的样本数; y_k, \hat{y}_k 分别为实际值、预测值。

预测模型具体流程见图 1。

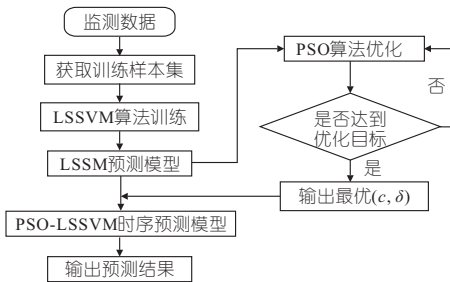


图 1 预测模型流程图

Fig. 1 Flow chart of the prediction model

2.2 预测模型评估

以 X 市 Y 区某日的分区入口实测流量数据为例。为测试模型的小样本预测能力, 采用 Y 区连续 7 d 历史数据评估模型。前 6 d 数据用来训练模型, 第 7 d 数据用来测试模型。数据样本采样频率为 5 min, 即每天有 288 个监测数据, 利用 Matlab 软件中 Wden 函数对流量数据进行小波变换处理, 采用小波阈值法降噪, 阈值设为启发式估计规则, 可通过不同的噪声实时调整阈值。考虑到用于漏损检测的流量数据的实际需求, 将每小时内的流量监测值做累计处理后作为各整点时刻的流量值, 从而放大管网漏损信号。采用 Matlab 建立预测模型, 得到第 7 d 整点流量预测结果见图 2。

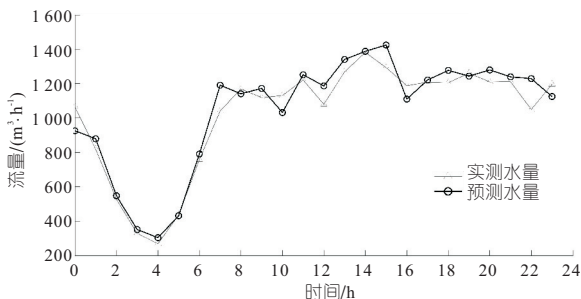


图 2 第 7 d 监测与预测流量对比

Fig. 2 Comparison of monitoring and predicted flow rates on day 7

为检验模型预测结果的准确性与稳定性, 选

用平均绝对误差 (e_{MAE})、平均相对误差百分比 (e_{MAPE})、Theil 不等系数 (U) 和纳什模型效率 (N_{NSE}) 4 种统计量^[6]作为评价指标对模型预测结果进行评估。计算结果见表 1。

表 1 模型评估结果

Tab. 1 Results of the model evaluation

评价指标	e_{MAE} /($m^3 \cdot h^{-1}$)	$e_{MAPE}/\%$	U	N_{NSE}
数值	63.2	0.07	0.001 28	0.998 7

由表 1 可看出, 在该算例中, 利用 LSSVM-PSO 时序预测模型进行小样本预测的平均预测误差百分比在 1% 以内, 且纳什模型效率很高, 满足预测模型的稳定性要求。可见基于 LSSVM-PSO 算法的时序预测模型可预测管网分区流量、检测管网漏损。

3 漏损检测和估算方法

3.1 检测方法

当利用预测模型预测管网分区流量时, 流量实际监测值和预测值之间关系为:

$$y_k = \hat{y}_k + e_k \quad (4)$$

式中, $k = 1, 2, \dots, n$ 为预测值样本时间 T 内的离散时刻, T 可取 12、24 h 等, 由预测模型输出样本的序列时长决定; e_k 为 k 时刻的模型误差预测。

当管网发生漏损时, 预测值与监测值之间的差值由误差 e_k 和漏损水量共同组成, 因此可通过设定模型预测误差在正常情况下的波动阈值来判断管网是否发生漏损。当监测值与预测值之间的残差超出预测模型误差阈值时, 可判定为管网发生漏损。

3.2 预测误差分析

利用数据分析方法, 确定管网正常运行情况下流量监测值与预测值之间的误差波动范围, 作为漏损检测的阈值。利用预测模型输出 4 d 的流量预测值, 计算模型预测值与实际监测值的误差并绘制误差直方图, 结果见图 3, 表 2 为误差值描述性统计。

利用数据分析软件 SPSS 对误差进行处理, 当偏度和峰度为 0 时, 样本分布为正态分布, 因此偏度和峰度越接近 0, 样本分布与正态分布越接近。由表 2 可知, 偏度 (-0.176)、峰度 (-0.216) 均非常接近 0, 可判断数据样本的误差频率分布近似服从正态分布。

为进一步验证误差分布是否服从正态分布, 对数据进行线性拟合分析, 拟合结果见图 4。由

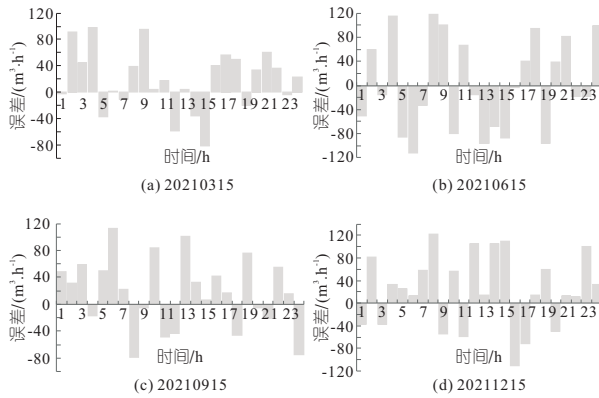


图 3 X 市 Y 区 4 d 流量预测误差直方图

Fig. 3 Histogram of 4-day flow prediction error in District Y, X City

表 2 误差值描述性统计

Tab. 2 Descriptive statistics of the error values

统计参数	统计量	标准误差
均值	18.695 86	9.47
标准差	46.395 01	
峰度	-0.215 65	0.35
偏度	-0.176 04	0.74

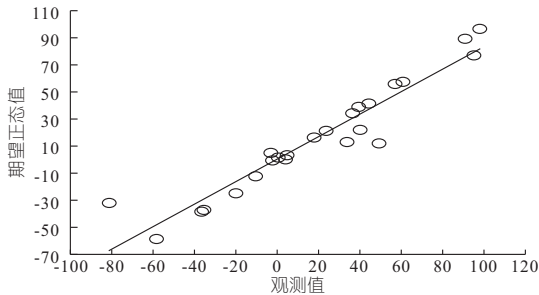


图 4 线性拟合 Q-Q 图

Fig. 4 Linear fit to the Q-Q plot

图 4 可看出,期望正态值与实际观测值大致分布在 $y=x$ 直线两侧,可将预测误差数据近似视为正态分布,进行误差阈值和漏损量的估算。

3.3 漏损量和阈值估算

当周期 T 为 24 h 时,预测误差 $e(k)$ 近似服从正态分布:

$$e(k) \approx N(0, \sigma^2(k)) \text{ 且 } \sigma^2(k) = \sigma^2(k+T) \quad (5)$$

若管网中发生漏损,则流量预测值与监测值之间的关系为:

$$y(k) = \hat{y}(k) + e(k) + f(k) \quad (6)$$

式中, $y(k)$ 为 k 时刻实际流量监测值; $e(k)$ 为 k 时刻漏损量的估计误差; $\hat{y}(k)$ 为 k 时刻流量预测值; $f(k)$ 为 k 时刻管网的漏损水量, $f(k) > 0$ 。

可推出漏损量 $f(k)$ 的估算公式为:

$$\hat{f}(k) = y(k) - \hat{y}(k) = f(k) + e(k) \quad (7)$$

当给定一定时间段内的流量数据样本值后,管网在该时段内的漏损量可用各时刻的漏损估计值进行近似:

$$f(k) \approx \bar{f}(k) = \sum_{i=0}^{N_w-1} \frac{f(k-i)}{N_w} \quad (8)$$

式中, N_w 为数据样本的数量。

由于预测误差近似服从正态分布,根据式(7),若管网中 k 时刻发生漏损,其平均漏损量可采用最大似然估计法求解一定时间段内 N_w 个估计值的联合概率分布来近似估计:

$$\bar{f}(k) = \frac{\sum_{i=0}^{N_w-1} \hat{f}(k-i)}{\sum_{i=0}^{N_w-1} \sigma^2(k-i)} \quad (9)$$

在未发生漏损时, $\bar{f}(k)$ 即为 k 时刻流量预测误差的估计值,因此利用无漏损状态下的流量样本数据,根据式(9)将一定时间段内误差估计值的最大值作为阈值:

$$\lambda = \max \bar{f}(k) \quad (10)$$

式中, λ 为误差阈值,用于判断管网是否发生漏损。

当利用式(10)确定阈值时,阈值会随样本数据的更新而变化,提高了管网漏损检测的灵敏度和准确度。

最终,管网漏损检测流程见图 5。

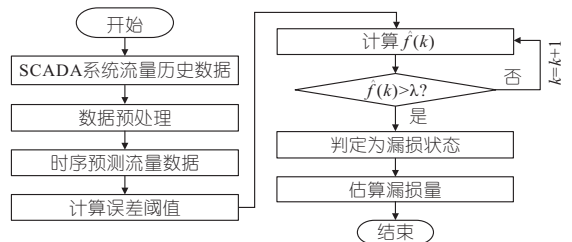


图 5 漏损检测流程图

Fig. 5 Flow chart of leakage detection

4 实际案例

4.1 案例管网概况

Z 镇人口约 10 000 人,其配水管网由 782 个节点、两个水库、一个水箱和 909 条管道组成,管网总长度约 42.6 km,最小服务水头为 20 m。整个管网由三个供水分区构成:A 区面积较大,共有 655 个用水节点,高程在 16~48 m 之间。B 区面积较小,仅有 31 个用水节点,高程均小于 16 m,与 A 区并网运行。C 区(图 6 星标左侧区域)由 92 个用水节点组成,高程均在 48 m 以上,由水箱和水泵联合供水,该区入口处装有流量计,并在 82 个用水节点上装有 AMRs 水量传感器。

采用 Epanet 建立 Z 镇管网水力模型,其供水管网的拓扑结构、监测点位置及 2018 年全年发生漏损的位置见图 7。Z 镇供水管网在 2018 年共检

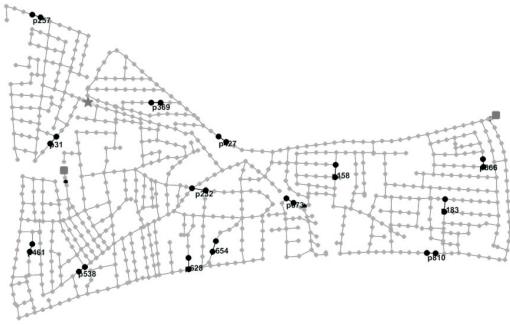


图 6 Z 镇供水管概况示意图

Fig. 6 Schematic diagram of the water supply pipe overview in Z Town

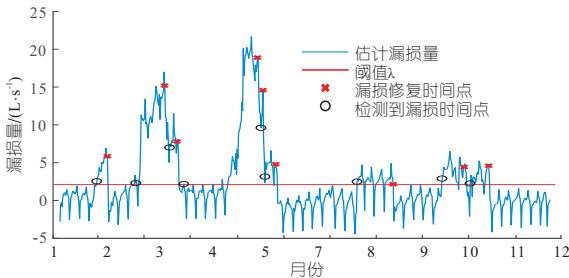


图 7 2018 年 Z 镇管网 A、B 区漏损检测结果
Fig. 7 Leakage detection results in Area A and B of Z Town in 2018

测到 14 处漏损点,按照漏失水量大小和持续时间分为爆管、管网较大漏失及背景漏失三类,其中漏失量较小且持续时间较长的背景漏失发生在管道 p257、p427、p810、p654 上,且一直未修复。

4.2 漏损检测方法验证

由于管网 A 区和 B 区并网运行,在检测漏损时将管网分为 A、B 区和 C 区两部分。A、B 区仅有管网分区入口的流量监测数据,故利用 LSSVM-PSO 算法来预测该分区流量。而 C 区由于 89% 的用户均装有 AMRs 水量传感器,因此无需复杂的算法,仅根据 AMRs 水量传感器数据和分区入口流量监测数据即可预测分区流量。

4.2.1 基于分区流量监测数据检测漏损

将 LSSVM-PSO 模型的训练数据集大小设为 7 d,然后计算各时刻预测值和实际监测值的残差,通过对比残差值和阈值大小判断管网是否发生漏损。若流量残差超出阈值,则判定为管网发生漏损,并计算漏损量;若残差未超出阈值,则作为历史数据,用来预测第 2 天各时刻的分区流量。

在初次模型训练时采用 2018 年 1 月 1~6 日各时刻的历史数据,得到阈值计算结果为 2.1 L/s。全年的漏损检测和估算结果见图 7。

由图 7 可看出,2018 年在 A、B 区发生的 9 次漏损均被检测到,且各漏点的检测时间及时,对漏损量大小的估计结果相对准确。图 7 中所示的阈值为全年漏损检测过程中各个时段所用阈值的最大

大值,若全年选用该阈值也可检测出全部漏损。

4.2.2 基于 AMRs 和分区流量监测数据检测漏损

由于 C 区 89% 的用户装有 AMR 传感器,可得到大量的用户用水量数据,利用该年 1 月份第 1 周的历史数据计算出 AMRs 测得的总用水量和分区入口流量计监测值之间的比值 K ,那么 C 区的流量预测值计算式为:

$$\hat{y}(k) = K \sum_{i=1}^{n_m} q_{AMR_i}(k) \quad (11)$$

式中, $\hat{y}(k)$ 为 k 时刻 C 区的流量预测值; n_m 为 C 区内 AMRs 的数量; $q_{AMR_i}(k)$ 为 k 时刻 C 区内各 AMRs 的监测值。

利用式(11)得到漏损检测的阈值仅为 0.3 L/s。C 区全年的漏损检测及漏损量估算结果见图 8。

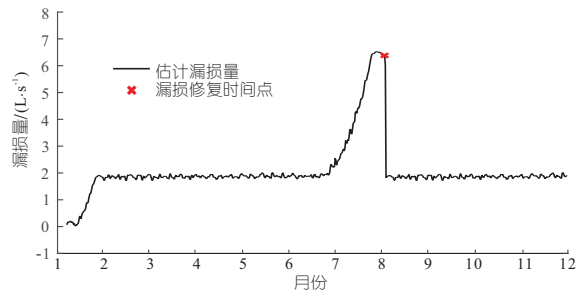


图 8 2018 年 Z 镇管网 C 区漏损检测结果
Fig. 8 Leakage detection results in Area C of Z Town pipe network in 2018

由图 8 可看出,由于 C 区监测数据较多,水量预测精度较高,2018 年 C 区发生的两次漏损均被检测到,漏损检测的灵敏度非常高,在漏点 p257 发生的漏损量较小的背景漏失也能被识别。

4.3 验证结果分析

由于 Z 镇不同分区流量传感器布置差异较大,将漏损检测过程分为两部分进行。管网的 A、B 区用基于 LSSVM-PSO 的预测模型进行流量预测,各漏点的检测时间及时且漏损量估计结果相对准确。C 区由于在 89% 的用水节点上都装有用水量传感器,利用历史数据求出用水量总和与分区流量监测值之间的比例关系即可进行流量预测,漏损检测的灵敏度非常高。

5 结论

a. 在样本数据较少时,利用 LSSVM-PSO 时序预测模型预测流量与实际监测流量的平均相对误差低,预测数据准确度较高,且纳什模型效率高,满足模型稳定性要求。根据预测结果计算阈值和流量预测误差可判断管网是否正常运行,为突发漏损检测和漏损点定位提供参考。

b. 流量监测数据的质量对基于数据驱动的管网漏损检测与定位方法有重要影响。在实际工程中,改善传感器测量效果和布局的同时,需有效利用监测数据并不断优化预测模型,提高基于数据驱动检测管网漏损等突发状况的能力。

参考文献:

- [1] 宋杰,吕谋,郝晨西,等. 基于 PSO-SVM 模型的供水管网漏损诊断[J]. 水电能源科学, 2020, 38(7): 122-125.
- [2] 吴以文,杜坤,吴汉清,等. 基于 LSSVM 交互预测的供水管网爆管检测[J]. 中国给水排水, 2022, 38(9):

58-63.

- [3] 陈磊,王培永. 基于遗传最小二乘支持向量机的初次给水管漏损时间预测模型[J]. 浙江工业大学学报, 2021, 49(5): 546-549.
- [4] 吴杰. 基于流量和压力数据分析的城市供水管网漏损定位研究[D]. 太原:太原理工大学, 2020.
- [5] 王彤,金赵归,杨瑞虎,等. 基于 PSO-LSSVM 时序预测模型的管网漏失信号识别[J]. 水电能源科学, 2022, 40(2): 132-135, 181.
- [6] 郭呈冠,刘书明,李俊禹. 基于双向长短时记忆神经网络的水量预测方法研究[J]. 给水排水, 2018, 44(3): 123-126.

Research on Leakage Detection Based on Flow Data of Pipeline Network

WANG Tong^{a,b}, LI Zhong-yu^{a,b}, KANG Bing-qing^{a,b}, ZHU Duo-lin^{a,b},

WANG Qing-yi^{a,b}, ZHAO Hong-bin^{a,b}, XU De-lun^{a,b}, HONG Lei^{a,b}

(a. School of Civil Engineering; b. Key Laboratory of Water Supply and Sewerage, Ministry of Housing and Urban-Rural Development, Chang'an University, Xi'an 710061, China)

Abstract: In order to alleviate the waste of resources caused by the leakage of urban water supply networks, taking predictive classification as the basic idea, the monitoring data after wavelet noise reduction processing was used in the prediction model of PSO seeking least squares support vector machine algorithm, and the prediction model was trained and evaluated, then combined with the model prediction error distribution law, the estimation method of threshold value and leakage volume was introduced for leakage detection. The results show that the average error between the model prediction and the actual water volume is low, the stability is high, and the combination of the prediction-threshold classification method can detect sudden leakage in a timely manner and estimate the leakage volume relatively accurately.

Key words: water supply network; flow data; leakage detection; prediction classification; error analysis

(上接第 149 页)

- [2] 赵彬如,戴会超,戎贵文,等. 竖缝位置对竖缝式鱼道水力特性的影响[J]. 水利水电科技进展, 2017, 37(5): 69-73, 83.
- [3] 张超,孙双科,李广宁. 竖缝式鱼道细部结构改进研究[J]. 中国水利水电科学研究院学报, 2017, 15(5): 389-396.
- [4] 徐体兵,孙双科. 竖缝式鱼道水流结构的数值模拟

[J]. 水利学报, 2009, 40(11): 1386-1391.

- [5] 郭维东,孙磊,高宇,等. 同侧竖缝式鱼道水力特性研究[J]. 水电能源科学, 2012, 30(3): 81-83.
- [6] 刘本芹,黄岳,宣国祥. 小长宽比鱼道池室水力学试验研究[J]. 水利水运工程学报, 2015(6): 101-106.
- [7] 国家能源局. 水电工程过鱼设施设计规范(NB/T35054-2015)[S]. 北京:中国电力出版社, 2015.

Design and Experimental Study of Vertical Slit Fishway in Actual Project

HAN Lei¹, LIU Fa-zhi², ZHAO Xing-long³, YE Kun-he³, LI Shu-hang¹, LI Yang¹

(1. Heilongjiang Province Hydraulic Research Institute, Harbin 150080, China;

2. Shuifa Planning and Design Co., Ltd., Jinan 250000, China;

3. School of Hydraulic and Electric-power, Heilongjiang University, Harbin 150080, China)

Abstract: The vertical slit fishway has the capacity to better cope with the fluctuation of water level in both upper and lower sections and favorable water flow conditions. Three schemes were designed for the shape of the fishway tank chamber of a certain project fishway. The physical models verified with a 1 : 5 scales were used to investigate the impact of the tank chamber's structure on the flow pattern, maximum flow velocity, and hydraulic properties of mainstream region under different size and type of baffle and guide plates. This was done through numerical simulation calculation. They performed a test of the physical model on the optimal shape with scale of 1 : 5. The physical model of the overall with scale 1 : 20 was used to analyze the maximum velocity along the water depth and the vertical slit, and the diversion flow of the fishway under different water depth conditions was proposed, providing reference and reference for similar projects.

Key words: hydraulics; fishway design; numerical simulation; physical model; test analysis