

# 基于数据驱动的金属供水管网铁腐蚀预测模型研究

李瑞莲<sup>1</sup>, 朱 斌<sup>2</sup>, 周秀美<sup>2</sup>, 白晓慧<sup>3</sup>, 舒诗湖<sup>4</sup>

(1. 太原理工大学 信息与计算机学院, 山西 太原 030024; 2. 上海城投水务集团有限公司, 上海 200002; 3. 上海交通大学 生命科学技术学院, 上海 200240; 4. 上海城市水资源开发利用国家工程中心有限公司, 上海 200082)

**摘 要:** 供水管网铁腐蚀是影响末端水质可接受性的重要因素。根据管网水质构建的管网铁腐蚀预测模型对有效发挥水厂水质调控功能, 保持管网水质化学稳定性及管网维护具有重要应用价值。根据上海市某供水系统管网水质和管壁铁腐蚀的大量监测数据, 分别采用多元回归法、逐步回归法和线性回归法构建了供水管网铁腐蚀预测模型, 并比较了模型误差。结果表明, 水体 pH 值、 $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  含量为管壁铁腐蚀的主要影响因素; 逐步回归模型的模拟效果最好, 该模型具有公式简单和预测误差小的优点, 在 95% 置信度下, 预测值和实测值之间不存在显著性差异, 可用于该供水系统铁腐蚀预测和出厂水质目标调控。

**关键词:** 供水管网; 铁腐蚀; 预测模型; 数据驱动; pH 值;  $\text{Cl}^-$ ;  $\text{SO}_4^{2-}$

**中图分类号:** TU991 **文献标识码:** A **文章编号:** 1000-4602(2019)11-0054-05

## Iron Corrosion Prediction Model Based on Data-driven in Water Distribution System

LI Rui-lian<sup>1</sup>, ZHU Bin<sup>2</sup>, ZHOU Xiu-mei<sup>2</sup>, BAI Xiao-hui<sup>3</sup>, SHU Shi-hu<sup>4</sup>

(1. College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China; 2. Shanghai Municipal Investment <Group> Corporation, Shanghai 200002, China; 3. School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China; 4. Shanghai National Engineering Research Center of Urban Water Resources Co. Ltd., Shanghai 200082, China)

**Abstract:** The phenomenon of iron corrosion in pipeline is an important factor affecting water quality acceptance at the tap. An iron corrosion prediction model based on pipe water quality can provide guidance to the control of pipe water chemical stability and pipeline maintenance, through controlling the effluent quality in waterworks. Based on the water quality data and iron corrosion detection data collected in a water distribution system in Shanghai, three prediction models, including the multivariate regression method, the stepwise regression method and the linear regression method, were evaluated for predicting iron corrosion in water distribution system. The results showed that pH,  $\text{Cl}^-$  and  $\text{SO}_4^{2-}$  were the major influencing factors to iron corrosion. The stepwise regression model demonstrated the best predictions in iron corrosion. This model had a simple formula and low residuals. Under 95% confidence level, the difference between predicted and measured values was insignificant, which could be utilized to predict

基金项目: 中国工程院重大咨询项目(2016-06-ZD-01); 国家水体污染控制与治理科技重大专项(2017ZX07207-004、2017ZX07207-005)

通信作者: 白晓慧 E-mail: xhbai@sjtu.edu.cn

iron corrosion and set control targets for effluent quality.

**Key words:** water distribution system; iron corrosion; prediction model; data-driven; pH;  $\text{Cl}^-$ ;  $\text{SO}_4^{2-}$

城市供水管网铁腐蚀是导致供水管网漏损、影响供水安全和供水企业健康发展的重要因素,也是导致供水末端水体浊度、色度升高的重要原因。根据国内外对铁腐蚀的相关研究,饮用水的 pH 值、碱度、氯化物、硫酸盐和有机物等都会影响铁质管材的氧化腐蚀速率<sup>[1-2]</sup>。目前关于管壁铁腐蚀的研究大多集中在管垢腐蚀机理<sup>[3-4]</sup>或管外及工业用水系统管道腐蚀预测与评价方面<sup>[5-9]</sup>,而利用实际管网水质针对城市供水管网管壁铁腐蚀进行模型预测的研究则极少,该模型的建立对通过调控水厂出水水质化学稳定性来调控管网铁腐蚀、改善末端饮用水水质可接受性和科学制定供水管道更新改造计划或运行维护具有重要意义。

笔者根据对实际供水管网水质和铁腐蚀检测的大量数据,采用基于数据驱动的多指标决策模型<sup>[10]</sup>,以不同建模方法构建并比较了供水管壁铁腐蚀预测模型的适用性,以为供水系统铁腐蚀控制提供技术支持。

## 1 研究对象与方法

### 1.1 研究对象

以上海黄浦江上游水源净水工艺及其供水主干系统为研究对象。该水厂采用常规工艺加臭氧—生物活性炭深度处理工艺,并以次氯酸钠作为消毒剂。其供水主干管材均为内含水泥涂衬的球墨铸铁管。采样点包括水厂出厂水及至供水末梢且水量覆盖的6个管网采样点(管径>800 mm),全长近8 km。采样时间从当年7月10日至次年5月22日,每月采样两次。

采样方法:准备灭菌并干燥的2.5 L塑料桶,在供水企业的协助下,通过设置为采样点的街边消防栓采集水样,采样时先放掉消防栓及其支管中存水,再关小水量待水流稳定后采样。所有样品应及时分析,或置于4℃冰箱冷藏保存不超过48 h。

### 1.2 研究方法

管网水质数据通过对上海市黄浦江上游水源供水系统实际管网采样点的水质检测获得。检测指标包括 pH 值、碱度、总硬度、氯化物、硫酸盐、总铁、溶解性有机碳(DOC),其中 pH 值、总铁采用美国哈希

便携式检测仪现场测定,其他水质指标测定方法主要依据《生活饮用水标准检验方法》。

出厂水的 pH 值为 6.91~7.16,均值为 7.03;碱度为 66.01~106.83 mg/L,均值为 78.42 mg/L;总硬度为 160.54~243.82 mg/L,均值为 199.98 mg/L;氯化物为 52.93~88.41 mg/L,均值为 73.73 mg/L;硫酸盐为 82.88~124.02 mg/L,均值为 100.32 mg/L;总铁为 0.03~0.32 mg/L,均值为 0.09 mg/L;DOC 为 4.62~5.38 mg/L,均值为 4.97 mg/L。

管网水的 pH 值为 6.68~7.20,均值为 6.96;碱度为 61.18~106.83 mg/L,均值为 76.94 mg/L;总硬度为 153.88~246.72 mg/L,均值为 195.32 mg/L;氯化物为 55.76~93.61 mg/L,均值为 73.71 mg/L;硫酸盐为 85.45~125.30 mg/L,均值为 100.74 mg/L;总铁为 0.04~0.65 mg/L,均值为 0.20 mg/L;DOC 为 4.45~15.85 mg/L,均值为 5.83 mg/L。

相关性分析采用双侧检验的 Pearson 相关分析。采用幂函数型和线性函数型预测模型,幂函数型预测模型包括多元回归模型和逐步回归模型,广泛应用于金属管道随时间变化的腐蚀预测和消毒副产物生成预测<sup>[5-9,11]</sup>;线性函数型预测模型多应用于金属管道腐蚀率预测。相关性分析和回归分析皆采用 SPSS11.5 软件。

## 2 结果与讨论

### 2.1 管网水质与管壁铁腐蚀相关性分析

管网水的 pH 值、碱度与铁含量呈显著负相关,相关系数分别为 -0.823、-0.744,说明 pH 值和碱度的下降会加重管壁铁腐蚀。国内外大量研究通过静态管材浸泡和动态模拟管网证实, pH 值降低可以增大铁腐蚀速率,碱度的增加能强化水质缓冲能力,并在管壁形成诸如  $\text{CaCO}_3$  和  $\text{FeCO}_3$  的保护层,减缓金属管壁的腐蚀<sup>[12-13]</sup>。

$\text{Cl}^-$  和  $\text{SO}_4^{2-}$  含量与铁含量呈显著正相关,相关系数分别为 0.884、0.702,表明水中  $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  含量的上升能加速管壁铁腐蚀。这是由于水体中  $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  的增加会提高水的电导率,加快腐蚀进程;

同时,  $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  离子半径小, 容易穿透、破坏金属表面的钝化膜, 促进腐蚀。Strehblow 对  $\text{Cl}^-$  引起铁钝化膜破坏机理的研究中, 提出  $\text{Cl}^-$  能促进铁钝化膜的溶解破坏; 利用 AES 和 XPS 等方法也能检测到在点蚀过程中  $\text{Cl}^-$  渗入钝化膜的现象<sup>[3]</sup>。

DOC 与铁含量呈正相关, 相关系数为 0.719。研究表明, DOC 参与铁管壁的电化学腐蚀, 主要是酸性基团与铁氧化表面反应。此外, DOC 中所包含的小分子羧酸, 诸如甲酸、乙酸等, 给微生物生长提供了营养环境, 利于微生物诱导的铁腐蚀<sup>[4]</sup>。

## 2.2 多元回归模型

多元回归模型包括的水质参数有 pH 值、碱度、 $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  含量以及 DOC, 其中  $\text{Cl}^-$ 、 $\text{SO}_4^{2-}$  和 DOC 含量都以  $\text{mg/L}$  为计量单位。水体硬度与铁腐蚀相关性较低, 因此模型不包括硬度。模型的拟合方程 ( $n=5$ ) 如下:

$$\text{Fe} = 0.025 \text{ pH}^{-13.393} [\text{Alk}]^{-0.194} [\text{Cl}^-]^{4.041} [\text{SO}_4^{2-}]^{2.121} [\text{DOC}]^{0.140} \quad (1)$$

模型的复相关系数  $R^2$  为 0.851, 拟合程度较高。各水质参数的指数系数表征参数对铁释放的影响程度。pH 值指数系数为负, 其绝对值在各项中最大, 表明 pH 值的降低是促进管壁铁腐蚀的重要因素。 $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  含量升高则是导致铁腐蚀的次要因素, 指数系数与相关性分析结果相符。

残差是实测值与回归估计值的差, 可通过残差值分析数据的可靠性和异常数据出现频率, 通常采用  $3\sigma$  准则<sup>[6]</sup>。

模型预测值与实测值的比较如图 1 所示, 模型残差值范围为  $(-0.078 \text{ } 0, 0.091 \text{ } 8)$ , 在 95% 置信度下, 即置信区间为  $(-0.090 \text{ } 1, 0.090 \text{ } 1)$ , 存在 1 个差异数值, 出现频率为 2.78%。

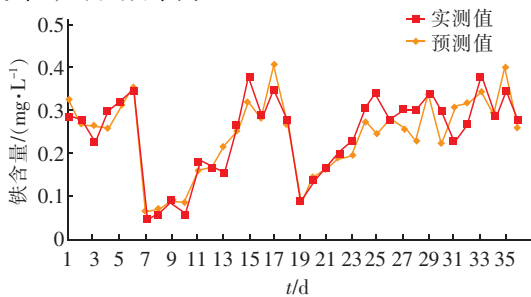


图1 多元回归模型预测值与实测值对比

Fig. 1 Comparison between predicted values and measured values by multivariate regression model

## 2.3 逐步回归模型

逐步回归模型通过对自变量的偏回归平方和 (即贡献) 在预先给定的  $F$  水平下进行显著性检验, 其优点在于能从大量自变量中选择出最重要的自变量。本研究采用  $\alpha=0.10$  水平的  $F$  检验, 得到主要参数为 pH 值、 $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  含量。模型的拟合方程 ( $n=3$ ) 如下:

$$\text{Fe} = 0.077 \text{ pH}^{-13.364} [\text{Cl}^-]^{4.062} [\text{SO}_4^{2-}]^{1.907} \quad (2)$$

模型的复相关系数  $R^2$  为 0.848, 拟合程度较高。经  $T$  检验, pH 值、 $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  的  $P$  值均为零, 按  $\alpha=0.10$  水平, 均有显著性意义。

模型预测值与实测值的对比如图 2 所示。模型残差值范围为  $(-0.072 \text{ } 2, 0.080 \text{ } 0)$ , 在 95% 置信度下, 即置信区间为  $(-0.087 \text{ } 0, 0.087 \text{ } 0)$ , 不存在差异数值, 出现频率为零。模型预测值接近实测值, 预测能力可信度高。

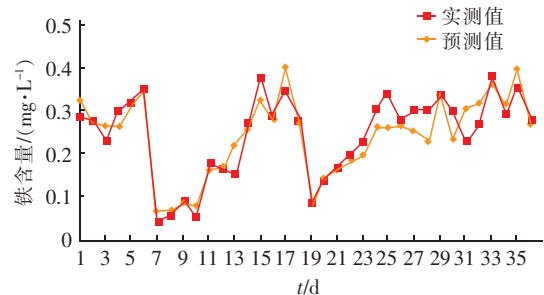


图2 逐步回归模型预测值与实测值对比

Fig. 2 Comparison between predicted values and measured values by stepwise regression model

## 2.4 线性回归模型

线性回归模型包含的水质参数有 pH 值、碱度、 $\text{Cl}^-$ 、 $\text{SO}_4^{2-}$ 、DOC。模型的拟合方程 ( $n=5$ ) 如下:

$$\text{Fe} = -0.390 \text{ pH} - 0.013 \text{ } 2 \text{ Alk} + 0.322 \text{ Cl}^- + 0.359 \text{ SO}_4^{2-} - 0.001 \text{ } 02 \text{ DOC} + 1.770 \quad (3)$$

模型的复相关系数  $R^2$  为 0.816, 拟合程度较高, 但较上述两个模型低。经  $T$  检验, pH 值、 $\text{Cl}^-$  和  $\text{SO}_4^{2-}$  的  $P$  值分别为 0、0 和 0.001, 按  $\alpha=0.10$  水平, 均有显著性意义; 而碱度和 DOC 参数的  $P$  值偏高, 不具有显著性。

模型预测值与实测值的对比如图 3 所示, 模型残差值范围为  $(-0.078 \text{ } 5, 0.088 \text{ } 3)$ , 在 95% 置信度下, 即置信区间为  $(-0.090 \text{ } 6, 0.090 \text{ } 6)$ , 不存在差异数值, 出现频率为零, 模型预测值接近实测值。

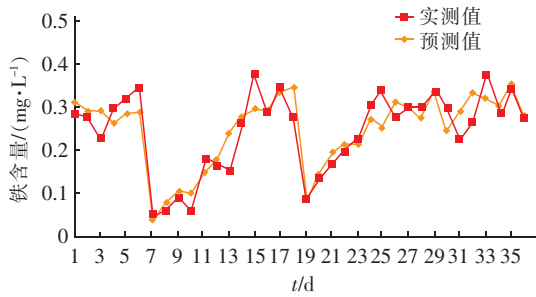


图3 线性回归模型预测值与实测值对比

Fig. 3 Comparison between predicted values and measured values by linear regression model

## 2.5 3种预测模型的比较

3种预测模型的比较见图4。

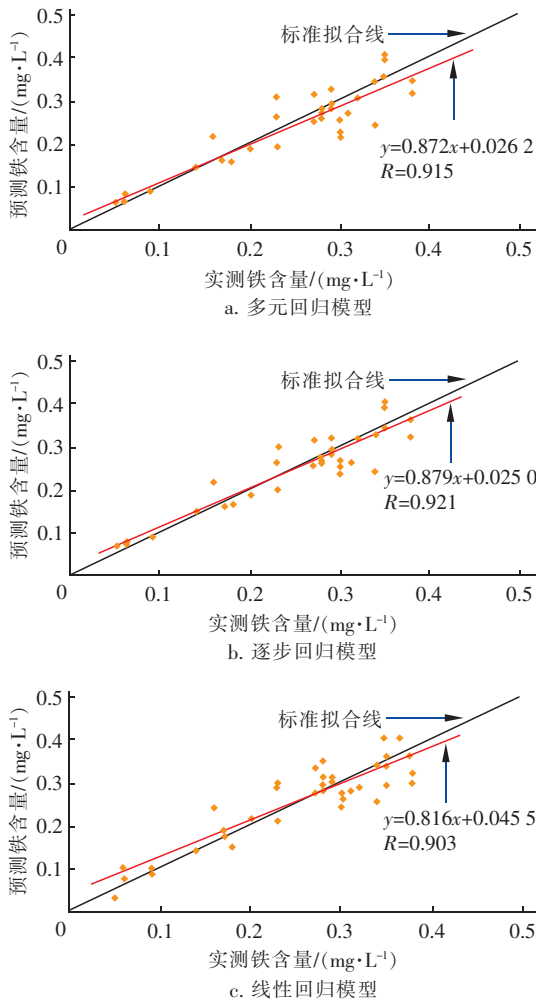


图4 3种预测模型的比较

Fig. 4 Comparison of three prediction models

通过对3种模型的实测值和预测值进行线性拟合,发现逐步回归模型的预测值向标准拟合线的集聚程度最高,拟合方程的 $R$ 值最高(0.921);多元回

归模型的预测值次之, $R=0.915$ ;线性回归模型的预测值最差, $R=0.903$ 。

本研究所建模型的预测误差都明显低于丁宏达所述3种常用球墨铸铁管道外腐蚀率预测模型的预测误差<sup>[5]</sup>,相关性也高于高强等采用NARX神经网络仿真建立的工业循环水腐蚀速率模型数据的相关性(0.879)<sup>[8]</sup>。

在实测铁含量 $<0.2$  mg/L的条件下,逐步回归模型和多元回归模型预测值与实测值的偏差明显较小,其中逐步回归模型的预测值偏差不超过0.05 mg/L,预测效果最好。在实测铁含量 $>0.2$  mg/L的条件下残差较大,预测值和实测值比值 $<1.0$ ,说明预测模型普遍低估。因此在铁含量较高( $>0.2$  mg/L)的条件下,可对预测值进行修正,以逐步回归模型为例, $Fe_{\text{实测}} = 1.138(Fe_{\text{预测}} - 0.025)$ 。

通过比较3种预测模型的优缺点(如表1所示),逐步回归模型公式简单、需要参数少、数据易获得且预测误差小,可作为最佳的供水管壁铁腐蚀预测模型。

表1 3种预测模型的特点

Tab. 1 Characteristics of three prediction models

项 目	优点	缺点
多元回归模型	拟合程度高,预测误差小	公式复杂,参数数据不易获得(如DOC)
逐步回归模型	公式简单,参数少;拟合程度高,预测误差小	
线性回归模型	公式简单	参数数据不易获得;拟合程度偏低,预测误差较大

## 3 结论

① pH值和碱度与供水管网铁腐蚀呈现显著负相关, $Cl^-$ 、 $SO_4^{2-}$ 和DOC含量与铁腐蚀呈现显著正相关。

② 通过逐步回归法拟合的预测模型公式简单、需要参数量少、预测误差小,优于多元回归模型和线性回归模型,能较准确地预测因管网水质变化导致的供水管壁铁腐蚀程度。

③ 由于管网水质化学稳定性主要由出厂水质决定,研究建立的模型可用于城市供水管网铁腐蚀预测与调控、水厂净水水质化学稳定性控制指标设



定及供水企业管网更新改造计划制定等。

### 参考文献:

- [1] Imran S, Dietz J D, Asce M, *et al.* Modified Larsons ratio incorporating temperature, water age, and electroneutrality effects on redwater release[J]. *J Environ Eng*, 2005, 131(11):1514-1520.
- [2] Lasheen M R, Sharaby C M, El-Kholy N G, *et al.* Factors influencing lead and iron release from some Egyptian drinking water pipes[J]. *J Hazard Mater*, 2008, 160(2/3):675-680.
- [3] Strehblow H H S. Breakdown of passivity and localized corrosion: Theoretical concepts and fundamental experimental results[J]. *Materials and Corrosion*, 1984, 35(10):437-448.
- [4] Ona-Nguema G, Abdelmoula M, Jorand F, *et al.* Iron (II, III) hydroxycarbonate green rust formation and stabilization from lepidocrocite bioreduction[J]. *Environ Sci Technol*, 2002, 36(1):16-20.
- [5] 丁宏达. 金属管道外腐蚀率预测模型的分析评价[J]. *水力采煤与管道运输*, 2008(1):1-4.  
Ding Hongda. Analysis and evaluation of the external corrosion rate of metal pipes using prediction model[J]. *Hydraulic Coal Mining & Pipeline Transportation*, 2008(1):1-4(in Chinese).
- [6] 张新生, 赵梦旭, 王小完. 尾段残差修正 GM(1,1) 模型在管道腐蚀预测中的应用[J]. *中国安全科学学报*, 2017, 27(1):65-70.  
Zhang Xinsheng, Zhao Mengxu, Wang Xiaowan. Application of modified empenage residual error GM(1,1) model in prediction of pipeline corrosion[J]. *China Safety Science Journal*, 2017, 27(1):65-70(in Chinese).
- [7] 彭世金, 汤海平, 丁雅萍, 等. 管道腐蚀安全评价模型特性研究[J]. *中国安全生产科学技术*, 2015(3):172-178.  
Peng Shijin, Tang Haiping, Ding Yaping, *et al.* Study on characteristics of safety assessment model on pipeline corrosion[J]. *Journal of Safety Science and Technology*, 2015(3):172-178(in Chinese).
- [8] 高强, 赵阳. 工业循环冷却水腐蚀速率模型的研究[J]. *工业水处理*, 2015, 35(6):65-67.  
Gao Qiang, Zhao Yang. Study on the corrosion rate model applied to industrial circulating cooling water[J]. *Industrial Water Treatment*, 2015, 35(6):65-67(in Chinese).
- [9] 董超, 张凌峰, 李荣, 等. 循环冷却水系统腐蚀速率的预测模型[J]. *化工自动化及仪表*, 2011, 38(6):676-678.  
Dong Chao, Zhang Lingfeng, Li Rong, *et al.* Corrosion rate prediction model for circulating cooling water system[J]. *Control and Instruments in Chemical Industry*, 2011, 38(6):676-678(in Chinese).
- [10] 李佳龙, 阎威武, 白晓慧, 等. 基于数据驱动与信息增益率多指标决策模型的生活饮用水水质评估[J]. *大连理工大学学报*, 2016, 56(1):92-97.  
Li Jialong, Yan Weiwu, Bai Xiaohui, *et al.* Multiple criteria decision-making model based on data driven and information gain ratio for drinking water quality evaluation [J]. *Journal of Dalian University of Technology*, 2016, 56(1):92-97(in Chinese).
- [11] Sohn J, Amy G, Cho J. Disinfectant decay and disinfection by-products formation model development: chlorination and ozonation by-products [J]. *Water Res*, 2004, 38(10):2461-2478.
- [12] Bai X H, Li R L, Zhu H F, *et al.* Simultaneous corrosion control and water quality improvement in water distribution systems [J]. *Water Practice and Technology*, 2017, 12(4):790-796.
- [13] Bai X H, Wang Y, Li R L, *et al.* Corrosion control and inhibition mechanism of sodium polyphosphate on highly eroded cement coating pipe surfaces [J]. *Desalination and Water Treatment*, 2018, 113:81-87.



作者简介:李瑞莲(1972-),女,山西阳泉人,硕士,讲师,主要研究方向为信号处理与模拟仿真。

E-mail: ruilian\_li@163.com

收稿日期:2018-12-13