doi:10.13582/j.cnki.1672-9102.2017.03.010

基于非线性映射的污水浓度软测量

张志飞[1],刘士亚[1],周少武[2]

(1.佛山科学技术学院 自动化学院,广东 佛山 528000;
2.湖南科技大学 信息与电气工程学院,湖南 湘潭 411201)

摘　要:基于污水处理厂减少监测污水装置的要求,提出了一种以相对误差最小为性能指标的污水浓度预测方法,该方法首先将低维空间的数据映射到高维空间,然后在高维空间上建立线性预测模型.最后给出了应用实例,并与传统的最小二乘法和当前热门的神经网络方法的结果进行了比较,结果表明本文方法结构简单而且有效.

关键词:软测量;相对误差;非线性映射;污水浓度

中图分类号:TP302.1　　**文献标志码**:A　　**文章编号**:1672-9102(2017)03-0054-06

Concentration soft-sensor of waste-water based on nonlinear mapping

Zhang Zhifei[1], Liu Shiya[1], Zhou Shaowu[2]

(1. College of Automatic Control, Foshan University, foshan 528000, China;
2. School of Information and Electronic Engineering, Hunan University of Science and Technology, Xiangtan 411201, China)

Abstract: A prediction approach of waste-water concentration was proposed, in order to reduce costs or decrease monitoring devices, by minimum relative error performance index. Firstly, non-linearly data on low dimensional space was mapped into high dimensional space, then established the linear prediction model on the high dimensional space, finally an application example was provided to illustrate the method, and comparisons was made with the traditional least squares method or the current popular neural network method. Simulation shows the method is simple and effective.

Keywords: Soft-sensing; Relative error; Nonlinear mapping; Sewage concentration

在污水处理过程中,监控污水浓度的变化是监管的必要,也是衡量污水处理设备效率的依据.由于测量某种浓度的设备价格昂贵,或者测量某种污水浓度由于技术方面的要求而严重滞后,有些指标在极端工作环境下,由于干扰使在线测量严重失真,因此,软测量技术显得非常必要,近年来得到了广泛的应用[1].目前,主要的软测量方法可分为统计模型和人工智能模型.建模的依据是样本能够充分体现污水处理系统的所有特性或者至少代表了系统的主要特性.但样本总量的有限性,待预测参数与已知样本数据间的非线性,给研究带来了困难,因而提高模型的泛化能力和自适应能力是确保模型抗干扰能力和确保模型可靠性的基本要求.各种模型各有特点,很多学者采用混合建模的方法进行软测量,其中神经网络由于其具有很强的非线性映射能力、自学习能力和鲁棒性,因此是目前软测量领域中研究中最为活跃的分支[2-3].神经网络的有效性和合理性,除了结果具有较好的精度外,还必须有合理的机理解释.值得指出的是,在神经网络如径向基核函数中,聚类中心的确定目前仍然没有规范可行的确定方法,在数据处理中,为了确保核函数的有效数字,常常需要对数据进行预处理,使所有数据处于[-1, 1]或者[0, 1]之内[4-5],通常操作运算的规则将各种浓度除以其中的最大浓度值,这显然是不符合量纲原理的,因为同性质的数据或无量纲数才

收稿日期:2016-07-13
基金项目:国家自然科学基金资助项目(51577057)
通信作者:张志飞(1963-),男,湖南益阳人,博士,教授,主要从事人工智能非线性系统方面的研究. E-mail:zhifeizhang@ sina.com

能进行算术运算.另一方面,对于采用高斯核函数的模型,只要输入与聚类中心不同,神经元均有输出,这是不合理的,因为神经元在较小的刺激下,是不会有反应的,因而没有输出.从方法的硬件实现来讲,神经网络的非线性规律给实现带来了难度.本文将文献[6]和文献[7]中低维空间的数据映射到高维空间的思想扩展到污水处理浓度的预测问题,然后以相对误差均方和最小为性能指标,获取高维空间的回归模型.最后给出的实例主要讨论了以多项式映射[6]表和对数映射[7]2 种映射下的预测效果.

1　模型结构及算法

给定污水浓度样本 $\boldsymbol{X}=[x_1,x_2,\cdots,x_l]^{\mathrm{T}}\in\boldsymbol{R}^{l\times n}$,$x_i=[x_{i1},x_{i2},\cdots,x_{in}]\in\boldsymbol{R}^{1\times n}$ 为第 $i(i=1,2,\cdots,n)$ 个样本,$y_i\in\boldsymbol{R}^{1\times k}(i=1,2,\cdots l)$ 为第 i 个样本的输出(相当于需要软测量的浓度参数),l 为样本总数,n 为已知的测量指标个数,k 为待测参数的个数.不失一般性,设 $\boldsymbol{X}$ 是列满秩的,即 $\mathrm{rank}(\boldsymbol{X})=n$. 设待测试样本 $\boldsymbol{T}=[x_1,x_2,\cdots,x_m]^{\mathrm{T}}\in\boldsymbol{R}^{m\times n}$,其中 m 为测试样本总数.

对任取向量 $x\in\boldsymbol{R}^n$,下列映射:

$$x:\rightarrow\varphi(x)\in\boldsymbol{R}^N \tag{1}$$

在 N 维空间上,如果输出 y 与 x 存在线性关系,即

$$\hat{y}=\varphi(x)\boldsymbol{\omega} \tag{2}$$

式中:$\boldsymbol{\omega}$ 为回归系数或权重系数.

定义相对误差

$$e_k=\frac{\hat{\boldsymbol{y}}_k-\boldsymbol{y}_k}{\boldsymbol{y}_k}. \tag{3}$$

式中:$\boldsymbol{y}_k$ 为样本 x 对应的实际输出.

取性能指标

$$J=\sum_{k=1}^{l}e_k^2=\sum_{k=1}^{l}\frac{(\hat{\boldsymbol{y}}_k-\boldsymbol{y}_k)^{\mathrm{T}}(\hat{y}_k-\boldsymbol{y}_k)}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k}. \tag{4}$$

将式(2)代入式(4),有

$$J(\omega)=\sum_{k=1}^{l}e_k^2=\sum_{k=1}^{l}\left(\boldsymbol{\omega}^{\mathrm{T}}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)\boldsymbol{\varphi}_k(x)}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k}\omega-2\boldsymbol{\omega}^{\mathrm{T}}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)}{\boldsymbol{y}_k}\right)+l. \tag{5}$$

最小化式(5)的条件是 $\frac{\partial J}{\partial\boldsymbol{\omega}}=0$,得到

$$\sum_{k=1}^{l}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)\boldsymbol{\varphi}_k(x)}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k}\boldsymbol{\omega}=\sum_{k=1}^{l}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)}{\boldsymbol{y}_k}. \tag{6}$$

记

$$\boldsymbol{\Phi}(X)=\sum_{k=1}^{l}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)\boldsymbol{\varphi}_k(x_k)}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k},\eta=\sum_{k=1}^{l}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)}{\boldsymbol{y}_k}. \tag{7}$$

借用式(7),重写式(6):

$$\boldsymbol{\Phi}(X)\boldsymbol{\omega}=\eta. \tag{8}$$

由线性方程理论,有如下结论[7]:定理:式(8)有解的充分必要条件是 $\mathrm{rank}(\boldsymbol{\Phi}(X))\leqslant N$,取等号时有唯一解:

$$\boldsymbol{\omega}=\boldsymbol{\Phi}^{-1}(X)\eta. \tag{9}$$

取<号时,有无穷多组解,其中具最小范数的解为

$$\boldsymbol{\omega}=\boldsymbol{\Phi}^{+1}(X)\eta. \tag{10}$$

这里 $\boldsymbol{\Phi}^-,\boldsymbol{\Phi}^+$ 分别表 $\boldsymbol{\Phi}$ 的逆和广义逆.

注 1:由于 $\boldsymbol{\Phi}(X)\in R^{N\times N}$,故 $\mathrm{rank}(\boldsymbol{\Phi}(X))\leqslant N$ 条件总是可以得到满足,$\boldsymbol{\omega}$ 总是有解的,理论上,若 $\boldsymbol{X}$ 是列满秩的,则 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 可逆,但实际计算中 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 的条件数很大,常常是奇异的,因而式(10)的使用更普遍.

注 2:定义 $e_k=\hat{y}_k-\boldsymbol{y}_k$,相应地性能指标变为

$$J(\omega)=\sum_{k=1}^{l}e_k^2=\sum_{k=1}^{l}(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)\boldsymbol{\varphi}_k(x)\boldsymbol{\omega}-2\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)\boldsymbol{y}_k+\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k). \tag{11}$$

使式(11)取最小的结论,就是著名的最小二乘法结果,此时式(7)相应地就为

$$\boldsymbol{\Phi}(X)=\sum_{k=1}^{l}\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)\boldsymbol{\varphi}_k(x_k),\boldsymbol{\eta}=\sum_{k=1}^{l}\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k)\boldsymbol{y}_k. \tag{12}$$

最小二乘法侧重于测量数值较大的数据处理,即当测量值出现异常时有较好的预测效果,因而适应于报警处理系统,但另一方面,系统正常工作时参数预测往往误差偏大.

注 3:采用不同的映射 $\varphi(x)$,定理就是神经网络的结果,如取 $\varphi(x)=\exp(-\frac{x-c^2}{2\sigma^2})$,定理即演变成径向基函数神经网络.采用径向基函数的优点之一是可以确保解的唯一性,另一方面可使模型误差任意小,但经验证明,后者的泛化能力相对较差.文后的实例研究了多项式映射与对数映射的差异.

注 4:试图通过对输入数据的线性处理(在神经网络中,相当于增加一隐含层专门处理输入的数据,我们这里指出,这种处理也许可提高预测模型的泛化能力,但理论上讲,对提高预测精度是不可靠的.考虑如下映射:

$$x:\rightarrow\varphi(x\lambda). \tag{13}$$

式中,$\lambda=diag(\lambda_1,\cdots,\lambda_n)$ 为对角矩阵;$\lambda_i(i\in n)$ 为对第 i 个输入参数处理因子.式(13)也可以看成是先对样本进行的预处理,以期望解决不同参数对输出的影响程度.此时,性能指标变为

$$J(\boldsymbol{\omega},\lambda)=\sum_{k=1}^{l}e_k^2=\sum_{k=1}^{l}(\boldsymbol{\omega}^{\mathrm{T}}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k\lambda)\boldsymbol{\varphi}_k(x\lambda)}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k}\boldsymbol{\omega}-2\boldsymbol{\omega}^{\mathrm{T}}\frac{\boldsymbol{\varphi}_k^{\mathrm{T}}(x_k\lambda)}{\boldsymbol{y}_k})+l. \tag{14}$$

式(14)极小的条件是

$$\frac{\partial J}{\partial\boldsymbol{\omega}}=0,\frac{\partial J}{\partial\lambda}=0. \tag{15}$$

为方便说明,特别地取 $\varphi(x)=(e^{x_1},\cdots,e^{x_n})$[6],并注意到

$$\frac{\partial[\boldsymbol{\varphi}(x\lambda)^{\mathrm{T}}\boldsymbol{\varphi}(x\lambda)]}{\partial\boldsymbol{\lambda}_i}=\frac{\partial\boldsymbol{\varphi}^{\mathrm{T}}}{\partial\boldsymbol{\lambda}_i}\boldsymbol{\varphi}+\boldsymbol{\varphi}^{\mathrm{T}}\frac{\partial\boldsymbol{\varphi}}{\partial\boldsymbol{\lambda}_i}=\boldsymbol{\omega}_i x_i\varphi_i\boldsymbol{\varphi}\boldsymbol{\omega}+\boldsymbol{\omega}_i x_i\varphi_i\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\varphi}^{\mathrm{T}}. \tag{16}$$

将式(16)代入式(15),有

$$\frac{\partial J}{\partial\boldsymbol{\lambda}_i}=2\boldsymbol{\omega}_i x_i\varphi_i\sum_{k=1}^{l}\frac{\varphi_k\boldsymbol{\omega}-\boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k}=2\boldsymbol{\omega}_i x_i\varphi_i\sum_{k=1}^{l}\frac{\hat{y}_k-\boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{y}_k}=2\boldsymbol{\omega}_i x_i\varphi_i J>0. \tag{17}$$

式(17)表明,试图通过对输入数据的处理来获取精度是不可能的.

需要说明的是,式(17)虽然是从取特例映射所获得的结果,其它形式的映射,如对径向基 $\varphi(x)=\exp(-\frac{x-c^2}{2\sigma^2})$,除推导过程稍复杂一些,结论相同.定理的算法流程如图 1.

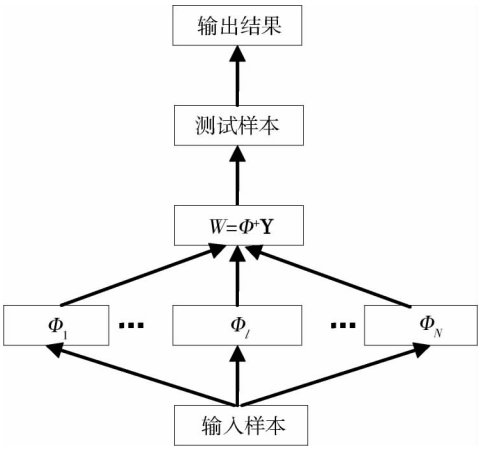

图 1 算法流程

2 污水处理浓度软测量

2.1 数据来源

活性污泥污水处理(Waste water Treatment Plant,WWTP)流程如图 2,包括预处理、初沉、曝气、二沉等

4 个部分,经过预处理和初沉池处理后,污水进入曝气池,经过好氧和厌氧生物的生化反应分解污水中的有机物,主要用来分解污水中的有机物并脱氮降磷,主要监测指标是反映水中可降解有机物参数的生化需氧量(BOD).由于污水处理的生化反应过程极其复杂,微生物数量与种类受到污水浓度、进水速率、天气、季节等变化的影响时刻变化,在线分析检测周期较长等原因,使用在线分析仪器检测 BOD 的效果并不理想[8].本节利用前面所述模型来预测 BOD_5,加州大学数据库(UCI)对某日均处理污水流量为 35 000 m^3/d,38 个与有机物和微生物相关的变量,共记录了 526 天的 38×526 个在线检测数据.本文通过分析聚类,选取 14 个参数作为预测 BOD_5,训练样本数 368 个,测试样本数 123 个.

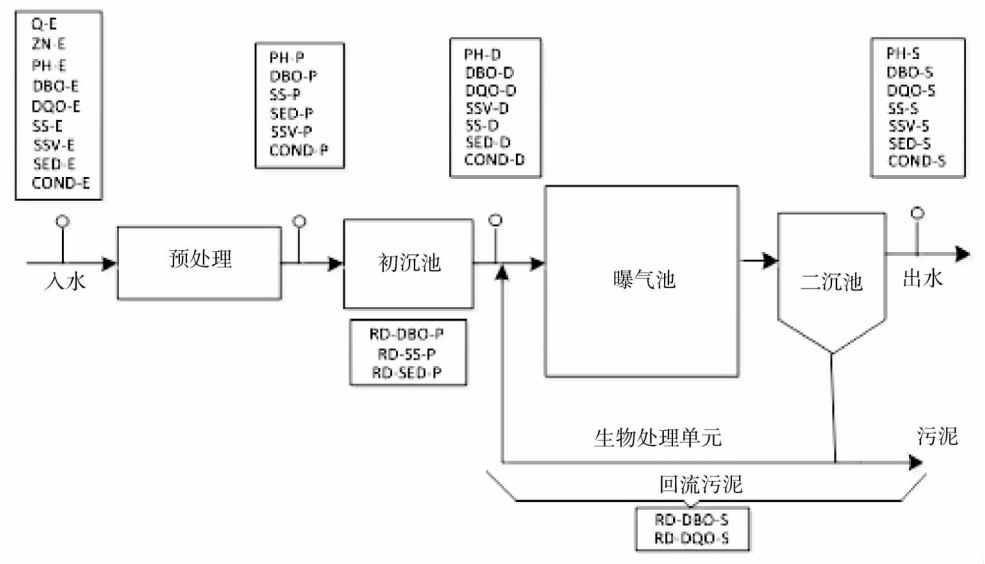

图 2　活性污泥污水处理流程

2.2　预测结果分析比较

对给定的预测精度 α,定义预测准确率 β 为

$$\beta = \frac{K}{L}. \tag{18}$$

式中:L 为预测的样本总数;K 为预测相对误差小于 $1-\alpha$ 的测试样本个数.

2.2.1　多项式映射最小相对误差理论与最小二乘结果的比较

从实践经验,BOD_5 与各参数指标三阶以上的关系不显著,因此本节选择二次多项式映射:$\varphi(x):x \to x \oplus x \cdot x$ 式中 $x \cdot x := \{x_i x_j\}, i,j \in n, x \oplus y = [x \quad y]$.

表 1 列出了以下 3 种二次多项式映射情形下测试结果.

(1) $\varphi(x) = x$, $(N=n)$;(2) $\varphi(x) = x \oplus x \cdot x(i=j)$ $(N-n=14)$;(3) $x \cdot x := \{x_i x_j\}, i,j \in n$, $(N-n=104)$.

表 1　多项式映射最小相对误差理论及预测准确率 β 分析结果

评价指标	N-n	平均相对误差	平均均方误差	精度>95%/%	精度>85%/%	精度>80%/%
最小二乘法	0	0.077 4	0.166 5	48	89	94
本文定理	0	0.073 6	0.164 6	46	91	97
最小二乘法	14	0.078 9	0.051 3	48	86	89
本文定理	14	0.078 7	0.051 3	51	88	95
最小二乘法	104	0.083 0	0.152 1	53	87	91
本文定理	104	0.090 4	0.205 0	53	83	91

表 1 结果表明:对于精度要求 85%以下的要求,本文方法比传统最小二乘法的准确率有较好的改进,对于精度要求 95%以上的要求,二者精度相当.当映射单元过多(相当于神经元多),效果反而略低于最小二乘法,这说明训练模型精度高时,会降低模型的泛化能力.最小二乘预测结果和多项式数映射的预测效

果参见图 3 和图 4.

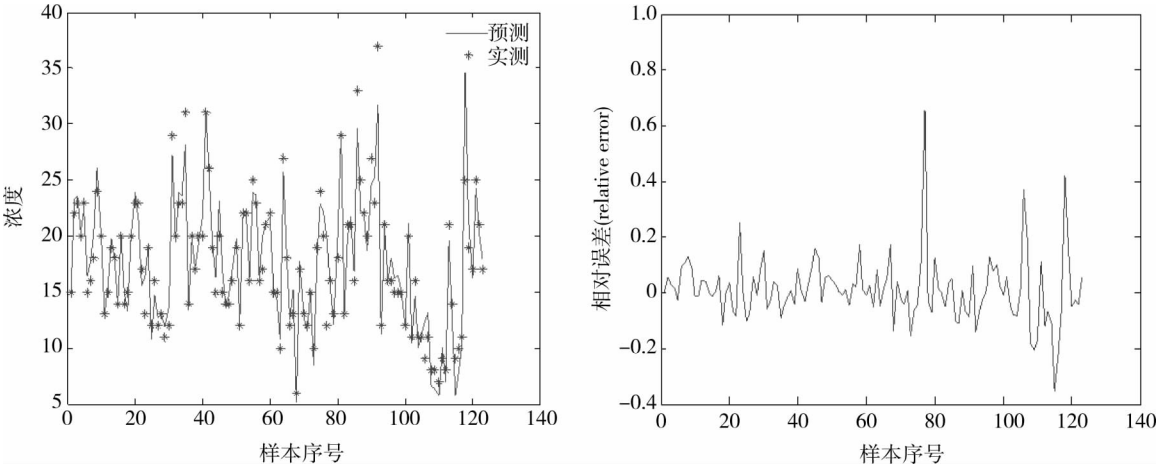

图 3 最小二乘法 BOD 浓度预测

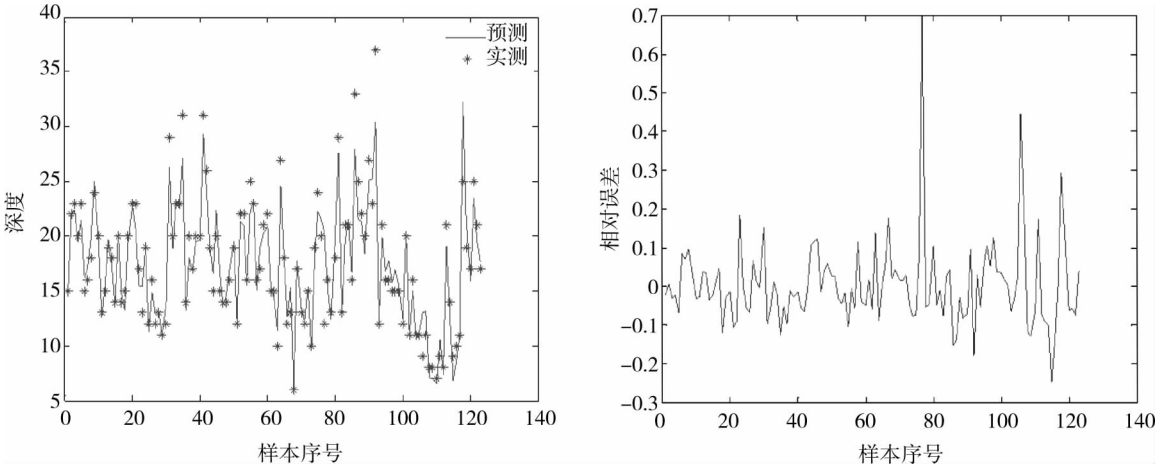

图 4 多项式映射 BOD 浓度预测

2.2.2 对数映射相对误差与最小二乘结果的比较

取类似文献[9]的对数非线性映射:

$$\varphi(x) = x \oplus 10\log 10(x). \tag{19}$$

分析结果如下表 2 所示.

表 2 对数映射分析结果分析表

评价指标	N-n	平均相对误差	平均均方误差	精度>95%/%	精度>85%/%	精度>80%/%
最小二乘法	0	0.077 4	0.055 6	*	*	*
本文定理	0	0.073 6	0.055 0	*	*	*
最小二乘法	14	0.088 7	0.159 9	48	86	89
本文定理	14	0.076 7	0.146 6	50	88	95
最小二乘法	104	0.084 7	0.155 6	44	85	91
本文定理	104	0.091 9	0.209 9	37	81	91

表 2 中 * 号同表 1 的结果,对于精度要求 85%以下的要求,对数映射比传统最小二乘法的准确率有较好的改进,对于精度要求 95%以上的要求,映射单元多时,效果反而不如简单的最小二乘法结果,再次说明训练模型精度高时,会降低模型的泛化能力.对数映射的预测效果参见图 5.

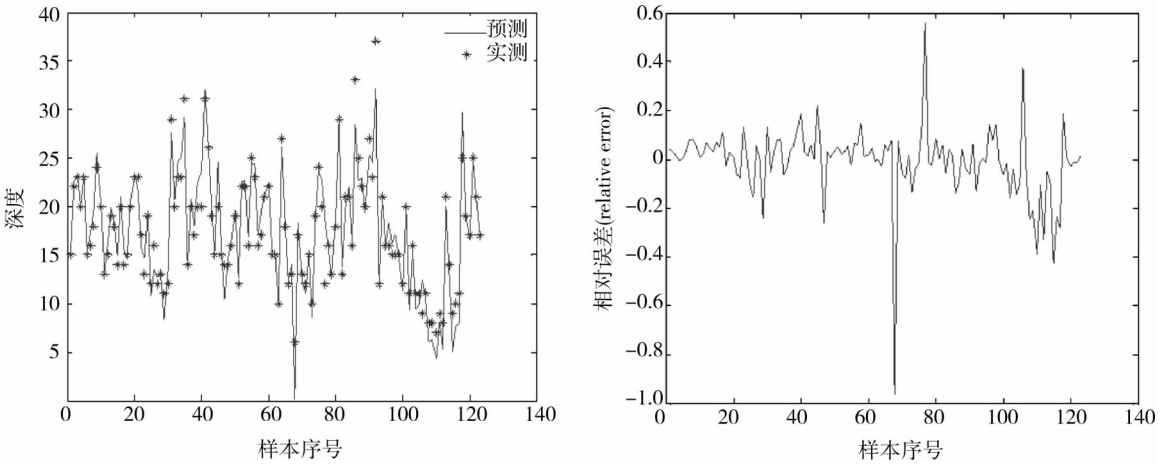

图 5　对数映射 BOD 浓度预测值和误差图

3　结论

本文提出了一种基于相对误差均和为性能指标的非线性映射污水处理浓度软测量方法,该算法与以往神经网络不同之处在于不必要求映射满足对称条件,解的不唯一性使提高模型的泛化功能成为可能.实际算例与经典方法比较表明,本文所提出的方法特别适应中等规模数据量的预测,如何进一步利用深度学习的方法提高模型的预测精度以及处理海量数据的在线预测是是今后需要继续进行深入的课题.

参考文献:

[1] Liu Y, Huang D, Li Y. Development of interval soft sensors using enhanced just-in-time learning and inductive[J]. Industrial & Engineering Chemistry Research. 2012,51(8):3356-3367.

[2] 朱群雄,甄玉山.用于化工软测量的基于移动窗的过程神经网络[J]. 清华大学学报(自然科学版), 2012,52(9):1165-1170.

[3] Qi H Y , Zhou X G , Liu L H , et al. A hybrid neural network-first principle model for fixed-bed reactor[J]. Chemical Engineering Science, 1999,54 (13/14) :2521-2526.

[4] 张勇,王介生. 基于 PCA-RBF 神经网络的浮选过程软测量建模[J]. 南京航空航天大学学报, 2006, 38(7):116-119.

[5] 刘瑞兰,苏宏业,褚健. 模糊神经网络的混合学习算法及其软测量建模[J].系统仿真学报,2005,17(12): 2878-2881.

[6] 张登峰,刘士亚,叶树林.基于多项式映射的分类器及其在变压器故障诊断中的应用研究[J].高压电器,2016(6):103-108.

[7] 张登峰, 张志飞, 章兢. 基于非线性映射的分类器及其在变压器故障诊断中的应用研究[J]. 湘潭大学自然科学学报, 2015, 37(3):82-92.

[8] 乔俊飞,郭楠,韩红桂.基于神经网络的 BOD 参数软测量仪表的设计[J].计算机与应用化学,2013,30(10): 1219-1222.

[9] Ganyun L V, Haozhong C, Haibao Z, et al. Fault diagnosis of power transformer based on multi-layer SVM classifier[J]. Electric Power Systems Research, 2005, 74(1): 1-7.