

DOI :10.3969/j.issn.1004-6933.2011.04.004

# 基于 GA 参数优选的 $\epsilon$ -SVR 地下水位预测方法

陈海洋,滕彦国,王金生

(北京师范大学水科学研究院,北京 100875)

**摘要** 选择径向基核函数建立地下水位  $\epsilon$ -SVR 预测模型,基于遗传算法实现惩罚因子  $C$ 、核函数参数  $\gamma$  和不敏感损失函数参数  $\epsilon$  的自适应优选,并运用建立的模型对某地傍河试验井地下水位进行预测。结果表明:基于 GA 参数优选的  $\epsilon$ -SVR 模型对训练样本的拟合误差平方和仅为 0.0022,回归系数达到 0.9933,检验样本拟合结果平均相对误差仅为 1.28%。这与人工神经网络模型相比,无论是对训练样本的拟合能力,还是对检验样本的泛化能力均有较大程度的提高,说明基于 GA 参数优选的  $\epsilon$ -SVR 模型可以很好地应用于进行地下水位预测。

**关键词** 支持向量回归机;遗传算法;地下水位预测;地下水管理

**中图分类号** :P641.2      **文献标识码** :A      **文章编号** :1004-6933(2011)04-0015-04

## Method for prediction of groundwater level based on epsilon-support vector regression machine with parameters optimized by genetic algorithm

CHEN Hai-yang, TENG Yan-guo, WANG Jin-sheng

(College of Water Sciences, Beijing Normal University, Beijing 100875, China)

**Abstract** : Taking the radial basis function (RBF) as a kernel function, a prediction model of groundwater level was developed based on the epsilon-support vector regression ( $\epsilon$ -SVR). The optimization of the model parameters, including punishment factor  $C$ , kernel function parameter  $\gamma$  and non-sensitivity loss function parameter  $\epsilon$ , was realized by genetic algorithm. The model was applied to predict the groundwater level for a riverside test well. The results show that the sum of error square of the model for training is only 0.0022 with regression coefficient of 0.9933, and the average relative error is 1.28% for prediction. Compared with the artificial neural network model, the model improves both fitting capacity for the training sample and generalization capacity for the test sample. Therefore, it can be used to better predict the groundwater level.

**Key words** : support vector regression machine; genetic algorithm; groundwater level prediction; groundwater management

地下水位预测可以看作是对地下水位及其影响因子间的复杂非线性函数关系的逼近分析,它是实现地下水资源动态均衡管理的重要方法之一。当前,发展较快、应用较广的地下水位预测方法主要有多元回归分析法、灰色聚类分析法、模糊模式识别法、季节性指数平滑法、人工神经网络法等。这些不同的预测方法均力图通过对地下水位动态及其影响因素之间的非线性关系的分析,建立随机性模型来

实现地下水位的预报。支持向量机(support vector machine,简称 SVM)是由统计学习理论发展起来的一种新型学习机器,它以结构风险最小化原理为理论基础,具有以下优点:逼近复杂非线性系统,较强的学习泛化能力和良好的分类性能,所需样本少,建模方便,计算简单,学习训练时间短,通用性强,可以用于解决非线性系统的模式识别问题<sup>[1]</sup>。 $\epsilon$ -SVR( $\epsilon$ -support vector regression machine)是将不敏感损失函数

引入支持向量机,用来解决非线性系统回归拟合问题的一个算法,可以用之来解决地下水位预测问题。

应用  $\epsilon$ -SVR 进行地下水位预测,需要重点考虑模型参数的优化选择。然而,当前对  $\epsilon$ -SVR 参数选择缺乏实质性的理论指导,只能是通过反复试验,人工选取参数,这不仅要求技术人员有丰富的实际经验,而且需要付出较高的时间代价。笔者选择径向基核函数建立了地下水位  $\epsilon$ -SVR 预测模型,基于遗传算法(genetic algorithm,简称 GA)实现了惩罚因子  $C$ 、核函数参数  $\gamma$  和不敏感损失函数参数  $\epsilon$  的自适应优选,并运用建立的模型对某地傍河试验井地下水位进行了预测。

## 1 支持向量回归机

SVM 是由 Vapnik 等在统计学习理论的基础上提出并发展起来的一种新型通用学习方法,它基于结构风险最小化原理,主要用于解决分类和回归问题<sup>[2]</sup>。SVM 避开了从归纳到演绎的传统过程,实现了高效的从训练样本到预报样本的转导推理,大大简化了回归问题的复杂性,具有较高的泛化能力。关于 SVM 的原理,有很多资料可以参阅,这里做简要介绍。

根据结构风险最小原理,回归估计问题就是寻找使下面函数风险最小的  $f(x)$

$$\min(1/2\|w\|^2 + CR_{emp})$$

式中  $1/2\|w\|^2$  反映回归函数  $f(x)$  的泛化能力,是正则化部分; $C$  为惩罚因子; $R_{emp}$  为经验风险,即样本损失函数的累积。对于  $D$  维空间的训练集  $\{X_i, Y_i\} i=1, 2, \dots, k, k$  为样本总数; $X_i \in R^d, Y_i \in R$ , 构造最优线性函数: $y = wx + b$  ( $w$  为分割平面的法向量,  $b$  为分割平面的偏移量),使得所有训练数据都可以在允许的精度内用该线性函数拟合,且使得所有样本点离超平面的“总偏差”最小。引入允许拟合误差松弛因子  $\xi_i$  与  $\xi_i^*$ , 有如下目标函数<sup>[3]</sup>:

$$\begin{cases} \min(1/2\|w\|^2 + C \sum_{i=1}^k (\xi_i + \xi_i^*)) \\ \text{s.t. } y_i - wx_i - b \leq \epsilon + \xi_i \\ wx_i + b - y_i \leq \epsilon + \xi_i^* \quad i = 1, 2, \dots, k \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases} \quad (1)$$

该问题是一个典型的凸二次优化问题,引入拉格朗日函数,可得对偶形式:

$$\max Q(\alpha, \alpha^*) = -\frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) (x_i \cdot x_j) - \epsilon \sum_{i=1}^k (\alpha_i + \alpha_i^*) + \sum_{i=1}^k (\alpha_i - \alpha_i^*) y_i$$

$$\text{s.t. } \sum_{i=1}^k (\alpha_i - \alpha_i^*) = 0 \quad 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, \dots, k \quad (2)$$

式中  $\alpha, \alpha^*$  是 Lagrange 乘子。估值函数为:

$$f(x) = \sum_{i=1}^k (\alpha_i - \alpha_i^*) (x_i \cdot x_j) + b \quad (3)$$

偏置量  $b$  可通过 KKT 条件计算:

$$b = \begin{cases} y_i - \sum_{i=1}^k (\alpha_i - \alpha_i^*) (x_i \cdot x) - \epsilon, \alpha_i \in (0, C) \\ y_i - \sum_{i=1}^k (\alpha_i - \alpha_i^*) (x_i \cdot x) + \epsilon, \alpha_i^* \in (0, C) \end{cases} \quad (4)$$

对应非线性回归支持向量机,可通过非线性映射把原始数据变换到高维特征空间,在特征空间设计线性回归支持向量机,运用原空间的函数来实现内积运算,从而取得在原空间非线性回归的效果。假设用非线性函数  $\Phi(x)$  把样本映射到高维空间,可得拟合样本集的估计函数为:

$$f(x) = \sum_{i=1}^k (\alpha_i - \alpha_i^*) (\Phi(x_i) \cdot \Phi(x_j)) + b \quad (5)$$

根据泛函理论,只要一种核函数满足 Mercer 条件,它就对应某一空间中的内积,因此只要设计相应的核函数,将原始不具备线性回归的样品数据映射到某一高维特征空间,在特征空间运用内积函数实现线性回归。假设利用核函数  $K(x_i, y_i) = (\Phi(x_i) \cdot \Phi(y_i))$  把原始数据映射到高维特征空间,那么特征空间的核函数线性回归向量机的估计函数为:

$$f(x) = \sum_{i=1}^k (\alpha_i - \alpha_i^*) K(x_i \cdot x_j) + b \quad (6)$$

很明显,不同的核函数可以构造实现输入空间不同类型的非线性决策面,它类似于一个 3 层前馈神经网络,其中隐节点对应于输入样本与一个支持向量的内积核函数,而输出节点对应于隐层输出的线性组合。常见的满足 Mercer 条件的核函数有线性核函数、多项式核函数、径向基核函数(简称 RBF)、多层感知机核函数(Sigmoid)。

## 2 基于 GA 的参数优化

笔者选取径向基函数,构建  $\epsilon$ -SVR 地下水位预测模型,基于遗传算法,对其中的惩罚因子  $C$ 、核函数参数  $\gamma$  和不敏感损失函数参数  $\epsilon$  等 3 个重要参数进行自适应寻优。

### 2.1 主要参数分析

惩罚因子  $C$  主要用来调和机器学习置信范围和经验风险的比例,以使学习机器的推广能力最好。由式(3)可知,惩罚因子  $C$  并没有出现在式(2)的对

偶形式中,而是改变了拉格朗日系数的取值范围,因此,对于一个  $\epsilon$ -SVR,如果无限制地增大惩罚因子  $C$ ,当  $\epsilon$ -SVR 中没有边界支持向量时, $C$  的改变不会再影响回归性能。通常,在确定的数据子空间中, $C$  的取值小表示对经验误差的惩罚小,学习机器的复杂度小而经验风险值较大;反之亦然。前者被称为“欠学习”现象,而后者则被称为“过学习”现象。每个数据子空间至少存在一个合适的  $C$  使得 SVM 推广能力最好。当  $C$  超过一定值时,SVM 的复杂度达到了数据子空间允许的最大值,此时经验风险和推广能力几乎不再变化。

核函数参数  $\gamma$  同样对模型精度有着重要意义。核函数是原始非线性样本在高维特征空间的映射函数,核参数的改变实际上是隐含地改变映射函数,从而改变决定最大 VC 维的样本数据子空间分布的复杂程度,也就决定了超优分类面能达到的最小经验误差。从径向基核函数的表现形式可以看出,参数  $\gamma$  相当于对样本间欧式距离的归一化,判定了特征空间中向量间的距离。当  $\gamma$  趋向于 0 时会发生严重的“过学习”现象,此时  $\epsilon$ -SVR 能实现允许精度下对训练样本的回归拟合,但对测试样本的泛化能力变得极差;当  $\gamma$  趋向于无穷大时会发生严重的“欠学习”现象,此时  $\epsilon$ -SVR 的训练回归拟合能力变弱<sup>[4]</sup>。

不敏感损失函数参数  $\epsilon$  控制不敏感区域的宽度,它的值影响用于建立回归估计函数支持向量的个数,影响对偶变量的稀疏性<sup>[5]</sup>。合适的参数值能确保全局最小解和可靠泛化界的优势,但参数  $\epsilon$  数值偏大将产生相对较少的支持向量,从而导致回归估计函数比较简单,性能降低。

## 2.2 基于 GA 的参数寻优算法

GA 是一种基于自然选择和群体遗传机理的新的全局优化求解方法,它模拟了自然选择和自然遗传过程中发生的繁殖、杂交和突变现象,能把自然界有机体的优胜劣汰的自然选择、适者生存的进化机制和在同一群体中个体之间的随机信息交换机制相结合,用以解决复杂问题<sup>[6]</sup>。GA 优化参数步骤如下:

a. 确定决策变量和约束条件。决策变量包括惩罚因子  $C$ 、核函数参数  $\gamma$  和不敏感损失函数参数  $\epsilon$ 。其中,核函数参数  $\gamma$  的搜索区间为  $[\min(\|x - x_j\|^2 \times 10^{-2}), \max(\|x - x_j\|^2 \times 10^{-2})]$ ;不敏感损失函数参数  $\epsilon$  的搜索区间为  $[0.01, 0.20]$ ;惩罚因子  $C$  的搜索区间采取如下方法确定<sup>[7-8]</sup>。选定某一足够大的  $C$  值,用该值训练  $\epsilon$ -SVR 求解出一组  $\alpha_i (i = 1, 2, \dots, k)$ ,令  $C_1 = \max(\alpha_i)$ 。若  $C_1 < C$  则以  $C_1$  作为  $C$  的搜索区间上界,若  $C_1 \geq C$  表明  $C$  仍对  $\alpha_i$  起约束作

用,需要选择一个更大的  $C$  训练 SVM,直至得到的  $C_1$  远小于  $C$  为止,由此可以确定  $C$  的搜索区间  $(0, C_1)$ 。

b. 染色体编码。染色体  $r$  采用二进制编码。每个算子由 3 段编码组成。即:  $r = A_1, A_2, A_3$ ,其中  $A_1, A_2, A_3$  分别为惩罚因子  $C$ 、核函数参数  $\gamma$  和损失函数参数  $\epsilon$  的代码表示段。

c. 初始群体的生成。随机产生  $N$  个初始串结构数据,每个串结构数据即为一个个体, $N$  个个体构成一个初始群体  $P(0)$ 。遗传算法以这  $N$  个串结构作为初始点开始迭代。

d. 适应度值评价检测。计算群体  $P(t) (t = 0, 1, 2, \dots, n; n$  表示遗传代数)中各个个体的适应度。遗传算法按与个体适应度成正比的概率来决定当前群体中各个个体遗传到下一代群体中的概率。适应度函数定义如下<sup>[9]</sup>

$$\begin{cases} \text{fit}(C, \gamma, \epsilon) = Y_{\min} - Y_i \\ Y_i = (y_i - y_i^*)^2 \end{cases} \quad (7)$$

式中:  $Y_{\min}$  为某染色体  $r_j$  所在世代的最小值;  $y_i$  为样品数据实际值;  $y_i^*$  为  $\epsilon$ -SVR 的计算结果。

e. 选择、交叉、变异算子。选择合适的算法,对群体中的个体使用选择、交叉、变异操作,使其产生经过遗传优化的下一代群体。

f. 终止条件判断。群体  $P(t)$  经过选择、交叉、变异后形成下一代群体  $P(t+1)$ 。若迭代次数大于某设定值或误差精度达到预定值,则中止迭代。

## 3 实例研究

利用  $\epsilon$ -SVR 进行地下水位预测,首先要确定影响地下水位的主要影响因素,构建地下水位  $\epsilon$ -SVR 预测模型;再通过调查或实测,选择包括影响因素在内的观测资料,建立训练样本集、验证样本集,利用  $\epsilon$ -SVR 模型进行学习训练,并通过参数优化确定惩罚因子及核函数参数。训练学习好的  $\epsilon$ -SVR 模型即可用于待测样本的回归拟合。

### 3.1 $\epsilon$ -SVR 模型构建

利用文献 [10] 中某地傍河试验井 24 个月地下水位实测序列值作为  $\epsilon$ -SVR 分析的样本数据,选取河道流量、气温、饱和差、降水量、蒸发量 5 个因子作为影响该区域地下水位的潜在因素,通过相关性分析进一步确定影响该区域地下水位的因素。利用 SPSS V13 对经过归一化后的样本数据进行相关分析,分析结果见表 1。从表 1 可以看出,气温与地下水位埋深相关性较小,为此,选取河道流量  $X_1$ 、饱和差  $X_2$ 、降水量  $X_3$  和蒸发量  $X_4$  作为  $\epsilon$ -SVR 模型的输入向量的分量,并以预测地下水水位值  $Y$  作为该预

测模型的输出值,选取径向基函数为核函数,建立该区域的地下水位预测  $\epsilon$ -SVR 模型:

已知该区域地下水水位实际测试数据样本点  $\{X_i, Y_i\} (i = 1, 2, \dots, k)$ ; 其中  $X_i \in R^4$  为影响地下水水位预测因素构成的向量,  $Y_i \in R$  为相应的地下水预测水位值。用  $\epsilon$ -SVR 对已知的样本集进行学习训练,得到对应的地下水位预测模型:

$$f(x) = \sum_{i=1}^k (\alpha_i - \alpha_i^*) \exp(-\gamma |x_i - x_j|^2) + b \quad (8)$$

表 1 某地傍河试验井 24 个月地下水位因素相关分析

相关因素	河道流量	气温	饱和差	降水量	蒸发量	地下水位埋深
河道流量	1					
气温	-0.57733	1				
饱和差	-0.88310	0.41612	1			
降水量	-0.65197	0.06915	0.77961	1		
蒸发量	-0.73750	0.77544	0.66122	0.40832	1	
地下水位埋深	-0.77435	0.26314	0.82219	0.87574	0.59004	1

### 3.2 $\epsilon$ -SVR 模型训练及预测

把 24 个样本数据分为 2 类:①训练样本 20 个(序号 1~20);②验证样本 4 个(序号 21~24)。首先对训练样本和验证样本进行归一化处理,以消除单位差异和异常数据的影响;再采用遗传算法对惩罚因子  $C$ 、模型参数  $\gamma$  和不敏感损失函数参数  $\epsilon$  进行寻优,以 10 类折叠交叉验证误差最小作为适应度函数,通过 matlab 编程实现参数自适应选取。实验结果为,惩罚因子  $C$  取值 92.5695;模型参数  $\gamma$  取值 0.3040;不敏感损失函数参数  $\epsilon$  取值 0.0100。最后用确定参数对模型进行训练,并用验证样本进行拟合预测,回归结果见表 2。作为比较,表 2 还列出了文献 [10] 中通过人工神经网络的训练拟合值。

### 3.3 $\epsilon$ -SVR 模型回归分析

从以上回归拟合结果可以看出,经过参数寻优后的  $\epsilon$ -SVR 模型对训练样本(序号 1~20)的拟合程度相当惊人,误差平方和仅为 0.0022,回归系数达到 0.9933。而利用训练好的模型对序号 21~24 的检验样本拟合结果也很好,绝对误差分别为: -0.008、0.02、0.08、-0.013,相对误差则分别为: 1.53%、0.31%、1.30% 和 1.99%,平均相对误差仅为 1.28%。与人工神经网络相比,无论是对训练样本的拟合能力,还是对检验样本的泛化能力均有较大程度的提高。

## 4 结 语

a. 基于结构风险最小化原理的支持向量机具有逼近复杂非线性系统、较强的学习泛化能力,可以很好的用来解决非线性系统分类和回归问题。而基

表 2 地下水位  $\epsilon$ -SVR 模型训练结果

序号	地下水位埋深/m	$\epsilon$ -SVR 拟合结果 <sup>①</sup>		神经网络拟合结果 <sup>②</sup>	
		拟合值/m	绝对误差/m	拟合值/m	绝对误差/m
1	6.92	6.9299	0.01	6.84	-0.08
2	6.97	6.7619	-0.21	6.90	-0.07
3	6.84	6.8295	-0.01	6.66	-0.18
4	6.50	6.4900	-0.01	6.27	-0.23
5	5.75	5.7450	0.00	5.87	0.12
6	5.54	5.5299	-0.01	5.45	-0.09
7	5.63	5.6404	0.01	5.83	0.20
8	5.62	5.6301	0.01	5.67	0.05
9	5.96	5.9702	0.01	6.19	0.23
10	6.30	6.2896	-0.01	6.32	0.02
11	6.80	6.7898	-0.01	6.63	-0.17
12	6.90	6.8901	-0.01	6.61	-0.29
13	6.70	6.7296	0.03	6.65	-0.05
14	6.77	6.7798	0.01	6.81	0.04
15	6.67	6.6801	0.01	6.54	-0.13
16	6.33	6.3404	0.01	6.39	0.06
17	5.82	5.8099	-0.01	6.25	0.43
18	5.58	5.5905	0.01	5.34	-0.24
19	5.48	5.4899	0.01	5.56	0.08
20	5.38	5.3899	0.01	5.43	0.05
21	5.51	5.4255	-0.08	5.52	0.01
22	5.84	5.8579	0.02	5.90	0.06
23	6.32	6.4024	0.08	6.52	0.2
24	6.56	6.4296	-0.13	6.57	0.01

注 ① $\epsilon$ -SVR 训练采用 Matlab7.9.0(R2009b) 编制程序,在配置为 CPU P8700 2.53Ghz 内存 2.5G 的微型计算机上运行通过。②BP 神经网络拟合结果来自文献 [10],本文未作验证。其中序号 6~24 为 BP 神经网络训练拟合值,序号 1~5 为待测样本的拟合值。

于遗传算法对 SVM 模型参数的自适应优选可以提升大大模型的泛化能力。

b. 选择径向基核函数建立了地下水位  $\epsilon$ -SVR 预测模型,基于 GA 实现了惩罚因子  $C$ 、核函数参数  $\gamma$  和不敏感损失函数参数  $\epsilon$  的自适应优选,并运用建立的模型对某地傍河试验井地下水位进行了预测。实验结果表明,基于 GA 参数优选的  $\epsilon$ -SVR 模型可以很好地用于地下水位预测,其回归拟合能力优于人工神经网络。

c. 需要说明的是,影响地下水埋深的因素较多,包括但不限于河道流量、饱和差、降水量、蒸发量等因素。笔者选取这 4 个变量作为模型输入,主要目的是为了验证模型方法的可行性。实际应用中,应根据研究区域的具体水文地质条件,充分考虑地下水埋深影响因素的时空分布特性,选取合适的影响因子进行地下水位预测。

参考文献:

[1] 祁亨年. 支持向量机及其应用研究综述[J]. 计算机工程 2004 30(10) 6-8.  
 [2] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报 2000 26(1) 32-42.

(下转第 73 页)

m<sup>3</sup>;累计引水 14 闸次,引水时间 79.45 h,引水量 2 147.2 万 m<sup>3</sup>。

两次调水期间,长江口南支水质基本达到Ⅱ类水质标准,为调水改善崇明岛内河水质提供了优质的引水水源,17 个水质监测断面的水质指标基本达到Ⅱ类水质标准。但不同区域河道水质改善效果不尽相同:引水口门河道和南横引河较接近长江口南支,水动力条件改善效果最为明显,而水质改善最为明显的为竖向河道,这主要是由于引水口门河道、南横引河已经在常态化的调水中获得较稳定的水质;北横引河位于排水末端,水质改善效果相对差一点。

### 3 结 语

通过两次崇明岛河网调水试验,获得了原型实况调度的水文、水质同步监测第一手资料,掌握了群闸调度时各监测控制断面流量、水位以及岛内河道水质改善情况。两次调水试验为下一阶段建立崇明岛河网水动力水质模型、优化崇明岛调水方案提供了基础资料,也为开展崇明岛河道环境综合治理提供了参考依据。

a. 调水试验表明,开展水资源调度可以改善内河水质。但由于岛内尚分布有数量众多的横向小河道,两次调水试验均没有明显改善小河道的水动力条件和水质。建议加大对小河道的综合整治力度,以达到全面改善岛内河道水环境的目标。

b. 由于常年失修,部分竖向河道坍塌、淤积严重,建议加大崇明岛骨干河道的整治力度,提升骨干河道的过水能力,进一步改善水动力条件,以保障区域防洪除涝安全和改善河网水质。

c. 长江口水体含沙量较高,引水时泥沙也同步引入内河,加剧内河的淤积,从而影响河道的蓄排能力,建议下一阶段开展泥沙处理方式的专题研究。

d. 崇明岛现状污水处理能力较低,大部分污染源不经处理直排入河,建议根据崇明生态岛建设的要求,加大截污治污力度,加强研究,因地制宜地采取一些生态治理措施,对水体中的污染物进行吸收、降解和转化,增强水体自净能力。

### 参考文献:

- [1] 刘光文. 水文分析与计算[M]. 北京:水利电力出版社, 1989.
- [2] 茅志昌. 21 世纪初长江入海流量变化及其对长江口水质、生态环境的影响[J]. 海洋科学, 2001, 25(4): 32-34.
- [3] 朱慧峰, 阮仁良. 长江口水源地的开发利用态势及保护对策[J]. 中国给水排水, 2004, 20(4): 91-93.
- [4] 李伯昌, 施慧燕. 长江口河段水环境现状分析[J]. 水资源保护, 2005, 21(1): 39-44.
- [5] 余国安, 王兆印, 谢小平. 长江口水质空间分布现状评价[J]. 人民长江, 2007, 38(1): 81-83.

(收稿日期 2010-10-30 编辑 徐 娟)

(上接第 4 页)

- [13] PICHAYA R, ERIK R C. Determination of PCB sources by a principal component method with nonnegative constraints[J]. Environmental Science and Technology, 1997, 31(9): 2686-2691.
- [14] BAKAC M, KUMRU M N. Factor analysis applied to distribution of elements in western Turkey[J]. Applied Radiation and Isotopes, 2001, 55(5): 721-729.
- [15] PEKEY H, KARAKAS D, BAKOGLU M. Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses[J]. Marine Pollution Bulletin, 2004, 48(9/10): 809-818.

(上接第 18 页)

- [3] 王景雷, 吴景社, 孙景生, 等. 支持向量机在地下水位预报中的应用研究[J]. 水利学报, 2003(5): 122-127.
- [4] 王睿. 关于支持向量机参数选择方法分析[J]. 重庆师范大学学报:自然科学版, 2007, 24(2): 36-38.
- [5] 陈前平, 徐斌. 基于粒子群算法的支持向量机参数优化方法的研究[J]. 中国水运:理论版, 2006, 6(1): 97-99.
- [6] 雷英杰. MATLAB 遗传算法工具箱及应用[M]. 陕西:西安电子科技大学出版社, 2005: 78-90.
- [7] 周兆永, 汪西莉, 曹艳龙. 基于 GA 优选参数的 SVM 水质

- [16] SINGH K P, MALIK A, SINHA S. Water quality assessment and apportionment of pollution source of Gomti river (India) using multivariate statistical techniques: a case study[J]. Analytica Chimica Acta, 2005, 538(1/2): 355-374.
- [17] 郭芬, 张远. 水环境中 PAHs 源解析研究方法比较[J]. 环境监测管理与技术, 2008, 20(5): 11-16.
- [18] 夏青, 陈艳卿, 刘宪兵. 水质基准与水质标准[M]. 北京:中国标准出版社, 2004.
- [19] 苏金明. 统计软件 SPSS12.0 for Windows 应用及开发指南[M]. 北京:电子工业出版社, 2004.

(收稿日期 2010-06-13 编辑 高渭文)

评价方法研究[J]. 计算机工程与应用, 2008, 44(4): 190-193.

- [8] WANG W, XU Z, LU W. Determination of the spread parameter in the Gaussian kernel for classification and regression[J]. Neurocomputing, 2003, 55(6): 643-663.
- [9] 于青, 赵辉. 基于 GA 的支持向量机参数优化研究[J]. 计算机工程与应用, 2008, 44(15): 139-131.
- [10] 郑书彦, 李占斌, 李喜安. 地下水位动态预测的人工神经网络方法[J]. 西北水资源与水工程, 2002, 13(2): 14-16.

(收稿日期 2011-01-19 编辑 徐 娟)