

DOI:10.3969/j.issn.1004-6933.2013.04.005

支持向量机在湖库营养状态识别中的应用

崔东文

(云南省文山州水务局,云南 文山 663000)

摘要:依据我国湖库富营养化评价标准和支持向量机(SVM)原理及方法,构建基于交叉验证(CV)的 CV-SVM 湖库营养状态识别模型,采用随机内插的方法在各分级标准阈值间生成训练样本和测试样本,在达到预期识别精度后将模型运用于全国 24 个湖库营养状态的识别,并与投影寻踪法、评价指标法和神经网络评价法的识别结果进行比较。结果表明:基于线性核函数的 CV-SVM 模型对于随机生成的训练样本和测试样本的正确识别率分别达到 97.8% 和 97.3% (5 次平均),对全国 24 个湖库营养状态的识别结果与采用投影寻踪法、评价指标法和神经网络评价法的识别结果基本相同,模型具有泛化能力强、识别精度高、收敛速度快、不易陷入局部极值等特点。

关键词:湖库营养状态;识别模型;支持向量机;交叉验证

中图分类号:X524 文献标志码:A 文章编号:1004-6933(2013)04-0026-05

Application of support vector machine to lake and reservoir trophic status recognition

CUI Dongwen

(Wenshan Water Conservancy Bureau of Yunnan Province, Wenshan 663000, China)

Abstract: According to China's lake and reservoir eutrophication assessment standards and the support vector machine (SVM) theory and method, a CV-SVM lake and reservoir trophic status recognition model was constructed based on cross-validation (CV). With the interpolation method, the training samples and testing samples were randomly generated within the classification threshold. After the model's desired accuracy was achieved, it was applied to the recognition of trophic status of 24 lakes and reservoirs nationwide and compared with the projection pursuit, evaluation index, and neural network evaluation methods. The results are as follows: the recognition rate using the CV-SVM model based on a linear kernel function reached 97.8% and 97.3% (for five times on average) for the randomly generated training and testing samples, respectively. The trophic status recognition results of the 24 lakes and reservoirs were basically consistent with those obtained by the projection pursuit, evaluation index, and neural network evaluation methods. The model has the advantages of high generalization ability and recognition accuracy, fast convergence, and the unlikely occurrence of a local minimum.

Key words: lake and reservoir trophic status; recognition model; support vector machine; cross-validation

湖库营养状态的识别,就是通过与湖库营养状态有关的一系列指标及指标间的相互关系,对湖库的营养状态做出准确的判断^[1-2],识别过程具有高维、非线性特征,适宜借助诸如人工智能、模糊识别、知识工程等方法建立模型,以处理多指标系统的综合识别问题^[3-4]。模拟智能方法目前已成为建立和

识别这类复杂系统最为有效的途径之一,而人工神经网络(artificial neural network,ANN)则是这类智能算法中运用最为广泛的算法之一。ANN 具有较强的非线性映射能力、鲁棒性、容错性和自适应、自组织、自学习等许多特性,适宜解决高维、非线性系统问题,BP 模型(back-propagation network,BP)无疑是

作者简介:崔东文(1978—),男,高级工程师,主要从事水资源水环境研究及水资源保护等工作。E-mail:cdwgr@163.com

ANN 最为常用的神经网络模型之一,广泛应用于湖库营养状态研究中^[5-11]。支持向量机(support vector machine, SVM)是 20 世纪 90 年代中后期发展起来的基于统计学习理论构建的典型神经网络^[12],它由 Vapnik 首先提出,是一种通用的前馈神经网络,用于解决模式分类和非线性映射问题。SVM 具有严谨的数学基础,通过统计学习中的 VC 维(vapnik-chervonenkis dimension)理论和寻求结构风险最小化原理来提高泛化能力,有效解决了传统 BP 神经网络模型存在着学习收敛速度慢、易陷入局部极值以及网络结构难以确定等缺点,在湖库营养状态识别中得到了应用^[13]。笔者依据我国湖库富营养化评价标准和 SVM 原理及方法,构建基于交叉验证法(cross validation, CV)的 CV-SVM 湖库营养状态识别模型,采用随机内插的方法在各分级标准阈值间生成训练样本和测试样本,在达到预期的识别精度后将其运用于全国 24 个湖库营养状态识别中,并与投影寻踪法^[1]、评价指标法^[2]和 RBF 等 4 种神经网络评价法^[3]的评价结果进行比较,为湖库营养状态的识别提供新的途径和方法。

1 基于 SVM 湖库营养状态识别模型

SVM 最初是为研究线性问题提出的,其用于模式分类的基本思想是通过非线性变换将输入空间变换到一个高维空间,在此新空间中通过求解凸二次规划问题,寻求最优线性分类超平面,使它能够尽可能多地将两类数据点正确地分开,同时使分开的两类数据点距离分类面最远,而这种非线性变换是通过定义适当的内积(核)函数实现的。从 SVM 分类判别函数的形式上看,它类似于一个 3 层前馈神经网络,其隐层节点数对应于输入样本与一个支持向量机的内积核函数,输出节点数对应于隐层输出的线性组合^[14-17]。SVM 神经网络结构如图 1 所示。

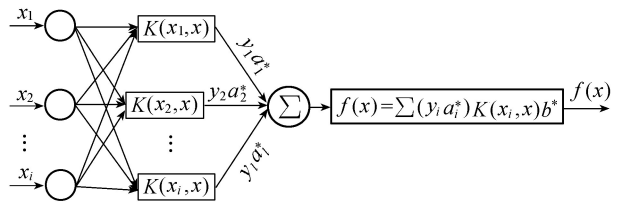


图 1 SVM 神经网络示意图

为不失一般性,设含有 l 个训练样本的集合 $\{(x_i, y_i), i=1, 2, \dots, l\}$ 由两个类别组成,若 x_i 属于第一类,则记为 $y_i=1$;若 x_i 属于第二类,则记为 $y_i=-1$,存在超平面 $w x+b=0$ 能够将样本正确分成两类,即相同类别的样本都落在分类超平面的同一侧,则称该样本集是线性可分的,满足:

$$\begin{cases} w x_i + b \geq 1 & y_i = 1 \\ w x_i + b \leq -1 & y_i = -1 \end{cases} \quad i = 1, 2, \dots, l \quad (1)$$

此时相应的两条极限直线之间的距离为 $2/\|w\|$,即在满足式(1)约束下寻求最优超平面,使得 $2/\|w\|$ 最大,即最小化 $\|w\|^2/2$,表示为:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s. t. } \begin{cases} y_i(w x_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \end{cases} \quad i = 1, 2, \dots, l \quad (2)$$

式中, C 为对分类错误的惩罚因子,用于调整置信区间和经验误差之间的均衡,它控制对错样本惩罚的程度,将式(2)折中考虑最少错分样本和最大分类间隔,就得到了广义最优分类面:

$$\begin{cases} \max(a) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i x_j) \\ \text{s. t. } \sum_{i=1}^l a_i y_i = 0 \end{cases} \quad a_i \geq 0 \quad (3)$$

求解式(3)可得到最优分类函数

$$f(x) = \text{sgn} \left[\sum_{i=1}^l a_i^* y_i (x_i x_j) + b^* \right] \quad (4)$$

式中, $f(x) = \text{sgn}(x)$ 为符号函数; a^* 为与每个样本对应的 Lagrange 乘子; b^* 是分类阈值,可以用任一支持向量在满足式(1)等号条件下求得。

然而,在实际应用中,绝大多数问题都是非线性的,这时对于线性可分 SVM 是无能为力的,对于此类线性不可分问题,是通过非线性映射将原输入空间的样本映射到高维的特征空间,再在高维特征空间构造最优分类超平面。VLAPIMIR 等^[12]提出采用满足 Mercer 条件的函数 $K(x_i, x_j)$ 来代替内积运算,此时 SVM 用于分类的最优分类函数为:

$$\begin{aligned} f(x) &= \text{sgn}(w^* \Phi(x) + b^*) = \\ &= \text{sgn} \left(\sum_{i=1}^l a_i^* y_i \Phi(x_i) \Phi(x) + b^* \right) = \\ &= \text{sgn} \left(\sum_{i=1}^l a_i^* y_i K(x_i, x) + b^* \right) \end{aligned} \quad (5)$$

常用的核函数有:

线性核函数 ($K(x, x_i) = x^T x_i$)、多项式核函数 ($K(x, x_i) = (\gamma x^T x_i + r)^p, \gamma > 0$)、径向基核函数 ($K(x, x_i) = \exp(-\gamma \|x - x_i\|^2), \gamma > 0$) 和两层感知核函数 ($K(x, x_i) = \tanh(\gamma x^T x_i + r)$)。

2 湖库营养状态模型的建立

2.1 评价指标及标准

依据 SL395—2007《地表水资源质量评价技术

表 1 我国湖库富营养化评价标准

富营养化程度	富营养化等级	评价指标				
		$\rho(\text{Chl-a})/(\text{mg}\cdot\text{L}^{-1})$	$\rho(\text{TP})/(\text{mg}\cdot\text{L}^{-1})$	$\rho(\text{TN})/(\text{mg}\cdot\text{L}^{-1})$	$\rho(\text{COD}_{\text{Mn}})/(\text{mg}\cdot\text{L}^{-1})$	$h(\text{SD})/\text{m}$
贫营养	1	≤ 1.0	≤ 2.5	≤ 30.0	≤ 0.3	≥ 10.0
贫中营养	2	(1.0,2.0]	(2.5,5.0]	(30.0,50.0]	(0.3,0.4]	[5.0,10.0)
中营养	3	(2.0,4.0]	(5.0,25.0]	(50.0,300.0]	(0.4,2.0]	[1.5,5.0)
中富营养	4	(4.0,10.0]	(25.0,50.0]	(300.0,500.0]	(2.0,4.0]	[1.0,1.5)
富营养	5	(10.0,64.0]	(50.0,200.0]	(500.0,2000.0]	(4.0,10.0]	[0.4,1.0)
重富营养	6	>64.0	>200.0	>2000.0	>10.0	<0.4
	极大值	128	400	4000	20	20
	极小值	0.5	1.25	15	0.15	0.2

规程》湖库营养状态评价标准^[18],选取 Chl-a、TP、TN、COD_{Mn}和透明度(SD)作为湖库营养状态评价指标,见表 1,将湖库营养程度按由轻到重分别记为 1、2、3、4、5、6 级^[1],选取重富营养临界值的 2 倍作为评价指标极大值(上限值),贫营养临界值的 0.5 倍作为评价指标极小值(下限值),并以其上下限值作为评价对象评价指标的极点值。

2.2 营养状态识别的实现

2.2.1 数据的标准化处理

表 1 中湖库富营养化程度评价指标分为正向指标和负向指标,为了消除不同量纲对评价结果的影响,首先需对评价指标数据进行标准化处理。对湖库富营养化程度评价等级起越小越优作用的指标,如 Chl-a、TP 等,其处理方法为:

$$\hat{x} = (x - x_{\min}) / (x_{\max} - x_{\min}) \tag{6}$$

对湖库富营养化程度评价等级起越大越优作用的指标,如水体透明度,其处理方法为:

$$\hat{x} = 1 - (x - x_{\min}) / (x_{\max} - x_{\min}) \tag{7}$$

式中: $\hat{x}$ 为经过标准化处理的数据, x 为原始数据, $x_{\max}$ 、 $x_{\min}$ 为数据序列中的最大、最小数。经过标准化处理后,数据处于[0~1]范围之内,有利于网络训练。

2.2.2 训练及测试样本设计

依据表 1,为不失一般性,采用随机内插的方法在各分级阈值间生成 30 个样本,随机选取 20 个作为训练样本,10 个作为测试样本,以此计算共随机内插得到 180 个样本,其中 120 个作为训练样本,60 个作为测试样本。

2.2.3 SVM 训练、测试及评价

SVM 用于处理模式分类或非线性映射问题时,在选定核函数条件下,模型中的惩罚因子 C 和核函数参数 g 的选取对模型的识别精度有着关键性影响^[14]。由于核函数、惩罚因子和核函数参数的选取目前尚无理论上的指导原则,最优参数的选取多凭经验、实验对比等进行搜寻,极大地制约了 SVM 模型精度^[16-17]。目前普遍采用 CV 法来搜寻合理参数。CV 法是用来验证分类器性能一种统计分析方法,基本思想是把在某种意义下将原始数据

(dataset)进行分组,一部分作为训练集(train set),另一部分作为验证集(validation set),首先用训练集对分类器进行训练,再利用验证集来测试训练得到的模型(model),以此来做为评价分类器的性能指标。常见的 CV 法有 hold-out method、K-fold cross validation (K-CV) 和 leave-one-out cross validation (LOO-CV),CV 法可以有效地避免模型“过学习”以及“欠学习”现象的发生。

本文基于 MATLAB 环境,运用 libsvm 工具箱编写相关算法程序,采用 CV 算法作为惩罚因子 C 和核函数参数 g 的寻优方法,在保证模型其他参数不变的条件下,分别采用基于不同核函数的 SVM 对训练样本和测试样本进行识别,以正确识别率作为模型的评价指标。由于训练样本和测试样本是随机内插生成,因此模型每次运行结果均不一样,笔者摘录某次随机运行的 5 次结果(表 2),以正确识别率的平均值来评价网络性能的优劣。

表 2 基于不同核函数的 SVM 正确识别率对比结果

样本类型	不同核函数识别正确率比较/%			
	线性	多项式	RBF	Sigmoid
训练样本	98.3(118/120)	80.0(96/120)	100(120/120)	21.7(26/120)
	100(120/120)	86.7(104/120)	100(120/120)	16.7(20/120)
	96.6(116/120)	88.3(106/120)	100(120/120)	16.7(20/120)
	96.6(116/120)	89.2(107/120)	100(120/120)	17.5(21/120)
	97.5(117/120)	84.2(101/120)	100(120/120)	14.2(17/120)
平均识别率	97.8	87.1	100	17.3

样本类型	不同核函数识别正确率比较/%			
	线性	多项式	RBF	Sigmoid
测试样本	98.3(59/60)	78.3(47/60)	90.0(54/60)	23.3(14/60)
	95.0(57/60)	81.7(49/60)	93.3(56/60)	16.7(10/60)
	96.6(58/60)	86.7(52/60)	98.3(59/60)	16.7(10/60)
	98.3(59/60)	91.7(55/60)	95.0(57/60)	18.3(11/60)
	98.3(59/60)	88.3(53/60)	96.7(58/60)	15.0(9/60)
平均识别率	97.3	85.7	94.7	18.0

注:括号中的数据表示训练或检测样本正确识别的个数除以训练或检验样本总数。

从表 2 可以看出,线性核函数和 RBF 核函数对应的训练样本和测试样本的正确识别率较高,而 Sigmoid 核函数对应的识别率最低,多项式核函数处

于中间水平。从模型的泛化能力考虑,即衡量测试样本的正确识别率,则线性核函数对应的模型性能最佳,因此本文采用线性核函数进行建模,用于对全国 24 个湖库营养状态的识别,交叉验证优化后得到的最优参数为: $C=0.435\ 28$, $g=18.379\ 2$ 。

3 应用实例

3.1 数据来源

选取我国主要湖库的调查资料进行实例分析^[1](表 3)。

表 3 我国主要湖库调查资料

湖 库(省、市、区)	$\rho(\text{Chl-a})/(\text{mg}\cdot\text{L}^{-1})$	$\rho(\text{TP})/(\text{mg}\cdot\text{L}^{-1})$	$\rho(\text{TN})/(\text{mg}\cdot\text{L}^{-1})$	$\rho(\text{COD}_{\text{Mn}})/h(\text{SD})/(\text{mg}\cdot\text{L}^{-1})$	m
洱 海(云南)	1.86	22	246	3.09	2.77
高州水库(广东)	1.49	46	358	1.47	1.72
博斯腾湖(新疆)	3.52	23	932	5.96	1.46
淀山湖(上海)	3.00	29	1086	2.87	0.67
于桥水库(天津)	10.79	25	1220	4.11	1.42
固城湖(江苏)	4.99	52	2374	2.75	0.28
南四湖(山东)	3.77	194	3201	6.96	0.44
磁 湖(湖北)	14.47	77	1000	3.74	0.36
达理湖(内蒙古)	7.24	153	1671	16.25	0.48
巢 湖(安徽)	11.80	115	1786	4.01	0.28
滇池外海(云南)	44.43	108	1309	7.11	0.49
滇池草海(云南)	298.86	931	15273	16.58	0.23
西 湖(浙江)	58.95	161	2478	6.94	0.43
甘棠湖(江西)	75.69	141	1417	7.23	0.38
蘑菇湖(新疆)	54.77	287	2206	10.38	0.53
麓 湖(广东)	119.51	372	3038	9.92	0.34
东山湖(广东)	149.45	428	5350	13.40	0.22
墨水湖(湖北)	153.59	232	15692	13.51	0.22
荔湾湖(广东)	162.92	743	7337	14.46	0.31
流花湖(广东)	323.51	643	6777	25.26	0.15
玄武湖(江苏)	168.14	663	4073	10.08	0.22
镜泊湖(吉林)	4.96	316	1270	5.96	0.73
南 湖(吉林)	120.60	228	2630	8.22	0.22
邛 海(四川)	0.88	130	410	1.43	2.98

3.2 评价结果及分析

依据表 1,利用上述训练及测试好的 CV-SVM 模型对表 3 中 24 个湖库营养状态进行识别,识别结果与文献[1-3]进行比较,见表 4。

分析表 3、表 4 可以得出以下结论:
CV-SVM 模型识别结果与文献[1-3]的识别(评价)结果基本相同,与投影寻踪法^[1]和评分指数法^[2]的评价结果较为接近,仅有 3 湖(库)之别,其中与投影寻踪法比较,有博斯腾湖、于桥水库和蘑菇湖偏小 1 个等级;与评分指数法比较,有博斯腾湖、固城湖和南四湖偏大 1 个等级。与文献[3]的 BP 神经网络法相比,有 5 湖(库)之别,均偏小 1 个等级;与文献[3]的 RBF、GRNN 和 Elman 神经网络法相比,有 6 湖(库)之别,均偏小 1 个等级。从与文献[1-3]的比较结果可以得出,CV-SVM 湖库营养状

态识别模型泛化能力强,识别精度高,表明研究建立的 CV-SVM 湖库营养状态识别模型和识别方法是合理可行的,可为湖库营养状态的识别提供新的途径和方法。

表 4 我国主要湖库营养状态识别结果

湖库(省、市、区)	识别结果比较						
	CV-SVM 法	投影寻踪法 ^[1]	评价指标法 ^[2]	RBF 网络法 ^[3]	GRNN 网络法 ^[3]	BP 网络法 ^[3]	Elman 网络法 ^[3]
洱 海(云南)	3	3	3	4	4	4	4
高州水库(广东)	3	3	3	4	4	4	4
博斯腾湖(新疆)	4	5	3	5	5	5	5
淀山湖(上海)	4	4	4	5	5	4	5
于桥水库(天津)	4	5	4	5	5	5	5
固城湖(江苏)	5	5	4	5	5	5	5
南四湖(山东)	5	5	4	5	5	5	6
磁 湖(湖北)	5	5	5	5	5	5	5
达理湖(内蒙古)	5	5	5	5	5	5	6
巢 湖(安徽)	5	5	5	5	5	5	5
滇池外海(云南)	5	5	5	5	5	5	5
滇池草海(云南)	6	6	6	6	6	6	6
西 湖(浙江)	5	5	5	5	5	5	5
甘棠湖(江西)	5	5	5	5	5	5	5
蘑菇湖(新疆)	5	6	5	6	6	6	6
麓 湖(广东)	6	6	6	6	6	6	6
东山湖(广东)	6	6	6	6	6	6	6
墨水湖(湖北)	6	6	6	6	6	6	6
荔湾湖(广东)	6	6	6	6	6	6	6
流花湖(广东)	6	6	6	6	6	6	6
玄武湖(江苏)	6	6	6	6	6	6	6
镜泊湖(吉林)	5	5	5	5	5	5	5
南 湖(吉林)	6	6	6	6	6	6	6
邛 海(四川)	4	4	3	4	4	4	3

4 结 语

SVM 是建立在统计学习 VC 维理论和结构风险最小化原则上的学习机,可避免局部极小值和“过学习”等问题,其构造分类函数的复杂程度取决于支持向量的个数,与特征空间的维数无关,能有效解决可能导致的“维数灾”问题,具有良好的推广性能和实用价值。从湖库营养状态识别结果可以看出,本文所构建的基于线性核函数的 CV-SVM 模型具有良好的泛化能力强和识别精度,其在解决小样本、非线性及高维系统问题时表现出更大的优势。当然,SVM 核函数、惩罚因子和核函数参数的合理选取是制约 SVM 进一步提高精度的关键性因素,采用遗传算法、粒子群算法以及针对遗传算法可能存在着早熟收敛、易陷入局部极值等缺点提出的基于改进的多种群遗传算法、量子遗传算法和免疫遗传算法^[17,19]对 SVM 核参数和误差惩罚因子进行优化,使 SVM 的识别精度得到更进一步的提高。

参考文献:

[1] 王贵作,任立良,王斌,等. 基于投影寻踪的湖泊富营养化程度评价模型[J]. 水资源保护,2009,25(5):14-18. (WANG Guanzuo, REN Liliang, WANG Bin, et al. Based on projection pursuit lake eutrophication evaluation model[J]. Water Resources Protection, 2009,25 (5):14-18. (in Chinese))

[2] 舒金华. 我国主要湖泊富营养化程度的评价[J]. 海洋与湖沼, 1993, 24 (6): 616-620. (SHU Jinhua. Assessment of eutrophication in main lakes of China[J]. Oceanologia Et Limnologia Sinica, 1993, 24 (6): 616-620. (in Chinese))

[3] 崔东文. 几种神经网络模型在湖库富营养化程度评价中的应用[J]. 水资源保护,2012,25(5):14-18. (CUI Dongwen. Applications of several neural network models to eutrophication[J]. Water Resources Protection, 2012, 25 (5):14-18. (in Chinese))

[4] 金菊良,王文圣,洪天求,等. 流域水安全智能评价方法的理论基础探讨[J]. 水利学报,2006,37(8):918-925. (JIN Juliang, WANG Wensheng, HONG Tianqiu, et al. Theoretical basis of intelligent evaluation methods of watershed water security [J]. Journal of Hydraulic Engineering,2006,37(8):918-925. (in Chinese))

[5] 邓大鹏,刘刚,李学德,等. 基于神经网络简单集成的湖库富营养化综合评价模型[J]. 生态学报,2007,27(2):725-731. (DENG Dapeng, LIU Gang, LI Xuede, et al. A model based on simple ensemble of neural networks for comprehensive eutrophication assessment of lake and reservoir[J]. Acta Ecologica Sinica, 2007,27(2):725-731. (in Chinese))

[6] 陈安,罗亚田. 人工神经网络方法在环境科学领域应用进展[J]. 重庆环境科学,2003,25(9):65-70. (CHEN An, LUO Yatian. Advances in artificial neural network method in the application of environmental science[J]. Chongqing Environmental Science, 2003, 25 (9): 65-70. (in Chinese))

[7] 任黎,董增川,李少华. 人工神经网络模型在太湖富营养化评价中的应用[J]. 河海大学学报:自然科学版, 2004,32(2):147-150. (REN Li, DONG Zengchuan, LI Shaohua. Application of artificial neural network model to assessment of Taihu Lake eutrophication[J]. Journal of Hohai University: Natural Sciences, 2004, 32 (2): 147-150. (in Chinese))

[8] 韩涛,李怀恩,彭文启. 基于 MATLAB 的神经网络在湖泊富营养化评价中的应用[J]. 水资源保护,2005,21(1):24-26. (HAN Tao, LI Huaen, PENG Wenqi. Application of neural network based on MATLAB toolbox to the evaluation of lake eutrophication [J]. Water Resources Protection,2005,21(1):24-26. (in Chinese))

[9] 吴利斌,尚士友,岳海军,等. 利用模糊神经网络对湖泊富营养化程度进行评价的研究[J]. 内蒙古农业大学学报,2004,25(4):67-70. (WU Libin, SHANG Shiyu, YUE Haijun, et al. Study on evaluation of eutrophication of lakes based on fuzzy neural network [J]. Journal of Inner Mongolia Institute of Agriculture and Animal Husbandry,2004,25(4):67-70. (in Chinese))

[10] 曾光明,卢宏伟,金相灿,等. 洞庭湖水体水质状况及运用小波神经网络对营养状态的评价[J]. 湖南大学学报:自然科学版, 2005, 32 (1): 91-94. (ZENG Guangming, LU Hongwei, JIN Xiangcan, et al. Assessment of the water quality and nutrition of the Dongting Lake with wavelet neural network [J]. Journal of Hunan University: Natural Science, 2005, 32 (1): 91-94. (in Chinese))

[11] 楼文高. 湖库富营养化人工神经网络评价模型[J]. 水产学报, 2001, 25 (5): 474-478. (LOU Wengao. Eutrophication assessment model using artificial neural networks for lakes and reservoirs[J]. Journal of Fisheries of China, 2001,25(5):474-478. (in Chinese))

[12] VLADIMIR N. 统计学习理论的本质[M]. 张学工,译. 北京:清华大学出版社,2000.

[13] 李正最,谢悦波. 洞庭湖富营养化支持向量机评价模型研究 [J]. 人民长江, 2010, 41 (10): 75-78. (LI Zhengzui, XIE Yuebo. Study on eutrophication evaluation model of Dongting Lake based on SVM [J]. Yangtze River,2010,41(10):75-78. (in Chinese))

[14] 田景文,高美娟. 人工神经网络算法研究及应用[M]. 北京:北京理工大学出版社,2006.

[15] 施彦,韩力群,廉小亲. 神经网络设计方法与实例分析 [M]. 北京:北京邮电大学出版社,2009.

[16] MATLAB 中文论坛. MATLAB 神经网络 30 个案例分析 [M]. 北京:北京航空航天大学出版社,2010.

[17] 史峰,王辉,郁磊,等. MATLAB 智能算法 30 个案例分析 [M]. 北京:北京航空航天大学出版社,2011.

[18] SL395—2007 地表水资源质量评价技术规程[S].

[19] 雷德明,严新平. 多目标智能优化算法及其应用 [M]. 北京:科学出版社,2009.

(收稿日期:2012 - 08 - 24 编辑:高渭文)