

# $\nu$ -支持向量机洪水预报模型研究

但灵芝<sup>1</sup>, 王建群<sup>1</sup>, 陈理想<sup>2</sup>, 陈红红<sup>1</sup>

(1. 河海大学水文水资源学院, 江苏 南京 210098; 2. 松辽水利委员会水文局, 吉林 长春 130021)

**摘要:** 为了提高洪水过程的非线性拟合能力和预报精度, 对支持向量机洪水预报模型进行了深入研究。针对鸭绿江流域临江站的实际情况, 建立了临江站洪水过程  $\nu$ -SVR 预报模型, 采用 1998~2014 年间的大水年份降水资料和洪水过程资料对  $\nu$ -SVR 预报模型进行了率定和验证, 并与线性动态系统模型、BP 神经网络模型和  $\varepsilon$ -SVR 模型进行了比较。结果表明:  $\nu$ -SVR 洪水预报模型比线性动态系统模型和 BP 神经网络具有较高的精度。 $\nu$ -SVR 洪水预报模型具有较好的非线性拟合能力和泛化能力, 能很好地控制支持向量个数、降低模型的复杂程度, 同时能保持良好的预报精度。

**关键词:** 洪水预报; 支持向量机; 线性动态系统; BP 神经网络

中图分类号: P338 文献标识码: A 文章编号: 1000-0852(2016)02-0007-05

## 1 引言

由于洪水过程是高度非线性的动态过程, 许多学者将神经网络方法用于洪水预报获得了良好的效果<sup>[1-3]</sup>。BP 神经网络具有很好的非线性逼近能力, 但还存在过学习、模型计算效率低、泛化能力差等固有的缺点, 在训练样本洪水代表性不足的条件下, 模型的外延能力有限<sup>[3]</sup>。支持向量机是 Vanpik 等人根据统计学习理论的 VC 维理论和结构风险最小原理提出的一种机器学习方法, 有效地解决了小样本、过学习、高维和局部极值等问题, 具有全局最优、良好的泛化能力、非线性处理等优越性能<sup>[4]</sup>。支持向量机方法已成功应用于年径流或月径流预测, 例如文献[5-10]等; 但用于洪水过程预报的研究较少, 而洪水过程的非线性特征更为显著。理论上, 基本的  $\varepsilon$ -支持向量机回归 ( $\varepsilon$ -SVR) 可以用于洪水过程预报, 但需要事先确定损失函数中的参数  $\varepsilon$  和平均损失常数  $C$ , 采用优化方法确定  $\varepsilon$  和  $C$  常常导致支持向量的个数太大而不实用。Scholkopf(2000)<sup>[11]</sup>提出了一种新的  $\nu$ -支持向量机回归 ( $\nu$ -SVR)。 $\nu$ -支持向量机回归能自动计算损失函数中的参数  $\varepsilon$  并能通过  $\nu$  的选择控制支持向量的个数。本文将对洪水过程  $\nu$ -支持向量机预报模型进行深入研

究, 研究建立鸭绿江上游临江站洪水过程  $\nu$ -支持向量机预报模型, 并通过实例与线性动态系统模型、BP 神经网络及  $\varepsilon$ -支持向量机模型进行比较。

## 2 $\nu$ -支持向量机回归模型

设样本集为  $(y_1, x_1), \dots, (y_l, x_l), x \in R^n, y \in R$  支持向量机回归的基本思想是<sup>[11]</sup>: 对于非线性样本, 通过非线性变换将原问题映射到某个高维特征空间中的线性问题上进行求解:

$$f(x) = w^T \phi(x) + b \quad (1)$$

式中:  $\phi(x)$  为  $R^n$  到高维特征空间的非线性映射向量;  $w$  为权值向量;  $b$  为阈值。传统的支持向量机回归模型  $\varepsilon$ -SVR 通过求解最优化问题(2)建立:

$$\begin{aligned} \min_{w, \zeta_i, \zeta_i^*, b} & \frac{1}{2} w^T w + C \frac{1}{l} \sum_{i=1}^l (\zeta_i + \zeta_i^*) \\ \text{s.t.} & (w^T \phi(x_i) + b) - y_i \leq \varepsilon + \zeta_i \\ & y_i - (w^T \phi(x_i) + b) \leq \varepsilon + \zeta_i^* \\ & \zeta_i, \zeta_i^* \geq 0 \end{aligned} \quad (2)$$

问题(2)的对偶形式为:

$$\max_{\alpha, \alpha^*} \sum_{i=1}^l \alpha_i (y_i - \varepsilon) - \alpha_i^* (y_i + \varepsilon) - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(x_i, x_j)$$

收稿日期: 2015-03-13

基金项目: 国家自然科学基金项目(41371047)

作者简介: 但灵芝 (1992-), 女, 江西九江人, 硕士研究生, 从事水文水资源研究。E-mail: danlzhi@163.com

通讯作者: 王建群 (1960-), 男, 江苏句容人, 博士, 教授, 从事水资源系统分析研究。E-mail: wangjq@hhu.edu.cn

$$\text{s.t.} \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad (3)$$

$$0 \leq \alpha_i, \alpha_i^* \leq C, \quad i=1, \dots, l$$

式中:  $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$  为核函数。

求解二次规划问题(3)可得拉格朗日待定系数  $\alpha_i$  和  $\alpha_i^*$ , 于是建立  $\varepsilon$ -SVR 回归函数  $f(x)$  为:

$$f(x) = \sum_{SVs} (\bar{\alpha}_i - \bar{\alpha}_i^*) K(x_i, x) + b \quad (4)$$

当采用径向基核函数时, 式中  $K(x_i, x) = \exp(-|x - x_i|^2 / (2\sigma^2))$ ;  $x_i$  为样本点;  $\sigma$  为径向基核函数参数。

在  $\varepsilon$ -SVR 中, 需要事先确定  $\varepsilon$ -不敏感损失函数中的参数  $\varepsilon$ 。本文采用能自动计算损失函数中的参数  $\varepsilon$  并能通过  $\nu$  的选择控制支持向量个数的  $\nu$ -SVR<sup>[11]</sup>。 $\nu$ -SVR 将  $\varepsilon$  作为决策变量, 求解最优化问题(5)建立:

$$\begin{aligned} \min_{w, b, \zeta, \varepsilon} & \quad \frac{1}{2} w^T w + C \left( v\varepsilon + \frac{1}{l} \sum_{i=1}^l (\zeta_i + \zeta_i^*) \right) \\ \text{s.t.} & \quad (w^T \phi(x_i) + b) - y_i \leq \varepsilon + \zeta_i, \\ & \quad y_i - (w^T \phi(x_i) + b) \leq \varepsilon + \zeta_i^*, \\ & \quad \zeta_i, \zeta_i^* > 0, i=1, \dots, l, \varepsilon > 0 \end{aligned} \quad (5)$$

式(5)的对偶形式:

$$\begin{aligned} \max_{\alpha, \alpha^*} & \quad \sum_{i=1}^l \alpha_i (y_i - \varepsilon) - \alpha_i^* (y_i + \varepsilon) - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(x_i, x_j) \\ \text{s.t.} & \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ & \quad \sum_{i=1}^l (\alpha_i + \alpha_i^*) \leq C \cdot v \\ & \quad 0 \leq \alpha_i, \alpha_i^* \leq C/l, \quad i=1, \dots, l \end{aligned} \quad (6)$$

求解二次规划问题(6)可得  $\nu$ -SVR。

可以证明<sup>[11]</sup>: 给定参数  $C$ , 如果  $\nu$ -SVR 的解为  $(\bar{\varepsilon}, \bar{w}, \bar{b})$ , 而  $\varepsilon$ -SVR 事先取  $\varepsilon = \bar{\varepsilon}$  并与  $\nu$ -SVR 取相同的  $C$ , 那么  $\varepsilon$ -SVR 得到的解为  $(\bar{w}, \bar{b})$ 。 $\nu$ -SVR 的性能取决于参数  $C, \nu$ 。惩罚因子  $C$  是模型复杂度和推广能力的折中, 用于调节样本空间中置信范围和经验风险的比例, 决定了对损失大于  $\varepsilon$  的样本的惩罚程度。参数  $\nu$  表示支持向量的下界与间隙误差的上界。不敏感损失参数  $\varepsilon$  控制着不敏感带的宽度, 影响着支持向量的数目,  $\varepsilon$  值越大, 支持向量数目减少, 回归估计的精度越低。由于  $\nu$  用于优化参数  $\varepsilon$ , 因此  $\nu$  也间接影响着支持向量的数目和预测的精度,  $\nu$  的减少导致支持向量的减少; 参数  $\nu$  的有效取值范围为  $\nu \in [0, 1]$ <sup>[11]</sup>。

### 3 实例分析

考虑鸭绿江上游云峰水库入库控制站临江站洪水过程预报问题。以临江站上游十三道湾站的流量过程为入流、考虑十三道站湾至临江站区间降雨径流影响, 预报临江站洪水过程。十三道站湾至临江站区间雨量站稀少, 难以建立新安江模型等概念性水文模型进行降雨径流预报。本文采用  $\nu$ -SVR 洪水过程预报模型方法进行临江站洪水过程预报。

以十三道湾站流量过程、临江站前期流量过程、十三道湾站至临江站区间降雨及前期影响雨量作为输入, 临江站流量过程作为输出, 建立如下系统响应模型:

$$Q_1(j) = f(Q_1(j - mp_{10} - 1), \dots, Q_1(j - mp_{10} - mp_1), Q_2(j - mp_{20} - 1), \dots, Q_2(j - mp_{20} - mp_2), P_w(j - mp_{30} - 1) \dots, P_w(j - mp_{30} - mp_3), P_a(j - mp_{40} - 1) \dots, P_a(j - mp_{40} - mp_4)) \quad (7)$$

式中:  $Q_1(j), j=1, 2, \dots$  为临江站流量过程;  $Q_2(j), j=1, 2, \dots$  为十三道湾站流量过程;  $P_w(j), j=1, 2, \dots$  为十三道湾站至临江站区间面平均雨量过程;  $P_a(j), j=1, 2, \dots$  为区间前期影响雨量过程;  $mp_{10}=3, mp_1=4; mp_{20}=3, mp_2=4; mp_{30}=0, mp_3=8; mp_{40}=0, mp_4=1$ , 时段长为 3h, 预见期为  $\tau = (1+3) \times 3 = 12\text{h}$ 。十三道湾站至临江站区间面平均雨量采用临江、八道沟、十三道湾、柳树河、闹枝沟 5 站算术平均计算。前期影响雨量采用以下公式计算:

$$P_a = k p_1 + k^2 p_2 + \dots + k^n p_n \quad (8)$$

$$\text{若 } P_a > I_m, \text{ 则取 } P_a = I_m \quad (9)$$

式中:  $P_t, t=1, 2, \dots, n$  为前  $t$  日日降水量, 取  $n=15$ ;  $k$  为日折减系数;  $I_m$  为最大初损值; 经分析, 取  $I_m=100\text{mm}$ ,  $k=0.88$ 。

对于回归模型(7), 可以采用混合线性回归模型<sup>[12-13]</sup>逼近、BP 神经网络模型逼近<sup>[3]</sup>或支持向量机回归模型逼近<sup>[4]</sup>。本文采用  $\nu$ -SVR 逼近回归模型(7), 建模步骤如下:

(1) 样本归一化处理。为避免各个因子之间的量级差异, 消除量纲和单位不同对各个因子的影响, 需对样本进行归一化处理, 将输入和输出统一归一化到  $[0, 1]$  内:

$$y^* = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \quad (10)$$

式中:  $y^*$  为标准化处理后的数据;  $y$  为原始数据;  $y_{\max}$  为样本系列最大值;  $y_{\min}$  为样本系列最小值。

(2) 确定核函数。选定径向基核函数, 核函数所含的参数为  $\sigma$ 。

(3) 参数选择。给定惩罚因子  $C$ 、核函数参数  $\sigma$  及

支持向量控制参数  $\nu$ , 求解二次规划问题(6)可得  $\nu$ -SVR 模型, 归一化后的实测值与预报值的均方误差为:

$$MSE = \frac{1}{l} \sum_{i=1}^l (f(x_i) - y_i)^2 \quad (11)$$

通过极小化均方误差, 可得求得  $\nu$ -SVR 模型参数  $C, \sigma, \nu$ 。通过极小化均方误差求解的  $\nu$ -SVR 模型, 其支持向量的个数可能很大。固定参数  $C, \sigma$ , 通过摄动参数  $\nu \in [0, 1]$ , 分析预报误差及支持向量机个数, 最终确定参数  $\nu$ 。

对于临江站洪水过程预报模型(7), 本文采用1998~2014年的大水年份即2004、2005、2007、2010、2012、2013年汛期洪水过程资料和相应的降水资料进行模型率定和验证, 资料情况见表1。

表1 模型率定和验证资料说明

Table1 The data of the calibration and validation of the model

资料年份	起止时间	备注
2004	6-19 8:00-9-18 8:00	率定
2005	5-8 8:00-9-19 8:00	率定
2007	7-30 8:00-9-30 8:00	率定
2010	7-17 8:00-9-30 8:00	率定
2012	5-1 8:00-9-30 8:00	验证
2013	5-1 8:00-9-30 8:00	验证

本文基于 libsvm 工具箱<sup>[14]</sup> 开源软件编写  $\nu$ -SVR 求解程序, 取  $\nu=1$ , 采用群体复合型进化算法<sup>[15]</sup>进行参数  $C, \sigma$  优选, 得最优参数为  $[C, \sigma] = [16.1255, 0.0753]$ 。摄动参数  $\nu$ , 求得  $\nu$ -SVR, 统计误差见表2。

表2  $\nu$ -SVR模型参数 $\nu$ 的摄动分析

Table2 The perturbation analysis of the parameter  $\nu$

序号	$\nu$	MSE	确定性系数	洪峰相对误差/%	支持向量数目
1	1	0.000226	0.98	6.37	2737
2	0.54	0.000228	0.97	6.60	1610
3	0.29	0.000232	0.97	6.89	879
4	0.19	0.000233	0.97	6.67	584
5	0.14	0.000232	0.97	7.39	439
6	0.11	0.000233	0.97	8.78	345
7	0.09	0.000241	0.97	8.43	284
8	0.07	0.000244	0.97	9.35	221
9	0.06	0.000258	0.97	9.91	194
10	0.04	0.000397	0.96	12.41	138

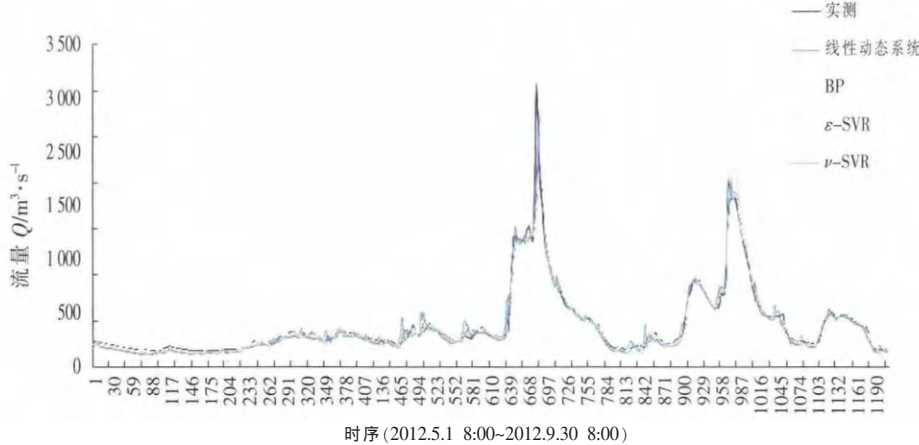


图1 临江站洪水过程(2012.5.1 8:00~2012.9.30 8:00,  $\Delta t=3h$ )

Fig.1 The flood hydrograph of the Linjiang station (2012.5.1 8:00~2012.9.30 8:00,  $\Delta t=3h$ )

由表2可以看出, 参数  $\nu$  减小, 支持向量机数目随之减少, 预测精度也降低; 参数  $\nu$  增加, 预测精度提高, 但支持向量数目也增加。当  $\nu=0.06$  时, 模型拟合精度可以接受, 支持向量的数目较少, 模型相对简单。最终确定  $\nu=0.06$ , 相应的支持向量数目为194。

将  $\nu$ -SVR 模型与线性动态系统模型、 $\epsilon$ -SVR 模型、BP 神经网络对比, 分析  $\nu$ -SVR 的性能。线性动态系统模型采用最小

表3 确定性系数及相对误差统计

Table3 The statistics of the deterministic coefficient and relative errors

年份	线性动态系统		BP神经网络		$\epsilon$ -SVR		$\nu$ -SVR		
	确定性系数	洪峰相对误差/%	确定性系数	洪峰相对误差/%	确定性系数	洪峰相对误差/%	确定性系数	洪峰相对误差/%	
率定期	2004	0.92	9.52	0.98	-0.35	0.98	-0.82	0.98	-2.14
	2005	0.92	-7.97	0.98	-6.23	0.96	-9.93	0.96	-8.56
	2007	0.95	-15.08	0.99	-6.86	0.98	-2.08	0.98	-10.49
	2010	0.92	9.13	0.98	0.66	0.99	0.93	0.98	1.87
验证期	2012	0.94	-10.49	0.96	-15.20	0.98	-8.90	0.97	-9.70
	2013	0.93	8.69	0.95	9.66	0.96	-3.26	0.96	-6.48

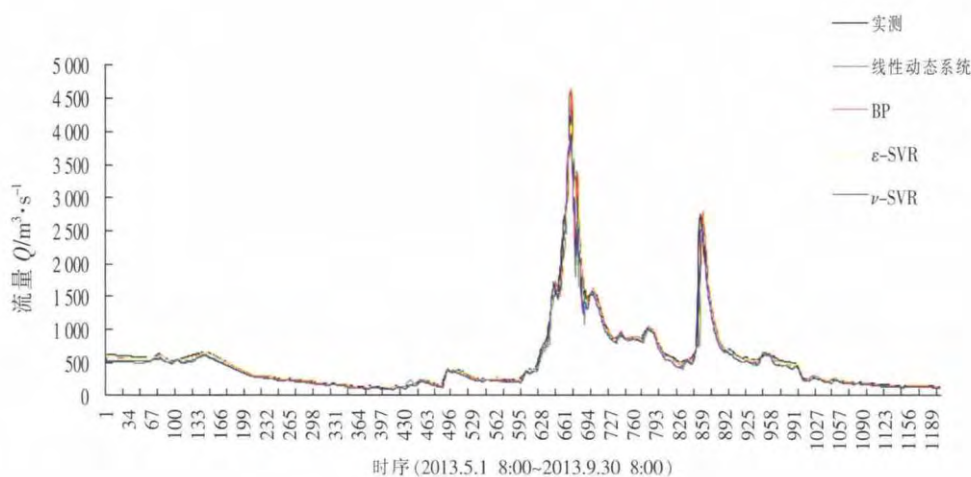


图2 临江站洪水过程(2013.5.1 8:00~2013.9.30 8:00,  $\Delta t=3h$ )

Fig.2 The flood hydrograph of the Linjiang station (2013.5.1 8:00~2013.9.30 8:00,  $\Delta t=3h$ )

二乘法求解;  $\epsilon$ -SVR 模型选取径向基核函数, 采用群体复合型进化算法优化参数  $C, \sigma, \epsilon$ ; BP 神经网络采用 3 层结构网络并采用 MATLAB 求解。4 种方法计算的洪水过程确定性系数及洪峰相对误差统计见表 3, 验证期洪水过程线拟合效果见图 1~图 2。其中,  $\epsilon$ -SVR 最优参数为  $[C, \sigma, \epsilon] = [18.1255, 0.07534, 0.01]$ , 支持向量个数为 447。

从表 3、图 1、图 2 可以看出:

(1) 4 种模型中, 无论是率定期还是验证期, 线性动态系统模型的确定性系数最低、洪峰相对误差较大, 且从过程线看, 线性动态系统模型的计算结果表现出不合理的跳动。

(2) 对于率定期, BP 神经网络模型的确定性系数及洪峰相对误差均好于线性动态系统模型、 $\epsilon$ -SVR 模型以及  $\nu$ -SVR 模型; 但对于验证期, BP 神经网络模型效果较差。这也印证了 BP 神经网络模型存在过学习、泛化能力差的固有缺陷。

(3) 无论是率定期还是验证期,  $\epsilon$ -SVR 模型的确定性系数及洪峰相对误差均好于  $\nu$ -SVR 模型, 但  $\epsilon$ -SVR 模型的支持向量个数为 447, 而  $\nu$ -SVR 模型的支持向量数目为 194。 $\nu$ -SVR 模型比  $\epsilon$ -SVR 模型简单, 精度稍低, 但在可接受的范围之内。此外, 从过程线看,  $\epsilon$ -SVR 模型以及  $\nu$ -SVR 模型均表现良好。

#### 4 结语

以鸭绿江流域临江站洪水过程预报问题为研究对

象, 建立了临江站洪水过程  $\nu$ -SVR 预报模型, 采用 1998~2014 年间的大水年份 2004、2005、2007、2010 年洪水过程资料和相应的降水资料及 2012、2013 年洪水过程资料和相应的降水资料对  $\nu$ -SVR 预报模型进行了率定和验证, 并与线性动态系统、BP 神经网络和  $\epsilon$ -SVR 模型比较, 结论如下:

(1)  $\epsilon$ -SVR 洪水过程预报模型及  $\nu$ -SVR 洪水过程预报模型较线性动态系统模型和 BP 神经网络具有较高的

精度、更好的非线性拟合能力和泛化能力。

(2)  $\nu$ -SVR 洪水过程预报模型, 相比  $\epsilon$ -SVR 能很好地控制支持向量个数, 降低模型的复杂程度的同时能保持良好的预报精度。

(3)  $\nu$ -SVR 洪水过程预报模型用于临江站洪水过程预测时, 能很好的反应十三道湾站及区间降雨对临江站洪水过程的影响; 将前期影响雨量纳入回归因子, 也降低了模型的复杂程度, 实例应用表明了模型的可行性和有效性。

参考文献:

- [1] HSU K L, GUPTA H V, SOROOSHIAN S. Artificial neural network modeling of the rainfall-runoff process [J]. Water Resources Research, 1995, 31(10): 2517-2530.
- [2] Qin JU, Zhongbo Yu, Zhenchun Hao, et al. Division-based rainfall-runoff simulations with BP neural networks and Xinanjiang model [J]. Neurocomputing, 2009, 72: 2873-2883.
- [3] 胡健伟, 周玉良, 金菊良. BP 神经网络洪水预报模型在洪水预报系统中的应用 [J]. 水文, 2015, 35(1): 20-25. (HU Jianwei, ZHOU Yuliang, JIN Juliang. Flood forecasting model on BP neural networks and its application in flood forecasting system [J]. Journal of China Hydrology, 2015, 35(1): 20-25. (in Chinese))
- [4] 杨志民, 刘广利. 不确定性支持向量机原理及应用 [M]. 北京: 科学出版社, 2007. (YANG Zhimin, LIU Guangli. The Principle and Application of Uncertainty Support Vector Machine [M]. Beijing: Science Press, 2007. (in Chinese))
- [5] 崔庆, 马孝义, 陈海涛. 最小二乘支持向量机模型在径流预报中的应用 [J]. 中国农村水利水电, 2012, 12: 63-65. (CUI Qing, MA Xiaoyi, CHEN Haitao. The least square support vector machine LS-SVM model applies in the prediction of the runoff [J]. China Rural Water and Hydropower, 2012, 12: 63-65. (in Chinese))



- [6] 卫太祥,马光文,黄炜斌. 基于惩罚加权支持向量机回归的径流预测模型[J]. 水力发电学报, 2012,31(6):35-38. (WEI Taixiang, MA Guangwen, HUANG Weibin. Runoff forecast based on weighted support vector machine regression model[J]. Journal of Hydroelectric Engineering, 2012,31(6):35-38. (in Chinese))
- [7] 李代华,崔东文. 相空间重构支持向量机在径流模拟中的应用研究[J]. 长江科学院院报, 2013,30(10):21-26. (LI Daihua, CUI Dongwen. Phase space reconstruction of support vector machine in runoff simulation[J]. Journal of Yangtze River Scientific Research Institute, 2013,30(10):21-26. (in Chinese))
- [8] 周秀平,李天翔,王文圣. 年径流预测的最小二乘支持向量机-马尔可夫链组合模型[J]. 水力发电学报, 2013,32(4):16-19. (ZHOU Xiuping, LI Tianxiang, WANG Wensheng. Annual runoff prediction based on least square support vectormachines -Markov chain combined model [J]. Journal of Hydroelectric Engineering, 2013,32(4):16-19. (in Chinese))
- [9] 张兰影,庞博,徐宗学,等. 基于支持向量机的石羊河流域径流模拟适用性评价[J]. 干旱区资源与环境, 2013,27(7):113-118. (ZHANG Lanying, PANG Bo, XU Zongxue, et al. Assessment on the applicability of support vector machine-based models for runoff simulation in Shiyang River basin [J]. Journal of Arid Land Resources and Environment, 2013,27(7):113-118. (in Chinese))
- [10] 李继伟,纪昌明,张新明,等. 基于支持向量机的水电站中长期径流组合预报[J]. 水电能源科学, 2013,31(11):13-16. (LI Jiwei, JI Changming, ZHANG Xinming, et al. Medium and long-term runoff combinational forecast based on support vector machine[J]. Water Resources and Power, 2013,31(11):13-16. (in Chinese))
- [11] 孙传文,钟平安,万新宇,等. 考虑季节因子的支持向量机径流预测模型 [J]. 中国农村水利水电, 2014,4:101-104. (SUN Chuanwen, ZHONG Pingan, WAN Xinyu, et al. Seasonal support vector mechine model of the runoff forecas[J]. China Rural Water and Hydropower, 2014,4:101-104. (in Chinese))
- [12] B. Scholkopf, A. Smola, R. C. Williamson, and P. L. Bartlett. New support vector algorithms[J]. Neural Computation, 2000,12: 1207-1245.
- [13] 张显扬,董增川,王建群,等. 雅鲁藏布江尼洋河流域洪水预报方法研究[J]. 河海大学学报, 2005,33(5):530-533. (ZHANG Xianyang, DONG Zengchuan, WANG Jianqun, et al. Method of flood forecasting for Niyang River basin of Yarlungzangbo River [J]. Journal of Hohai University (Nature Sciences), 2005,33(5):530-533. (in Chinese))
- [14] 王建群,董增川. 太湖流域平望水位和米市渡潮位过程预报研究[J]. 湖泊科学,2003,(3): 229-235. (WANG Jianqun, DONG Zengchuan. On the forecasting of Pingwang water level and Mishidu tide level hydrograph in Taihu basin [J]. Journal of Lake Sciences, 2003,(3):229-235. (in Chinese))
- [15] Chang Chih-Chung, Lin Chih-Jen. LIBSVM:a library for support vector machines [EB/OL]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [16] 王建群,卢志华,哈布哈琪. 求解约束非线性优化问题的群体复合进化算法[J]. 河海大学学报, 2001,29(3):29-43. (WANG Jianqun, LU Zhihua, Hapuarachchi H A P. Multi-complex evolution algorithm for constraint nonlinear optimization problems[J]. Journal of Hohai University (Nature Sciences), 2001,29(3):29-43. (in Chinese))

## Flood Forecasting Model with $\nu$ -Support Vector Machine

DAN Lingzhi<sup>1</sup>, WANG Jianqun<sup>1</sup>, CHEN Lixiang<sup>2</sup>, CHEN Honghong<sup>1</sup>

(1. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China;

2. Bureau of Hydrology, Songliao Water Resources Commission, Changchun 130021, China)

**Abstract:** In order to improve the capacity of nonlinear fitting and accuracy of flood forecasting, the support vector machine model was studied. For the actual situation of the Linjiang Station on the Yalujiang River, the flood forecasting model for the Linjiang Station with  $\nu$ -SVR forecasting model was established. The precipitation and discharge process data in the flood years from 1998 to 2014 in the watershed were used to calibrate and validate the model. The flood forecasting model for the Linjiang Station with  $\nu$ -SVR was compared with the linear dynamic system model, BP neural network model and  $\varepsilon$ -SVR model. The results show that the flood forecasting model with  $\nu$ -SVR has a high precision than the linear dynamic system model and BP neural network model. The flood forecasting model with  $\nu$ -SVR has better nonlinear fitting ability and generalization ability, can control the number of support vectors and reduce the complexity of the model, while maintaining a good prediction accuracy.

**Key words:** flood forecasting; support vector machine; linear dynamic systems; BP neural network