

自适应变异粒子群算法与支持向量机在 农业用水预测中的应用

孙艳, 刀海娅

(云南省水利水电勘测设计研究院, 云南 昆明 650021)

摘要: 为提高农业用水预测精度以及改善基本粒子群算法(PSO)的收敛性能,提出基于自适应变异(Adaptive Variation, AV)算法改进的 PSO-SVR 多元变量农业用水预测模型,以 2000-2011 年全国农业用水量预测为例进行实例研究。首先,选用 3 个典型函数测试 AVPSO 算法性能,并与基本 PSO 算法比较;其次选取粮食作物播种面积、水灾成灾面积等 4 个指标作为农业用水预测的影响因子,采用 AVPSO 算法优化 SVR 惩罚因子 C 和核函数参数 g , 构建 AVPSO-SVR 农业用水预测模型,并构建基本 PSO-SVR、GA-SVR、GA-BP 和传统 BP 模型作为对比模型;最后,利用实例前 8 年和后 4 年资料分别对各模型进行训练和预测。结果表明:① AVPSO 算法的全局搜索能力有了明显提高,有效避免了早熟收敛问题。② AVPSO-SVR 模型对实例后 4 年农业用水量预测的平均相对误差绝对值和最大相对误差绝对值分别为 0.48%、0.78%,预测精度及泛化能力均优于 PSO-SVR、GA-SVR、GA-BP 和传统 BP 模型,AVPSO 算法能有效对 SVR 惩罚因子 C 和核函数参数 g 进行寻优。

关键词: 需水预测; 自适应变异; 粒子群算法; 遗传算法; 支持向量机; 神经网络

中图分类号: S274.4

文献标识码: A

文章编号: 1672-643X(2015)03-0231-06

Application of AVPSO-SVR and support vector machine to agricultural water prediction

SUN Yan, DAO Haiya

(Yunnan Water Conservancy and Hydropower Survey and Design Institute, Kunming 650021, China)

Abstract: order to improve agricultural water precision and improve the basic particle swarm optimization (PSO) convergence performance, the paper proposed PSO-SVR multivariate agricultural water prediction model based on the adaptive mutation (Adaptive Variation, AV). It took national agricultural water consumption forecast as a case study from 2000 to 2011. First of all, it selected 3 typical function test the performance of AVPSO algorithm, and compared with the basic PSO algorithm; secondly choose 4 influence factors such as grain crops sown area, the flood area as that of agricultural water forecast. AVPSO algorithm is used to optimize the SVR penalty factor and kernel parameter, construction of AVPSO-SVR agricultural water consumption forecast model, and construct the basic PSO-SVR, GA-SVR, GA-BP and traditional BP model as a model for comparison; finally, it used the case before 8 years and after 4 years of data for the training and prediction of each model. The results showed that ① the global searching ability of AVPSO algorithm has been significantly improved, which can avoid the premature convergence problem. The AVPSO-SVR model of the relative mean absolute error of prediction of water with 4 years of agricultural instance and the maximum relative error absolute values were 0.48%, 0.78%. The prediction accuracy and generalization ability are better than those of PSO-SVR, GA-SVR, GA-BP and traditional BP model. AVPSO algorithm can effectively carry on the optimization of SVR penalty factor and kernel parameter.

Key words: water demand prediction; adaptive variation; particle swarm optimization; genetic algorithm; support vector machine; neural network

1 研究背景

提高农业用水预测精度对于合理制定水资源调配方案和调度计划,发挥有限水资源的最大效益具有重要意义。由于农业用水受气候条件、粮食作物播种面积、农作物种植结构、用水价格和用水管理水平等多种因素影响,用水量在年际间常表现出较大的随机性和波动性。提高农业用水预测精度一直是农业用水管理研究中的热点和难点问题。目前用于农业用水预测的方法主要有:趋势法、回归分析法、时间序列法、指标分析法、灰色预测法、系统动力学法和组合预测法等,均在农业用水预测中取得了一定的效果,但由于农业用水受多种因素的影响,其预测具有较大的不确定性、随机性和模糊性,使得上述方法的预测精度受到不同程度的影响。趋势法仅在系统比较稳定和短期预测条件下才具有较好的预测精度;回归分析法的预测精度对于自变量的选取及自变量统计值的准确性较为敏感,针对我国基本数据短缺、预测及决策体系不完善的现状,回归分析法较难达到理想的预测效果;时间序列法较适用于短期预测,且距离起点越近,预测精度越高;指标分析法受限于各区域用水定额差别较大,该法对用水系统历史数据精度要求较高,不能满足实际应用需求;灰色预测法在短期预测及具有确定性变化趋势的预测中具有较好的预测精度,对于长期预测或年际间波动较大的预测问题,其预测精度并不理想;系统动力学法具有较好的预测精度,但存在系统复杂、分析能力要求高以及预测工作量大等不足,并不适用于短期预测;组合预测法是目前广泛采用的提高预测精度的有效方法,其预测精度主要取决于各单一模型预测精度及各模型间是否具有良好的“互补性能”,在实际应用中受到一定的制约。人工神经网络(Artificial Neural Network, ANN)是一种非线性统计性数据建模工具,能以任意精度逼近任何非线性连续函数,在农业用水预测中有着广泛的应用^[1-3]。然而,由于传统 ANN 算法是基于渐近理论,仅在样本容量趋向于无穷大时其经验风险才趋近于实际风险,在实际应用中面临泛化能力差、收敛速度慢及存在局部极值等问题^[4-5]。回归支持向量机(Support Vector Regression, SVR)是 Vapnik 等人提出的一种新型机器学习方法,其优良的“理论全局最优”、“结构简单”及“克服维数灾”等性能在各行业领域具有广泛应用,尤其在解决小样本容量时,很大程度上解决了传统 BP 等网络在模型选择、过学习、高维和局

部极值等方面的问题^[6-7]。在实际应用中, SVM 学习参数——惩罚因子 C 、核函数参数 g 和不敏感系数 ε 对 SVM 性能有着关键性影响^[5],传统试凑法、经验选择法等人工方法选取 SVM 学习参数由于存在选取范围小、计算效率低、预测误差大等弊端已不能满足实际应用需求。粒子群优化算法(Particle Swarm Optimization, PSO)是近年来发展起来的一种新的进化算法(Evolutionary Algorithms, EA),同遗传算法(Genetic Algorithm, GA)类似, PSO 通过迭代寻找最优解及适应度来评价解的优劣,算法简单,具有实现容易、精度高、收敛快等优点,并被证明在大多数情况下比 GA 更有效^[8]。将 PSO 应用于 SVR 参数优化,可以快速有效地能获得大范围内的最佳参数 C 和 g ,使得优化目标全局最优。然而,基本 PSO 算法在实际应用中存在收敛速度慢、精度低和易陷入局部极值等不足^[9-10]。

针对上述分析及问题,笔者提出基于自适应变异(Adaptive Variation, AV)改进的 PSO 优化算法,并基于 SVR 算法原理,构建 AVPSO-SVR 多元变量农业用水预测模型,以 2000-2011 年全国农业用水量预测为例进行实例研究,并构建基本 PSO-SVR、GA-SVR、GA-BP 和传统 BP 模型作为对比模型。验证结果表明改进 PSO-SVR 模型具有较好的预测精度和泛化能力。

2 预测模型

2.1 回归支持向量机(SVR)

SVR 是通过核函数将低维空间中非线性回归问题映射到高维特征空间,然后在这个特征空间中求解凸优化问题,实现某一非线性变换后的线性回归。其回归算法原理如下^[5-8]:

设含有 l 个训练样本的集合为: $\{(x_i, y_i), i = 1, 2, \dots, l\}$, 其中, $x_i (x_i \in R^d)$ 是第 i 个训练样本的输入列向量, $x_i = [x_i^1, x_i^2, \dots, x_i^d]^T$, $y_i \in R$ 为对应输出值。SVM 回归用核函数将数据映射到高维特征空间,再在高维特征空间中进行线性回归,依据结构风险化最小化原则,将其学习过程转换为凸优化问题,即:

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

式中: $\Phi(x)$ 为非线性映射函数; ε 为线性不敏感损失函数; ξ_i, ξ_i^* 为松弛变量; C 为惩罚因子。

其回归方程最终表述为:

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) K(x_i, x_j) + b \quad (2)$$

式中: $K(x_i, x_j) = \Phi(x_i) \Phi(x_j)$ 为满足 Mercer 条件的核函数; a_i^*, a_i 为二次规划中 Lagrange 乘子。

$$b = \frac{1}{N_{nsv}} \left\{ \sum_{0 < a_i < C} [y_i - \sum_{x_j \in SV} (a_i - a_i^*) K(x_i, x_j) - \varepsilon] + \sum_{0 < a_i < C} [y_i - \sum_{x_j \in SV} (a_i - a_i^*) K(x_i, x_j) + \varepsilon] \right\} \quad (3)$$

式中: N_{nsv} 为支持向量机个数。

本文选择径向基核函数作为 SVM 核函数, 径向基核函数表达式为:

$$K(x, x_i) = \exp(-g \|x - x_i\|^2) \quad (4)$$

式中: $g > 0$ 。

Vapnik 等人研究表明, 在给定样本和核函数条件下, SVR 的性能主要受惩罚因子 C 、核函数参数 g 和不敏感系数 ε 的影响。惩罚因子 C 决定着由训练样本产生的经验风险对模型性能的影响, 即经验风险随着 C 值的增加而增加、减小而减小, 当 C 值无穷大时, SVR 结构风险趋于经验风险; 而当 C 值趋于零时, 由于 SVR 模型无法获取训练样本信息, 模型失去解决问题的能力。核函数参数 g 的改变意味着改变映射函数, 因此它主要影响样本数据在高维特征空间中分布的复杂程度; 不敏感系数 ε 用于控制支持向量的个数, 平衡模型的复杂程度与模型对训练样本维数的依赖程度。在实际应用中, 若 ε 值过小, 可能导致模型“过拟合”, 并且增加训练时间; ε 值过大, 则可能导致模型“欠拟合”。

因此, 本文在给定不敏感系数 ε 的条件下, 基于改进 PSO 算法对 SVR 惩罚因子 C 和核函数参数 g 进行寻优, 以期获得更好的预测效果。

2.2 基本粒子群优化算法

PSO 是近年来发展起来的一种新型启发式进化算法, 最早由 Kennedy 和 Eberhart 于 1995 年受鸟群觅食过程中迁徙和群聚行为而提出的群体智能全局随机搜索算法, 具有简单、容易实现、收敛速度快和全局寻优等特点, 并被证明在大多数情况下比遗传算法(GA)更有效。PSO 算法可描述为^[9]:

设在一个 S 维的搜索空间中, 由 n 个粒子组成的种群 $W = (W_1, W_2, \dots, W_n)$, 其中第 i 个粒子表示为一个 S 维的向量 $W_i(w_{i1}, w_{i2}, \dots, w_{is})^T$, 代表第 i 个粒子在 S 维搜索空间中的位置, 表示一个问题的潜

在解。根据目标函数可计算出每个粒子位置 W_i 所对应的适应度值, 并将第 i 个粒子的速度记为 $V_i(V_{i1}, V_{i2}, \dots, V_{is})^T$, 其个体极值记为 $P_i(P_{i1}, P_{i2}, \dots, P_{is})^T$, 种群全局的极值记为 $P_g(P_{g1}, P_{g2}, \dots, P_{gs})^T$ 。在每一迭代过程中, 粒子通过个体极值和全局极值更新自身的速度和位置, 更新公式为:

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - W_{id}^k) + c_2 r_2 (P_{gd}^k - W_{id}^k) \quad (5)$$

$$W_{id}^{k+1} = W_{id}^k + V_{id}^{k+1} \quad (6)$$

式中: ω 为惯性权重; $d = 1, 2, \dots, S$; $i = 1, 2, \dots, n$; k 为当前迭代次数; V_{id} 为粒子的速度; c_1 和 c_2 为非负常数, 称为加速因子; r_1 和 r_2 为分布于 $[0, 1]$ 之间的随机数。

2.3 自适应变异改进粒子群算法

粒子群算法虽然收敛速度快, 但同时存在易早熟收敛、搜索精度低、后期迭代效率不高等缺点。借鉴遗传算法中变异思想, 在 PSO 算法中引入变异操作对粒子以一定的概率重新初始化。变异操作拓展了在迭代中不断缩小的种群搜索空间, 使粒子能够跳出先前搜索到的最优值位置, 在更大的空间中进行搜索, 同时保持了种群多样性, 提高算法寻找到更优值的可能性^[8,11]。本文按照式(7)~(9)进行算法改进:

$$f_{avg} = \frac{1}{m} \sum_{i=1}^m f_i \quad (7)$$

$$f = \begin{cases} \max\{|f_i - f_{avg}|\}, & \max\{|f_i - f_{avg}|\} > 1 \\ 1, & \text{其它} \end{cases} \quad (8)$$

$$e^2 = \sum_{i=1}^m \left(\frac{f_i - f_{avg}}{f} \right)^2 \quad (9)$$

$$p_m = \begin{cases} \mu, & e^2 < e_d^2 \text{ 且 } f(g_{best}) > f_d \\ 0, & \text{其它} \end{cases} \quad (10)$$

式中: f_i 为第 i 个粒子的适应度; e^2 为群体适应度方差; μ 取 $[0.1, 0.3]$ 之间的任意数值; e_d^2 取远小于 e^2 的最大值; f_d 可设为理论最优值; 对于 g_{best} 的变异操作, 采用增加随机扰动的方法, 设 g_{besti} 为 g_{best} 的第 i 维取值, η 是服从 Gauss(0,1) 分布的随机变量, 则

$$g_{besti} = g_{besti} (1 + 0.5\eta) \quad (11)$$

2.4 AVPSO-SVR 预测模型

定义目标函数:

$$\begin{cases} \min f(C, g) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2 \\ \text{s. t. } C \in [C_{\min}, C_{\max}], g \in [g_{\min}, g_{\max}] \end{cases} \quad (12)$$

式中： y_i 为第 i 个样本实测值； $\hat{y}$ 为第 i 个样本模拟值。SVR 模型参数优化的思想就是通过迭代算法搜寻一组参数 (C, g) ，通过 SVR 学习使公式 (12) 目标函数达到最小。

本文选取公式 (12) 作为 AV - PSO 算法的目标函数，利用 AV - PSO 算法对参数 (C, g) 进行优化。具体算法步骤描述如下^[11-13]：

(1) 选取 SVR 模型的训练样本和检验样本，设定惩罚因子 C 和核函数参数 g 搜寻范围；确定目标函数，初始化改进 PSO 算法中的种群 $W = (W_1, W_2, \dots, W_n)$ 、最大迭代次数，惯性权重 ω 、加速因子 c_1 和 c_2 等参数取值，随机初始化待优化参数 C 和 g 的可行解位置 $W_i = (w_{i1}, w_{i2}, \dots, w_{is})^T$ 和速度 $V_i = (V_{i1}, V_{i2}, \dots, V_{is})^T$ 。

(2) 根据当前参数 C 和 g ，由式 (12) 计算适应度值，记忆个体与群体所对应的最佳适应度的位置 $P_i = (P_{i1}, P_{i2}, \dots, P_{is})^T$ 与 $P_g = (P_{g1}, P_{g2}, \dots, P_{gs})^T$ 。

(3) 根据式 (5)、式 (6) 更新粒子速度和位置。

(4) 根据式 (7) ~ 式 (9) 计算群体适应度方差 e^2 和定标因子 f 。

(5) 根据式 (10) 计算变异概率 p_m ，产生随机数，对满足变异条件的 g_{best} 按一定的概率 p_m 变异。

(6) 判断算法迭代终止条件是否满足，如果满足，转向 g ；否则，执行 c 。

(7) 获得最佳惩罚因子 C 和核函数参数 σ^2 ，利用 LSSVM 模型对检验样本进行预测。

3 算法验证

为验证 AVPSO 算法的性能，选取 3 个典型函数的极值优化来测试，并与基本 PSO 算法比较。函数如下：

(1) Rosenbrock 函数：

$$f_1(x) = \sum_{i=1}^n [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$$
$$x_i \in [-5, 5]$$

(13)

(2) Rastrigin 函数：

$$f_2(x) = \sum_{i=1}^n (x_i^2 - 10\cos(2\pi x_i) + 10)$$
$$x_i \in [-5.12, 5.12]$$

(14)

(3) Ackley 函数：

$$f_3(x) = -20\exp(0.2\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^n \cos(2\pi x_i)) + 22.71282$$

(15)

上述 3 个测试函数的全局极小值都为 0，其中

Rosenbrock 函数是很难极小化的病态单峰二次函数；Rastrigin 函数和 Ackley 函数为多峰函数。测试在 20 维空间中进行，对于所有测试，设定算法最大迭代次数为 2000 次，种群大小设为 50，粒子搜索范围设定为测试函数的取值范围，基本 PSO 算法的学习因子 c_1 、 c_2 均设为 1.49945，惯性权重 ω_{strat} 取 0.9， ω_{end} 取 0.4；AVPSO 算法学习因子 c_1 、 c_2 和惯性权重 ω 同基本 PSO 算法， e_d^2 取 6.5； f_d 取 10^{-6} 。对每个函数连续寻优 50 次，并以平均最优适应度值作为算法的评价指标，见表 1。

表 1 2 种算法寻优测试结果比较

函数	基本 PSO 算法	AVPSO 算法
$f_1(x)$	15.5063	8.5725
$f_2(x)$	0.8780	0.7731
$f_3(x)$	11.8105	11.7050

从表 1 可以看出，对所测试的 3 个函数 AVPSO 算法均获得比基本 PSO 算法更优的结果。虽然 AVPSO 算法对于 3 个函数的寻优结果与函数本身的全局极值还存在一定的差距，但相对于基本 PSO 算法的寻优结果，AVPSO 算法具有较好的优势。因此，本文采用 AVPSO 算法对 SVR 惩罚因子 C 和核函数参数 g 进行寻优。

4 实例应用

4.1 数据来源与分析

由于农业用水受气候条件、粮食作物播种面积、农作物种植结构、灌溉制度、用水价格和用水管理水平等多种因素影响，用水量在年际间常表现出较大的随机性和波动性。本文利用 SPSS 软件从农业总产值、年均降水量、粮食作物播种面积、水灾成灾面积等 10 余个统计指标中遴选出与农业用水量显著的粮食作物播种面积、稻谷播种面积、小麦播种面积和水灾成灾面积 4 个统计指标作为农业用水预测的影响因子，资料来源于《2012 年中国统计年鉴》，见表 2。

4.2 应用

将 AVPSO - SVR、PSO - SVR、GA - SVR、GA - BP 和传统 BP 模型用于农业用水预测的步骤分为：训练及预测样本的选取、数据的预处理、模型的构建、模型训练及预测、预测结果分析 5 个步骤，详述如下：

(1) 训练及预测样本的选取。以表 2 中与农业用水具有较好相关关系的粮食作物播种面积、稻谷

播种面积、小麦播种面积和水灾成灾面积作为农业用水预测的影响因子,并以 2000 – 2007 年统计数据作为训练样本,2008 – 2011 年作为预测样本。

(2)数据的预处理。利用下式(16)对表 2 中数据进行归一化处理:

$$\hat{x} = (x - x_{\min}) / (x_{\max} - x_{\min}) \tag{16}$$

式中: $\hat{x}$ 为经过标准化处理的数据; x 为原始数据; $x_{\max}$ 和 $x_{\min}$ 分别为数据序列中的最大数和最小数。

表 2 全国 2000 – 2011 年农业用水量及

影响因子统计值 亿 m³, 万 hm²

年份	农业	粮食作物	稻谷播种	小麦播种	水灾成灾
	用水量	播种面积	面积	面积	面积
2000	3783.5	10846.25	2996.17	2665.33	432.10
2001	3825.7	10608.00	2881.24	2466.38	361.40
2002	3736.2	10389.08	2820.16	2390.83	738.79
2003	3432.8	9941.04	2650.78	2199.69	1228.90
2004	3585.7	10160.60	2837.88	2162.60	374.65
2005	3580.0	10427.84	2884.72	2279.26	604.70
2006	3664.4	10495.80	2893.79	2361.30	456.88
2007	3599.5	10563.84	2891.88	2372.06	510.49
2008	3663.5	10679.26	2924.11	2361.72	365.56
2009	3723.1	10898.57	2962.69	2429.08	316.20
2010	3689.1	10987.61	2987.34	2425.65	702.42
2011	3743.6	11057.30	3005.70	2427.04	283.95
相关系数	0.719 **	0.707 *	0.810 **	–0.661 *	

注: “**”、“*” 分别表示在 0.01 水平和 0.5 水平(双侧)上显著相关。

(3)模型的构建。本文基于 Matlab 软件环境和 libsvm 工具箱、谢菲尔德(Sheffield)遗传算法工具箱构建模型,以粮食作物播种面积、稻谷播种面积、小麦播种面积和水灾成灾面积作为模型输入向量,农业用水量作为输出向量,创建 4 输入 1 输出的 AVPSO – SVR、PSO – SVR、GA – SVR、GA – BP 和传统 BP 预测模型。

(4)模型的训练及预测。利用所选取的训练样本和预测样本对 AVPSO – SVR、PSO – SVR、GA –

SVR、GA – BP 和传统 BP 模型进行训练及预测,并选取平均相对误差绝对值 MRE 和最大相对误差绝对值 $\max RE$ 对各模型的预测效果进行评价,公式如下:

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \times 100 \tag{17}$$

$$\max RE = \max_{1 \leq i \leq n} \frac{|\hat{y}_i - y_i|}{y_i} \times 100 \tag{18}$$

式中: $\hat{y}_i$ 为第 i 个样本预测值; y_i 为第 i 个样本实测值; $i = 1, 2, \dots, n$, n 为预测样本数。

各模型参数设置如下:AVPSO – SVR 模型:惩罚因子 C 和核函数参数 g 的搜索空间分别设为 0.1 ~ 100、0.01 ~ 1000;局部和全局搜索学习因子 $c_1 = 1.5$ 、 $c_2 = 1.7$;最大迭代次数 = 100;种群规模 = 30;惯性权值 $\omega_{\text{first}} = 0.9$ 、 $\omega_{\text{end}} = 0.4$; $e_d^2 = 6.5$; $f_d = 10^{-6}$;交叉验证参数 $C_v = 3$;不敏感系数 ε 设为 0.01。PSO – SVR 模型:除自适应变异相关部分外,其他参数设置同 AVPSO – SVR 模型;GA – SVR 模型:惩罚因子 C 和核函数参数 g 的搜索空间分别设为 0 ~ 100、0 ~ 1000;最大迭代次数 = 100;种群规模设 = 30;变异系数 = 0.9; $CV = 5$;不敏感系数 ε 设为 0.01。传统 BP 模型:参考文献[8,14],确定该实例 BP 模型结构为 4 – 7 – 1,隐含层和输出层传递函数分别采用 logsig 和 purelin,训练函数采用 traingdx,学习速率 $lr = 0.05$,设定期望误差 = 0.01,最大训练轮回 = 100 次。GA – BP 模型:网络结构及参数设置同 BP 传统模型,在此条件下运用 GA 来优化 BP 神经网络的初始权值和阈值。GA 参数设置为:种群规模 = 100,进化次数 = 200 次。

利用训练好的 AVPSO – SVR、PSO – SVR、GA – SVR、GA – BP 和传统 BP 模型对全国 2008 – 2011 年农业用水量进行预测,结果见表 3。并绘制全国 2000 – 2011 年农业用水量拟合、预测结果及相对误差图,见图 1 ~ 图 2。

表 3 2008 – 2011 年农业用水量预测结果及比较

预测样本	实际值/ (m ³ · s ⁻¹)	AVPSO – SVR 模型		PSO – SVR 模型		GA – SVR 模型		GA – BP 模型		传统 BP 模型	
		预测值/	相对	预测值/	相对	预测值/	相对	预测值/	相对	预测值/	相对
		亿 m ³	误差/%	亿 m ³	误差/%	亿 m ³	误差/%	亿 m ³	误差/%	亿 m ³	误差/%
2008	3663.5	3692.1	–0.78	3696.1	–0.89	3693.7	–0.82	3719.2	–1.52	3795.0	–3.59
2009	3723.1	3739.4	–0.44	3748.8	–0.69	3741.0	–0.48	3783.0	–1.61	3813.8	–2.44
2010	3689.1	3697.4	–0.22	3698.5	–0.25	3706.3	–0.47	3705.3	–0.44	3657.3	0.86
2011	3743.6	3761.0	–0.46	3756.1	–0.33	3766.0	–0.60	3795.8	–1.39	3773.1	–0.79
$MRE / \%$		0.48		0.54		0.59		1.24		1.92	
$\max RE / \%$		0.78		0.89		0.82		1.61		3.59	

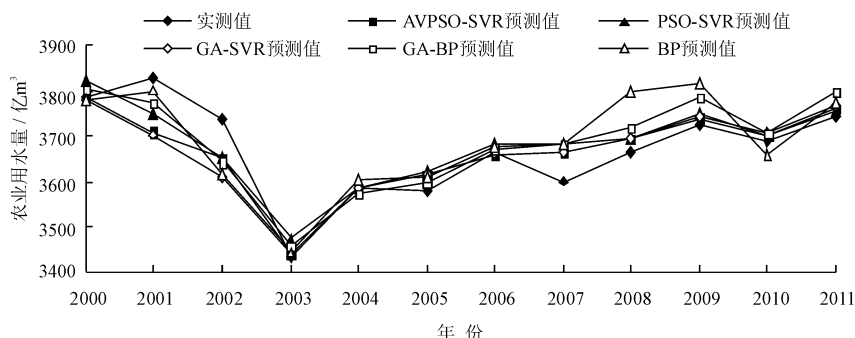


图1 2000 - 2011 年全国农业用水量拟合及预测值比较

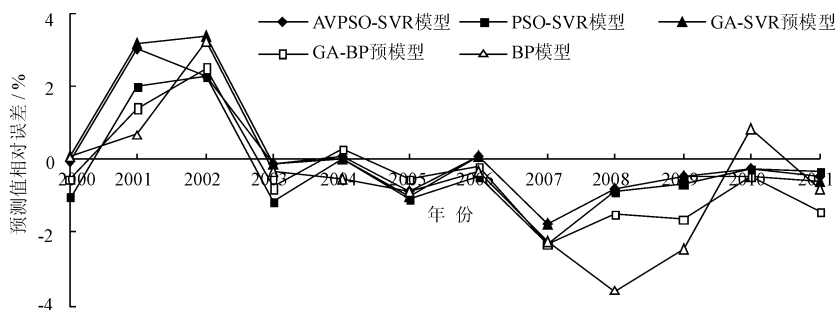


图2 2000 - 2011 年全国农业用水量拟合及预测相对误差

(5) 预测结果分析。分析表3及图1~图2可以得出以下结论:① AVPSO-SVR模型对于实例2008-2011年预测的 MRE 和 $\max RE$ 分别为0.48%、0.78%,预测精度及泛化能力均优于 PSO-SVR、GA-SVR、GA-BP 和传统 BP 模型,表明 AV-PSO 能有效对 SVR 惩罚因子 C 和核函数参数 g 进行优化,模型具有较好的预测精度和泛化能力,完全能满足农业用水预测的精度要求;②由于 SVR 模型是基于结构风险最小化原则,由经验风险和置信区间共同决定支持向量的实际风险,因此在本实例应用中,AVPSO-SVR、PSO-SVR 及 GA-SVR 模型的预测精度和泛化能力均优于 GA-BP 和传统 BP 神经网络模型。③对于 BP 神经网络而言,在相同网络结构、传递函数、训练函数及期望误差等条件下,GA-BP 模型预测精度、泛化能力及稳健性能均优于传统 BP 模型,表明 GA 能有效优化 BP 网络初始权值和阈值。从图1~2也可看出,传统 BP 模型拟合精度远优于预测精度,模型已呈现“过拟合”趋势。④从图1~2来看,由于受气候条件、粮食作物播种面积、农作物种植结构等多种因素影响,农业用水量呈现出较大的随机性和波动性,本文除传统 BP 模型预测精度不够理想外,其余模型均能对全国2000-2011年农业用水量进行有效拟合及预测,达到实际应用需求。

5 结 语

农业用水量约占全国用水总量的70%,提高农业用水量预测精度对于国家制订相关水资源战略规划及实行最严格水资源管理制度均具有重要意义。本文分析了 ANN、SVR 及基本 PSO 算法在实际应用中存在的问题和不足,针对问题与不足提出 SVR 惩罚因子 C 和核函数参数 g 的 AVPSO 优化方法,以3个典型函数测试了 AVPSO 算法的全局优化性能,并分析了农业用水的主要影响因子,提出 AVPSO-SVR 多元变量农业用水预测模型,构建基本 PSO-SVR、GA-SVR、GA-BP 和传统 BP 模型作为对比模型,以2000-2011年全国农业用水量预测为例进行实例研究,研究表明 AVPSO-SVR 模型兼顾了启发式优化算法与结构风险最小化算法二者的优点,克服了基本 PSO 算法收敛速度慢、易陷入局部极值等的不足,模型具有较好的预测精度高和泛化能力,可为相关预测研究提供借鉴。

参考文献:

- [1] 宋巧娜,唐德善. 基于灰色理论和 BP 神经网络的农业用水量预测[J]. 农机化研究,2007(9):53-55.
- [2] 赵振国,刘丽,徐建新. 基于 BP 神经网络预测区域农业用水量[J]. 人民黄河,2007,29(9):59-60+62.

在 $1\ 800\ \text{m}^3/\text{hm}^2$ 的灌水量时, ΔM 大于零,说明 $0 \sim 60\ \text{cm}$ 土壤土层整体脱盐,达到了淋洗的目的, K 值为 0.033 。在 $2\ 250\ \text{m}^3/\text{hm}^2$ 的灌水量时,与 $1\ 800\ \text{m}^3/\text{hm}^2$ 的一样都达到了淋洗效果, K 值为 0.052 。 K 值越大说明淋洗效果越不好,即在 $1\ 350$ 、 $1\ 800$ 和 $2\ 250\ \text{m}^3/\text{hm}^2$ 的灌水量下, $1\ 800\ \text{m}^3/\text{hm}^2$ 的灌水量的淋洗效果最好。

3 结 语

(1) 在土壤土层 $60\ \text{cm}$ 以下,灌水量对不同行距的盐分变化无影响。在土壤土层 $0 \sim 60\ \text{cm}$ 处,宽行、窄行和膜间不同行距的脱盐效果不同,其中宽行的脱盐效果最好。

(2) 比较 $1\ 350$ 、 $1\ 800$ 、 $2\ 250\ \text{m}^3/\text{hm}^2$ 3 个灌水量的效果, $1\ 800\ \text{m}^3/\text{hm}^2$ 的淋洗效果最优; $1\ 350\ \text{m}^3/\text{hm}^2$ 灌水量太小,达不到淋洗目的; $2\ 250\ \text{m}^3/\text{hm}^2$ 灌水量过大会导致地下水位上升,在没有排水的情况下,盐分会随着地下水位的上升而上升,也达不到淋洗盐分的目的。

本结论可为该地区棉田春灌提供参考,对于有土壤盐渍化影响的地区制定灌溉制度也有借鉴作用。但是影响盐分变化的因素有很多,如土质、地下水位及矿化度、初始含盐量、灌水量和灌水时间等

等,本次试验只考虑灌水量一种因素进行分析,建议把灌水、地下水位和排水这一灌排相结合的系统,综合考虑,确定合适的棉田春灌制度。

参考文献:

- [1] 亓沛沛,冉圣宏,张 凯. 不同灌溉方式和作物类型对西北干旱区耕地土壤盐渍化的影响[J]. 农业环境科学学报,2012,31(4):780-785.
- [2] 胡明芳,田长彦,赵振勇,等. 新疆盐碱地成因及改良措施研究进展[J]. 西北农林科技大学学报(自然科学版),2012,40(10):111-117.
- [3] 张 伟,吕 新,李鲁华,等. 新疆棉田膜下滴灌盐分运移规律[J]. 农业工程学报,2008,24(8):15-19.
- [4] 孙三民,安巧霞. 塔里木灌区冬灌对盐分运移的影响[J]. 塔里木大学学报,2008,20(4):34-37.
- [5] 刘 磊. 干旱区棉花膜下滴灌水盐运移规律及数值模拟研究[D]. 乌鲁木齐:新疆农业大学,2011.
- [6] 单鱼洋. 干旱区膜下滴灌水盐运移规律模拟及预测研究[D]. 北京:中国科学院研究生院(教育部水土保持与生态环境研究中心),2012.
- [7] 王在敏,何雨江,靳孟贵,等. 运用土壤水盐运移模型优化棉花微咸水膜下滴灌制度[J]. 农业工程学报,2012,28(17):63-70.
- [8] 杨鹏年,董新光,刘 磊,等. 干旱区大田膜下滴灌土壤盐分运移与调控[J]. 农业工程学报,2011,27(12):90-95.
- [9] 李 飞,罗明文,张运鑫,等. 基于 RBF 神经网络的农业需水量预测[J]. 人民黄河,2008,30(8):62-63+66.
- [10] 董晓华,刘 超,喻 丹,等. 基于平均线性粒子群算法的人工神经网络在径流预报中的应用[J]. 水文,2013,33(5):10-15.
- [11] 田文奇,和敬涵,姜久春,等. 基于自适应变异粒子群算法的电动汽车换电站充电调度多目标优化[J]. 电网技术,2012,36(11):25-29.
- [12] 龙 文,梁昔明,龙祖强,等. PSO-LSSVM 灰色组合模型在地下水埋深预测中的应用[J]. 系统工程理论与实践,2013,33(1):243-248.
- [13] 白春华,周宣赤,林大超,等. 消除 EMD 端点效应的 PSO-SVM 方法研究[J]. 系统工程理论与实践,2013,33(5):1298-1306.
- [14] 崔东文. 多隐层 BP 神经网络模型在径流预测中的应用[J]. 水文,2013,33(1):68-73.
- [15] 田雨波. 混合神经网络技术[M]. 北京:科学出版社,2009.
- [16] 王 雷. 支持向量机在汽轮机状态监测中的应用[M]. 北京:北京师范大学出版社,2012.
- [17] 崔东文. 支持向量机在湖库营养状态识别中的应用[J]. 水资源保护,2013,29(4):26-30.
- [18] 崔东文. 支持向量机在水资源类综合评价中的应用研究——以全国 31 个省级行政区水资源合理性配置为例[J]. 水资源保护,2013,29(5):20-27.
- [19] 王小川,史 峰,郁 磊,等. MATLAB 神经网络 43 个案例分析[M]. 北京:北京航空航天大学出版社,2013.
- [20] 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.