

城市小时级需水量的改进型引力搜索算法- 最小二乘支持向量机模型预测

戢 钢, 王景成[†], 葛 阳, 刘华江, 杨丽雯

(上海交通大学 自动化系; 系统控制与信息处理教育部重点实验室, 上海 200240)

摘要: 本文研究利用最小二乘支持向量机(least squares support vector machine, LS-SVM)算法建立城市小时级需水量预测模型. 采取精英策略, 自适应的速度更新权重系数, 同时引入粒子历史最优信息对引力搜索算法(gravitational search algorithm, GSA)进行了改进. 最后采用改进型引力搜索算法(ameliorated gravitational search algorithm, AGSA)优化LS-SVM水量预测模型的正规化参数和核参数来提高模型的预测精度及预测速度. 理论测试与实例分析表明, 基于AGSA比基于GSA, 遗传算法(genetic algorithms, GA)和粒子群优化算法(particle swarm optimization, PSO)的LS-SVM水量预测模型具有更好的预测精度, 从而验证了基于AGSA的LS-SVM算法适用于小时级需水量预测问题, AGSA适用于多领域的模型参数的优化过程.

关键词: 智能控制; 需水量预测; 最小二乘支持向量机; 改进的引力搜索算法

中图分类号: TP181 文献标识码: A

Gravitational search algorithm-least squares support vector machine model forecasting on hourly urban water demand

Ji Gang, WANG Jing-cheng[†], GE Yang, LIU Hua-jiang, YANG Li-wen

(Department of Automation, Shanghai Jiao Tong University; Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China)

Abstract: We investigate the model of hourly urban water demand forecasting with least squares support vector machine (LS-SVM). The convergence performance of gravitational search algorithm (GSA) is improved by employing an elite strategy and an adaptive velocity with updated weighting factor. Furthermore, the historical optimal information is introduced to speed up the convergence of GSA. The ameliorated GSA, called AGSA, is adopted to optimize the regularization parameters and kernel parameters of LS-SVM used in the hourly urban water demand prediction model. Theoretical analysis and experimental results show that the AGSA-based hourly urban water demand forecasting model achieves better regression precision than other models respectively based on particle swarm optimization (PSO), genetic algorithms (GA), and GSA. This implies that AGSA-based LS-SVM algorithm can be successfully used to build the model of hourly urban water demand forecasting. In fact, AGSA can also be applied to process of parameter optimization in many other fields.

Key words: intelligent control; water demand forecasting; least squares support vector machine; ameliorated gravitational search algorithm

1 引言(Introduction)

城市用水需求的日益增加, 使得如何有效经济地维持供需平衡成为当今社会的重大课题. 近年来, 需水量的短期预测研究取得了显著性成果^[1-3]. 传统预测算法的可靠性是基于大量的样本数据, 实际中诸多因素的信息是难以获取的. 最小二乘支持向量机(least squares support vector machine, LS-SVM)的基于小样本数据的特点使得其在水量预测问题中具有显著的优势.

LS-SVM在解决小样本, 高维模式识别等问题中有着独特的优势, 具有较高的泛化性能. 其将模型的训练过程归结为线性方程组求解问题, 大大提高了训练的速度^[4-5]. LS-SVM的正规化参数和核参数影响模型的预测精度和泛化能力. 网格法的参数寻优效率太低, 无法满足生产的需求. 蚁群算法^[6]、教学算法^[7]和引力搜索算法^[8]等一批启发式搜索算法在LS-SVM的参数寻优中逐渐得到应用, 解决了常规寻优算法搜索速度慢的难题. 综合算法收敛速度与收敛精度两方

收稿日期: 2013-12-06; 录用日期: 2014-09-23.

[†]通信作者. E-mail: jcwang@sjtu.edu.cn; Tel.: +86 21-34205374.

基金项目: 国家自然科学基金资助项目(61174059, 61233004, 61433002); 国家“973”计划资助项目(2013CB035406); 上海市经信委重大技术装备研制专项基金资助项目(ZB-ZBYZ-01112634); 上海市经信委引进技术与创新项目资助(12GA-31).

面的性能指标,相对于其他的启发式算法,引力搜索算法(gravitational search algorithm, GSA)具有相对的优势.目前,GSA算法已被应用于水轮机调速系统和汽轮机热耗率等多领域的参数辨识和预测建模问题中,结果都表明GSA算法具有很好的优化性能,因此,GSA算法适用于LS-SVM参数的优化.

本文对GSA算法主要有3个方面的改进:1)采取精英策略,选取最优粒子更新粒子群的每个粒子加速度,加快收敛过程.2)速度更新中,采取基于粒子重量的自适应惯性权重系数.惯性权重系数决定了对粒子当前速度继承的多少,合适的选择可以使粒子具有平衡的探索能力.采取自适应的惯性权重,能使得GSA具有较好的收敛速度和收敛精度.3)传统的GSA只考虑当前粒子群中个体之间的相互作用,而不考虑个体历史最优和群体历史最优粒子对当前粒子群的作用.结合粒子群优化(particle swarm optimization, PSO)算法的记忆功能改进GSA的优化性能,加速粒子的学习过程.本文最后将AGSA应用于LS-SVM的参数寻优,并基于实际数据,建立城市小时级需水量预测模型,为城市供水系统的实时优化调度提供依据.

2 引力搜索算法和最小二乘支持向量机(Gravitational search algorithm and least squares support vector machine)

2.1 引力搜索算法(Gravitational search algorithm)

GSA是一种新型的基于随机种群的启发式优化算法.在GSA种群中,个体是具有质量的粒子,用粒子的质量大小来评价它的优劣,粒子质量越大,则适应度值越小,粒子越优.粒子之间通过万有引力的作用,向质量较大的粒子靠近,从而使种群达到逐步收敛的过程.

设 D 维空间中 L 个粒子组成一个群体 X ,第 i 个粒子的位置和飞行速度分别为

$$\begin{aligned} x_i &= (x_{i1}, x_{i2}, \dots, x_{iD}), \\ v_i &= (v_{i1}, v_{i2}, \dots, v_{iD}). \end{aligned}$$

根据粒子适应度更新粒子的质量得到

$$\begin{cases} q_i(t) = \frac{\text{fit}_i(t) - \text{worst}(t)}{\text{best}(t) - \text{worst}(t)}, \\ M_i(t) = \frac{q_i(t)}{\sum_{j=1}^L q_j(t)}, \end{cases} \quad (1)$$

其中: $\text{fit}_i(t)$ 和 $M_i(t)$ 分别是第 t 代种群中粒子 i 的适应度和质量.最优适应度和最差适应度的定义如下面的式(2)–(3):

$$\text{best}(t) = \min_{j \in \{1, \dots, L\}} \text{fit}_j(t), \quad (2)$$

$$\text{worst}(t) = \max_{j \in \{1, \dots, L\}} \text{fit}_j(t). \quad (3)$$

通过引力的叠加性原理,更新粒子的加速度、速度和位置:

$$\begin{cases} \alpha_i^d(t) = \sum_{j \in L, j \neq i} \text{rand}_j G(t) \frac{M_j(t) M_i(t)}{R_{ij}(t)} \times \\ \quad (x_j^d(t) - x_i^d(t)), \\ v_i^d(t+1) = \text{rand}_{it} * v_i^d(t) + \alpha_i^d(t) \Delta t, \\ x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \Delta t, \\ d = 1, 2, \dots, D, \\ i = 1, 2, \dots, L, \end{cases} \quad (4)$$

式中: α_i^d , v_i^d , x_i^d 是粒子 i 在 d 维空间中的加速度、速度和位置; rand_{it} 和 rand_j 是 $[0, 1]$ 之间的随机数; $R_{ij}(t)$ 是粒子 i 和 j 之间的欧氏距离,计算公式为

$$R_{ij}(t) = \|X_i(t), X_j(t)\|_2,$$

$X_i(t)$ 和 $X_j(t)$ 分别表示第 t 代粒子群 $X(t)$ 中第 i 个和第 j 个粒子; $G(t)$ 是第 t 代种群的重力加速度,其计算公式为

$$g(t) = g(t_0) * \left(\frac{t}{\max_t}\right)^\beta, \quad \beta < 1,$$

\max_t 是最大循环次数; Δt 表示每次参数优化过程中粒子加速度对粒子的速度与位置的作用时间,定义 $\Delta t = 1$.

2.2 最小二乘支持向量机(Least squares support vector machine)

LS-SVM继承了SVM的结构风险最小化原则,但在优化目标中采取误差的二范数作为它的损失函数.在预测方面,SVM与LS-SVM都具有较好的预测精度但LS-SVM大大提高了预测的速度.

设训练样本为 $(x_i, y_i)_{i=1}^l$,其中: $x_i \in \mathbb{R}^n$ 为输入数据, $y_i \in \mathbb{R}$ 是相应的目标值.Suykens将误差 e_i 的二范数定义为损失函数 $J(\omega, e)$,利用损失函数最小化原则确定最优线性决策函数 y 的参数,则LS-SVM优化问题变为

$$\begin{cases} \min_{\omega, e} J(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^l e_i^2, \quad \gamma > 0, \\ \text{s.t. } y_i = \omega^T \phi(x_i) + b + e_i, \quad i = 1, 2, \dots, l, \end{cases} \quad (5)$$

式中: γ 是可调的归一化参数,影响模型的泛化能力;误差变量 $e_i \in \mathbb{R}$; b 为偏差量; $\phi(i)$ 为LS-SVM的核空间映射函数,把低维样本向量 x_i 从原空间 \mathbb{R}^n 映射到高维特征空间 \mathbb{R}^m ;权矢量 $\omega \in \mathbb{R}^m$.采用核函数可解决高维特征空间点积运算时存在的“维数灾难”.

LS-SVM在对偶空间上采取 ω ,将不等式约束变为等式约束.引入拉格朗日函数:

$$\begin{aligned} L(\omega, b, e, \alpha) &= \\ J(\omega, e) - \sum_{i=1}^l \alpha_i [\omega^T \phi(x_i) + b + e_i - y_i], \end{aligned} \quad (6)$$

式中 α 是拉格朗日乘子. 式(6)获得最优解的条件是:

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \Rightarrow \omega = \sum_{i=1}^l \alpha_i \phi(x_i), \\ \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^l \alpha_i = 0, \\ \frac{\partial L}{\partial e_i} = 0 \Rightarrow \alpha_i = \gamma e_i, \\ \frac{\partial L}{\partial \alpha_i} = 0 \Rightarrow \omega^T \phi(x_i) + b + e_i = y_i. \end{cases} \quad (7)$$

上式可简化为矩阵方程:

$$\begin{bmatrix} 0 & E^T \\ E & \Omega + \frac{1}{\gamma} E \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}, \quad (8)$$

式中: $E = [1, \dots, 1]^T$, $\Omega \in \mathbb{R}^{l \times l}$. Ω_{ij} 计算公式为

$$\begin{aligned} \Omega_{ij} &= \langle \phi(x_i), \phi(x_j) \rangle = \\ \phi(x_i)^T \phi(x_j) &= K(x_i, x_j), \quad i, j = 1, \dots, l, \end{aligned} \quad (9)$$

则公式的线性解为

$$\begin{cases} b = \frac{E^T \psi^{-1} y}{E^T \psi^{-1} E}, \\ \alpha = \psi^{-1} (y - Eb), \end{cases} \quad (10)$$

式中 $\Psi = \Omega + \gamma^{-1} E$.

LS-SVM常用的核函数有多项式核函数, RBF核函数和Sigmoid函数. 其中RBF核函数的参数选择容易, 且当参数在有效范围内改变时, 空间复杂度变化小, 易于实现, 因此本文采用RBF核函数.

3 改进的GSA(The ameliorated GSA)

标准GSA算法根据个体的适配信息进行搜索, 粒子的加速度是根据其他粒子的叠加引力获得的. 忽略质量较小的粒子的影响, 加快收敛速度.

根据牛顿万有引力定律, 引力是与粒子的质量成正比的. 当粒子质量很小时, 可以忽略其对其他粒子的作用. 在加速度更新中采取精英策略, 只有最优粒子 k_{best} 会影响其他粒子的加速度.

$$\alpha_i^d(t) = \sum_{j \in k_{best}, j \neq i} \text{rand}_j G(t) \frac{M_j(t) M_i(t)}{R_{ij}(t)} \times (x_j^d(t) - x_i^d(t)). \quad (11)$$

在粒子速度迭代中, 速度权重系数对算法的收敛性具有重大作用, 权重系数越大, 收敛速度越快, 有利于全局搜索, 但不易得到精确的解; 较小的权重系数有利于局部搜索, 能得到更为精确的解, 但收敛速度慢.

GSA速度的迭代更新权重系数采用基于适应度的自适应的惯性权重策略:

$$\text{rand}_{it} = 1 - \frac{M_i(t)}{M_{best}(t)}, \quad (12)$$

式中 M_{best} 为当前种群粒子的质量最大值.

GSA是无记忆算法, 粒子学习能力较弱. 通过将个体历史最优和群体历史最优信息引入速度更新策略, 使得群体更快地达到最优解:

$$\begin{aligned} v_i^d(t+1) &= (1 - \frac{M_i(t)}{M_{best}(t)}) \times v_i^d(t) + c_1 \times l_1 \times \\ &\quad (X_{pbestt_i}^d - x_i^d(t)) + c_2 \times l_2 \times \\ &\quad (X_{gbestt_i}^d - x_i^d(t)) + \alpha_i^d(t), \end{aligned} \quad (13)$$

式中: $l_1 = \text{rand}(0, 1)$, $l_2 = \text{rand}(0, 1)$; c_1, c_2 是正常数系数; $X_{pbestt_i}^d$ 是第 i 个粒子在第 d 维空间的历史最优值; $X_{gbestt_i}^d$ 是群体在第 d 维空间的历史最优值.

基于AGSA优化LS-SVM正规化参数和核参数, 具体步骤如下:

- 1) 初始化基于AGSA的LS-SVM模型的参数.
- 2) 将正规化参数和核参数种群用于LS-SVM水量预测模型, 得到训练样本点的预测值, 通过训练样本的预测值与实际值, 计算参数种群粒子适应度

$$f = \sqrt{\sum_{i=1}^l (S_{ireal} - S_{ipredict})^2 / l},$$

其中: S 表示训练样本集; S_{ireal} 表示第 i 个样本点的实际值, $S_{ipredict}$ 表示第 i 个样本点的预测值, l 表示训练样本数.

- 3) 根据循环次数更新 $g(t)$; 根据粒子适应度, 计算粒子的质量, 确定个体的最优位置和群体最优位置.

- 4) 根据式(11)更新粒子的加速度; 根据式(13)更新粒子速度, 根据式(4)更新粒子位置.

- 5) 判断是否满足约束条件. 若达到预设精度或设定的最大迭代次数, 则结束AGSA寻优过程, 得到最优解; 否则转至步骤2), 继续下一次寻优迭代.

4 算法测试(Algorithm test)

4.1 参数设置(Parameters setting)

利用常用的基本测试函数从理论上测试函数的性能; 同时比较AGSA与算法GSA, GA和PSO的收敛精度与收敛速度. 测试函数的形式如表1所示; 函数变量的维数, 取值范围以及最优解等参数设置如表2所示.

4.2 算法测试研究(Algorithm test research)

表3给出了各种算法的测试结果. 采用迭代次数CI, 最优解的平均值Mean以及标准差SD两个指标来评估算法的收敛性能. 算法的收敛速度利用平均优化时间 t_{run} 确定.

从表3可以看出, 针对表1中的5个基本测试函数, AGSA的最优解比GSA, GA, PSO的最优解更优; GSA与GA的收敛精度相差不大; PSO的收敛精度相对较差. 在收敛速度方面, 可以发现PSO的收敛速度最快, 而GSA的收敛速度相对很慢. 借鉴PSO的个体历史最优和群体历史最优策略改进GSA之后, 不仅算法的收敛速度大大加快, 而且收敛精度得到了进一步的提高.

表1 经典的基准测试函数

Table 1 The classical benchmark test functions

测试函数	最优值	最优解	变量范围
$f_1(\bar{x}) = \sum_{i=1}^n x_i^2$	0	$[0]^n$	$[-100, 100]^n$
$f_2(\bar{x}) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	0	$[0]^n$	$[-10, 10]^n$
$f_3(\bar{x}) = \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i) + 10)$	0	$[0]^n$	$[-5.12, 5.12]^n$
$f_4(\bar{x}) = \frac{1}{4000} \sum_{i=1}^n -x_i^2 - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}}) + 1$	0	$[0]^n$	$[-600, 600]^n$

表2 参数设置

Table 2 The setting of parameters

算法	群规模	迭代次数	变量维数	其他
PSO	20	1000	10, 30, 50, 100	$\omega = 0.7, c_1 = 1.5, c_2 = 1.5$
GA	20	1000	10, 30, 50, 100	$GAP = 0.95, pr = 0.7, pm = 0.1$
GSA	20	1000	10, 30, 50, 100	$G_0 = 100, \beta = 0.8$
AGSA	20	1000	10, 30, 50, 100	$G_0 = 100, \beta = 0.8, c_1 = 1.5, c_2 = 1.5$

表3 4种算法的优化结果比较

Table 3 Comparison of optimization results of four algorithms

函数	f_1				f_2				f_3				f_4				
	10	30	50	100	10	30	50	100	10	30	50	100	10	30	50	100	
PSO	CI	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	Mean	1.97	5.74	6.93	4.42	0.58	-0.65	0.36	0.52	0.20	-0.29	-0.54	-0.40	13.4	-34.2	22.2	6.37
	SD	15.0	20.6	20.9	23.7	1.21	3.74	2.51	2.55	1.11	1.47	1.78	1.84	157	116	148	11
	t_{run}	0.03	0.02	0.02	0.02	0.09	0.11	0.13	0.18	0.09	0.12	0.14	0.21	0.11	0.16	0.21	0.42
GA	CI	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	Mean	0.20	0.50	0.58	0.57	0.30	0.43	0.46	0.49	0.10	0.50	0.56	-0.48	0.40	0.43	0.46	0.48
	SD	4.75	5.78	4.97	15.5	5.77	6.89	11.5	10.8	1.35	0.50	4.96	4.99	5.99	7.11	4.95	6.23
	t_{run}	1.67	3.48	5.19	9.84	1.86	4.15	6.62	13.1	1.87	4.16	6.51	13.1	1.91	4.28	6.91	13.6
GSA	CI	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	Mean	-1.65	2.82	4.82	-3.88	0.29	0.65	0.54	0.44	0.53	0.16	0.55	-0.17	-0.47	-0.43	3.36	-2.58
	SD	13.4	14.9	24.9	25.2	5.47	5.34	5.68	5.23	1.54	2.56	2.82	2.70	4.99	5.01	15.45	5.95
	t_{run}	2.18	6.01	10.2	20.0	2.28	6.31	10.3	20.3	2.29	6.34	10.4	20.7	2.37	6.48	10.5	20.5
AGSA	CI	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	Mean	0.05	-0.03	-0.06	0.12	0.30	0.05	0.19	0.09	0.08	0.43	-0.08	0.15	0.01	0.02	0.01	-0.06
	SD	0.49	0.80	1.12	1.30	1.12	1.22	1.29	1.47	1.33	1.19	1.47	1.64	0.11	0.11	0.11	1.27
	t_{run}	1.72	3.77	7.71	13.1	1.59	4.01	7.85	13.2	1.54	4.80	7.91	14.5	0.12	0.38	0.54	13.0

5 试验研究(Experimental research)

5.1 模型输入变量(Input variables in mode)

选取2012年上海市城市水网某支线的数据. 考虑需水量的周期性特征, 随机选取2012年5天的每小时级流量数据进行比较. 从图1可以看出, 曲线的变化趋势整体上较为明显, 每天有一个用水高峰和一个用水低峰, 高峰期为凌晨8:00至晚上10:00, 上海市时用水量具有较强的日周期性. 考虑流量的自相关性, 计算时用水量序列的自相关系数. 相关性的

计算公式为

$$r_k = \frac{\sum_{t=1}^{N-k} (Q_t - \bar{Q})(Q_{t+k} - \bar{Q})}{\sum_{t=1}^N (Q_t - \bar{Q})^2}, \quad (14)$$

式中:

$$\bar{Q} = \frac{1}{N} \sum_{t=1}^N Q_t, \quad k = 1, 2, \dots, N - 1.$$

设定当 $|r_k| \in (0.75, 1)$ 时属于高度相关.

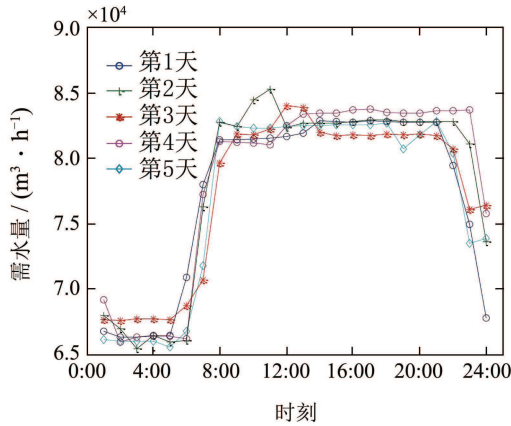


图 1 日流量变化趋势

Fig. 1 The trend of daily flow

选取2012年全年的时用水量数据作自相关性分析, 计算 $r = 1\ 264$ 的 r_k . 图2是进行流量自相关性分析的结果, 选取 $|r_k| > 0.75$ 的历史流量作为模型的输入. 通过计算得到

$$\begin{aligned} r_1 &= 0.875, r_{23} = 0.768, r_{24} = 0.821, \\ r_{25} &= 0.766, r_{48} = 0.798, r_{72} = 0.783, \\ r_{96} &= 0.77, r_{120} = 0.771, \\ r_{144} &= 0.764, r_{168} = 0.760. \end{aligned}$$

本文将预测时刻前 1 h, 23 h, 24 h, 25 h, 48 h, 72 h, 96 h, 120 h, 144 h, 168 h 的时用水量纳入模型的输入变量中.

城市短期需水量通常会受到气象(温度、降水量等)及节假日的影响, 这些影响不能规律性地表现在历史数据序列中. 本文选取温度作为气象对时用水量预测的影响. 分析出最高温度与流量的相关性为 0.75, 最低温度与流量的相关性为 0.71, 平均温度

与流量的相关性为 0.85, 因此选取平均温度作为模型的输入. 检修, 异常事故等对流量的影响不纳入模型的输入变量中, 作为模型的后期修正.

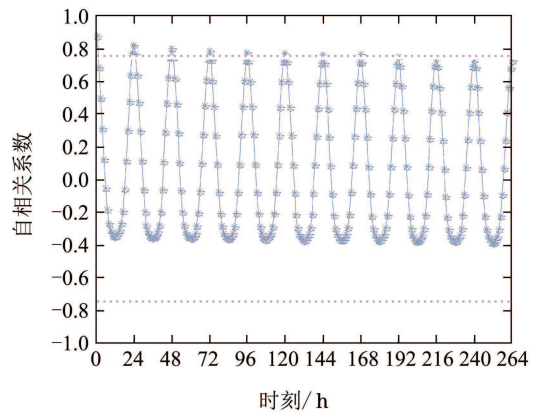


图 2 流量的自相关性分析

Fig. 2 Autocorrelation analysis of flow

综上所述, 预测某日 t 时刻的时用水量, 时用水量预测模型如下:

$$Q_t = f(Q_{t-1}, Q_{t-23}, Q_{t-24}, Q_{t-25}, \dots, Q_{t-168}, T_{\text{mean}}),$$

式中: Q_t 为 t 时刻的需水量预测值; $Q_{t-1}, Q_{t-23}, Q_{t-24}, Q_{t-25}, \dots, Q_{t-168}$ 为 $t-1, t-23, t-24, t-25, \dots, t-168$ 时刻的时用水量; T_{mean} 为小时级预报平均温度.

5.2 时用水量预测模型参数设置 (Parameters setting of hourly water demand forecasting)

通过算法的理论测试和对时用水量预测模型算法的反复调试, 使各算法的模型参数达到较优的状态. 各参数设置如表 4 所示.

表 4 模型参数设置

Table 4 Parameters setting in model

算法	群规模	迭代次数	变量维数	其他
PSO	20	1000	10, 30, 50, 100	$\omega = 0.7, c_1 = 1.5, c_2 = 1.5$
GA	20	1000	10, 30, 50, 100	$GAP = 0.95, pr = 0.7, pm = 0.1$
GSA	20	1000	10, 30, 50, 100	$G_0 = 100, \beta = 0.8$
AGSA	20	1000	10, 30, 50, 100	$G_0 = 100, \beta = 0.8, c_1 = 1.5, c_2 = 1.5$
预测算法	样本数	预测点数	输入变量维数	正规化参数及核参数
LS-SVM	300	24	13	$\gamma \in [0.01, 1000], \sigma \in [0.01, 100]$

5.3 实验结果分析 (Analysis of experimental results)

本文分别采用 AGSA, GSA, GA 和 PSO 预测支线 2012 年 10 月 9 日的 24 个小时级流量. 预测结果如图 3-4 所示.

从图 3-4 可以看出, 基于 AGSA 的 LS-SVM 模型具有最好的回归精度, 预测结果的均方根误差 MSE 为 1.22×10^6 , 平均相对误差为 0.58%; 而基于 GSA, GA 和 PSO 的 LS-SVM 模型的预测结果的均方根误差

差及平均相对误差分别为 2.06×10^6 , 2.10×10^6 和 3.76×10^6 ; 0.92%, 0.94%和1.74%。利用AGSA优化LS-SVM的参数可以得到最好的预测精度。因此AGSA适用于城市短期需水量预测问题。

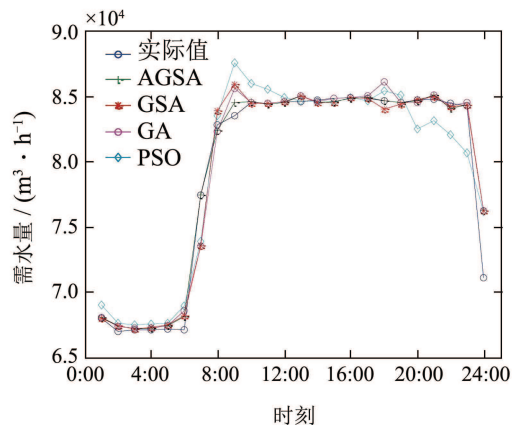


图3 各种算法的真实值与预测值比较

Fig. 3 Comparisons of real and predicted values in various algorithms

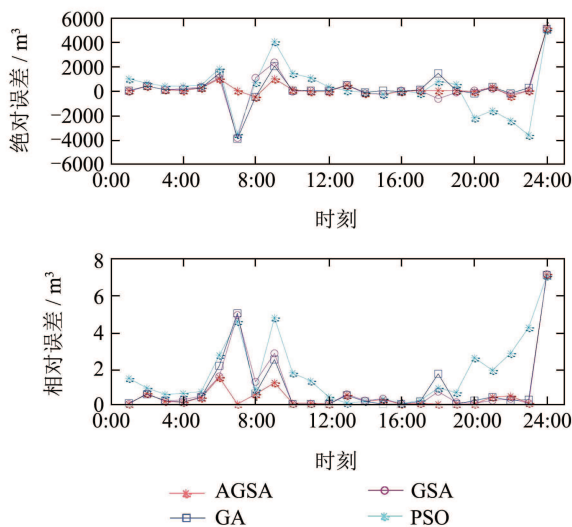


图4 多种算法的误差分析

Fig. 4 Error analysis in various algorithms

6 结论(Conclusions)

本文介绍了采用LS-SVM建立城市短期需水量预测模型,同时利用AGSA对LS-SVM的参数进行了训练以改进模型的预测精度。为了验证AGSA的有效性,将其与GSA, GA和PSO进行了比较,结果显示,基于AGSA的LS-SVM模型具有最好的预测

效果,但在过渡时期,影响因素较多且不易分析,流量变化毫无规律性,某些单点预测效果还不是很理想,后期须修正预测模型,进一步提高单点预测精度。通过实例分析表明,AGSA适用于多领域的参数优化。

参考文献(References):

- [1] HERRERA M, TORGO L, IZQUIERDO J, et al. Predictive models for forecasting hourly urban water demand [J]. *Hydrology*, 2010, 387(1): 141 – 150.
- [2] FENG K, ZHANG Y L. Water consumption forecasting using BP neural based on factor analysis in Chengdu [J]. *Water Resources Research*, 2011, 32(2): 8 – 11.
- [3] CUTORE P, CAMPISANO Y, KAPELAN Z, et al. Probabilistic prediction of urban water consumption using the SCEM-UA algorithm [J]. *Urban Water Journal*, 2008, 5(2): 125 – 132.
- [4] NISHA M G, PILLAI G N. Nonlinear model predictive control with relevance vector regression and particle swarm optimization [J]. *Control Theory & Applications*, 2013, 11(4): 563 – 569.
- [5] 刘建伟, 黎海恩, 刘瑗, 等. 迭代再权共轭梯度 q 范数正规化线性最小二乘支持向量机分类算法 [J]. *控制理论与应用*, 2014, 31(3): 334 – 342.
(LIU Jianwei, LI Haien, LIU Yuan, et al. Q -norm regularizing least-square-support-vector-machinelinear classifier algorithm via iterative reweighted conjugate gradient [J]. *Control Theory & Applications*, 2014, 31(3): 334 – 342.)
- [6] NIU D, WANG Y, WU D D. Power load forecasting using support vector machine and ant colony optimization [J]. *Expert Systems with Applications*, 2013, 37(3): 2531 – 2539.
- [7] RAO R V, SAVSANI V J, VAKHARIA D P. Teaching - learning-based optimization: A novel method for constrained mechanical design optimization problems [J]. *Computer-Aided Design*, 2011, 43(3): 303 – 315.
- [8] ZHANG W, NIU P, LI G, et al. Forecasting of turbine heat rate with online least squares support vector machine based on gravitational search algorithm [J]. *Knowledge-Based Systems*, 2013, 39: 34 – 44.

作者简介:

戴 钢 (1989-), 女, 硕士研究生, 从事供水管网系统的研究,

E-mail: j996268034@126.com;

王景成 (1972-), 男, 博士生导师, 从事实时控制系统的研究,

E-mail: jcwang@sytu.edu.cn;

葛 阳 (1986-), 男, 博士研究生, 从事网络控制系统的研究,

E-mail: geyang888@163.com;

刘华江 (1989-), 男, 硕士研究生, 从事工业过程的建模与控制研究, E-mail: 893042136@qq.com;

杨丽雯 (1990-), 女, 硕士研究生, 从事供水管网系统的研究,

E-mail: ylw0225@gmail.com.