

基于支持向量回归和小波包的供热负荷预测^{*}

沈阳建筑大学 黎展求[☆] 朱栋华
沈阳辰宇建筑有限责任公司 刘冬岩

摘要 通过分析影响热网负荷变化的各种因素,对热负荷数据进行预处理,运用小波包变换对负荷序列进行分解,对各子序列分别建立支持向量回归预测模型,最后通过序列重构,得出预测结果。仿真结果表明,该方法比传统 BP 神经网络和未作小波包分解的支持向量回归法具有更高的预测精度。

关键词 供热系统 热负荷预测 支持向量回归 小波包

Heating load prediction for heating systems based on support vector regression and wavelet packet

By Li Zhanqiu[★], Zhu Donghua and Liu Dongyan

Abstract Based on an analysis of the factors causing changes in the load of heat-supply network, preprocesses the data of the heating load, and decomposes the load sequences into different scales through the wavelet packet transform. Develops respective support vector regression predicting models for these sub-sequences. After reconstructing sequences, obtains the predicting results. Simulation results show that the method is superior in predicting accuracy to the traditional BP neural network and the support vector regression method without wavelet packet decomposition.

Keywords heating system, heating load prediction, support vector regression, wavelet packet

[★] Shenyang Jian Zhu University, Shenyang, China

①

0 引言

供热系统具有强非线性、大滞后、热惯性大、供热管网复杂等特点。供热负荷受供水流量、供回水温度、室外气温、光照、室外风速、节假日等多种因素影响,很难对其建立数理模型。实施供热计量后,热网的流量和热负荷将具有更大波动,必须及时获取用户用热量的变化信息,将负荷预报值作为控制对象,对供热过程进行跟踪调节控制,才能满足供热需求并节省能源。对供热参数进行及时准确的预测是提高系统调节性能、促使供热控制达到用户需求的重要前提。目前国内外采用的负荷预

测方法多为以下方法的独立或组合运用:时间序列、回归分析、神经网络、灰色系统模型、小波分析、粗糙集、支持向量机^[1-6]等。回归分析和时间序列法对历史数据要求高,预测误差较大;神经网络收敛速度慢,泛化能力差,训练容易过拟合;灰色系统模型在原始数据具有指数增长趋势时才具有较高的预测精度。根据各种方法的优缺点,组合多种方法进行研究和应用,可以获得更高的预测精度。

①[☆] 黎展求,女,1981年2月生,在读硕士研究生
110168 沈阳建筑大学 295 井
(024) 24694005 (O) 13709821607
E-mail: lizhanqiu@yahoo.com.cn
收稿日期:2006-08-28
修回日期:2007-01-10

^{*} 建设部研究开发项目(编号:06-K9-63)

Vapnik 等人根据统计学习理论提出的支持向量机器学习方法^[7], 广泛应用于模式识别和回归估计。将支持向量回归 (SVR, support vector regression) 与小波包分析结合, 对供热系统热负荷进行预测, 仿真结果表明该方法在小样本情况下有较好的预测效果。

1 小波包分析

小波分析是将信号分解成一系列小波函数的叠加, 而小波函数都是由一个母小波函数经过平移与尺度伸缩得来的^[8]。

母小波函数经过伸缩和平移后, 得到一个小波序列, 实际运用中通常采用离散小波序列。小波包分析能对信号进行更精细的局部分析, 是小波分析的进一步完善和发展, 对多分辨分析没有细分的高频部分进一步分解。

选定一个母小波函数, 将信号按小波包分解算法分解得到各个小波包子空间上的小波包系数。设平方可积的函数空间 $L^2(R)$ 存在一个子空间 U_j^n , 定义子空间 U_j^n 和 U_j^{2n} 分别是函数 $u_n(t)$ 和 $u_{2n}(t)$ 的闭包空间。给定的信号序列 $f(t) \in U_j^n$, $f(t)$ 表示为 $f(t) = \sum_l d_l^{j,n} u_n(2^j t - l)$, n 为采样点数 (任意非负整数), j 为尺度参数, l 为给定的分解层数, $j, l \in \mathbf{Z}$, $d_l^{j,n}$ 为小波包系数。对原始信号 $f(t)$ 进行小波包分解的结构如图 1 所示。小波包分解采用一组低通与高通共轭正交镜像滤波器组, 对低频信号与高频信号在通频范围内不同层次分解序列。

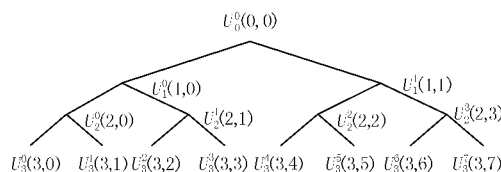


图 1 小波包分解树

当 $n=0$ 时

$$\left. \begin{aligned} u_0(t) &= \varphi(t) = \sqrt{2} \sum_{k \in \mathbf{Z}} h_k u_0(2t - k) \\ u_1(t) &= \psi(t) = \sqrt{2} \sum_{k \in \mathbf{Z}} g_k u_0(2t - k) \end{aligned} \right\} \quad (1)$$

式中 k 为位置参数; $\varphi(t)$ 为尺度函数; $\psi(t)$ 为小波基函数; h_k, g_k 是一组共轭镜像滤波器。

由尺度函数 $\varphi(t)$ 和小波基函数 $\psi(t)$ 构造的序列 $u_n(t)$ 为正交小波包。由离散小波包分解算法得

到小波包系数。由 $d_l^{j-1,n}$ 求 $d_l^{j,2n}$ 与 $d_l^{j,2n+1}$ 的分解算法为

$$\left. \begin{aligned} d_l^{j,2n} &= \sum_{k \in \mathbf{Z}} h_{k-2l} d_k^{j-1,n} \\ d_l^{j,2n+1} &= \sum_{k \in \mathbf{Z}} g_{k-2l} d_k^{j-1,n} \end{aligned} \right\} \quad (2)$$

式中 h_{k-2l} 为低通分解滤波器系数; g_{k-2l} 为高通分解滤波器系数。

在各个尺度 j 上利用小波包系数进行各个小波包子空间上的信号重构, 得到各个小波包空间上的信号分量。由 $d_l^{j,2n}$ 和 $d_l^{j,2n+1}$ 求 $d_l^{j-1,n}$ 的重构算法为

$$d_l^{j-1,n} = \sum_{k \in \mathbf{Z}} (\bar{h}_{l-2k} d_k^{j,2n} + \bar{g}_{l-2k} d_k^{j,2n+1}) \quad (3)$$

式中 $\bar{h}_{l-2k}$ 为低通重构滤波器系数; $\bar{g}_{l-2k}$ 为高通重构滤波器系数。

2 支持向量回归

支持向量机具有严格的理论基础, 利用了结构风险最小化原则^[9]使经验风险与置信范围 (该参数用来衡量函数集的复杂性) 之和最小, 通过非线性变换将低维空间的非线性问题转换为高维空间的线性问题, 能较好地解决小样本、非线性、高维数和局部极小点等实际问题。

给定训练样本集 $\{(x_i, y_i) | i=1, 2, \dots, s\}$, $x_i \in \mathbf{R}^d$, $y_i \in \mathbf{R}$, 其中 x_i 为 d 维输入向量, $d=1, 2, \dots, n$, y_i 为 x_i 对应的输出值, s 为样本数。支持向量回归实质是寻求能拟合训练样本的最优函数 $f(x)$ 。利用非线性映射 Φ 将数据 x_i 映射到多维特征空间 F , 线性回归函数 $f(x)$ 为 $\omega^T \cdot \Phi(x) + b$, 其中 ω 为权向量, $\omega \in \mathbf{R}^d$, b 为偏置量, $b \in \mathbf{R}$ 。

控制函数的复杂性可以通过使线性回归函数最平坦来实现, 即使 $\frac{1}{2} \|\omega\|^2$ 最小化。考虑求最优回归函数等价于使 $\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^s (\xi_i + \xi_i^*)$ 最小化 (其中 C 为误差惩罚系数, $C > 0$, 用来平衡函数复杂度和训练误差, 调节置信范围和经验风险的比例, 提高泛化能力; ξ_i, ξ_i^* 为非负的松弛变量, 分别表示损失函数参数 ϵ (回归允许的最大误差) 的上限误差和下限误差), 其约束条件为

$$\left. \begin{aligned} y_i - [\omega \Phi(x_i) + b] &\leq \epsilon + \xi_i \\ [\omega \Phi(x_i) + b] - y_i &\leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* &\geq 0, i = 1, \dots, s \end{aligned} \right\} \quad (4)$$

$\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^s (\xi_i + \xi_i^*)$ 中的第一项反映回归函数的复杂程度,第二项反映经验风险(训练误差)。损失函数参数 ϵ 为 Vapnik- ϵ 不敏感损失函数所定义的误差。

$$\max W(\alpha_i, \alpha_i^*) = -\epsilon \sum_{i=1}^s (\alpha_i + \alpha_i^*) + \sum_{i=1}^s y_i (\alpha_i - \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^s (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) (x_i \cdot x_j) \quad (5)$$

其中

$$\sum_{i=1}^s (\alpha_i - \alpha_i^*) = 0, 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, s \quad (6)$$

求出拉格朗日乘子 α_i 和 α_i^* 及 ω , 由 KKT (Karush-Kuhn-Tucker) 条件^[7] 计算出偏置 b , 得到回归函数:

$$f(x) = \sum_{i=1}^s (\alpha_i - \alpha_i^*) [\Phi(x) \cdot \Phi(x_i)] + b \quad (7)$$

α_i, α_i^* 不为 0 时对应的样本就是支持向量。引入满足 Mercer 条件^[7] 的核函数 $K(x, x_i)$, 定义 $K(x, x_i) = \Phi(x) \cdot \Phi(x_i)$, 将高维空间的内积运算转换为原空间的运算, 来拟合非线性函数映射 Φ , 得出回归函数:

$$f(x) = \sum_{i=1}^s (\alpha_i - \alpha_i^*) K(x, x_i) + b \quad (8)$$

常用的核函数有: 多项式核函数 $K(x, x_i) = (x \cdot x_i + 1)^d, d = 1, 2, \dots, n$; sigmoid 核函数 $K(x, x_i) = \tanh[c_1(x \cdot x_i) + c_2], c_1, c_2$ 为常数, $c_1 > 0, c_2 < 0$; 径向基核函数(RBF, radial basis function) $K(x, x_i) = \exp(-\gamma |x - x_i|^2)$, 其中 γ 为核函数参数, $\gamma = \frac{1}{2\sigma^2}$, σ^2 是高斯核宽度。

3 供热负荷预测模型

冬季供热系统的热负荷预测参考的因素^[10] 主要有供回水温度、供水流量、室外温度、室外风速、风向、太阳辐射。由于实际数据采集困难,且考虑过多因素会使预测模型更复杂,因此本文主要依据室外温度、供回水温度、供水流量、热负荷的历史数据来对热负荷、供水温度、回水温度作预测。以沈阳市某住宅小区 2005—2006 年部分供热数据为依据,根据用户 24 h 内用热量的变化规律,对供水温度、回水温度、热负荷分别建立预测模型,输入参数均为供水温度 t_g (°C)、回水温度 t_h (°C)、室外温度 t_w (°C)、供水流量 m_v (t/h)、热负荷 Q (GJ/h) 和节假日影响因素 h_0 。节假日用户的用热量增加,调

引入拉格朗日乘子 α_i, α_i^* 构造拉格朗日泛函,根据拉格朗日函数满足的极值条件,得到 $\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^s (\xi_i + \xi_i^*)$ 优化问题的对偶形式:

节增多,因此,设定节假日 h_0 为 1,工作日 h_0 为 0。模型输入数据为 3 月 1~25 日供热数据,采样时间为 1 h,选择 RBF 进行训练,训练样本为 500,测试样本为 100。图 2 为预测算法原理图。

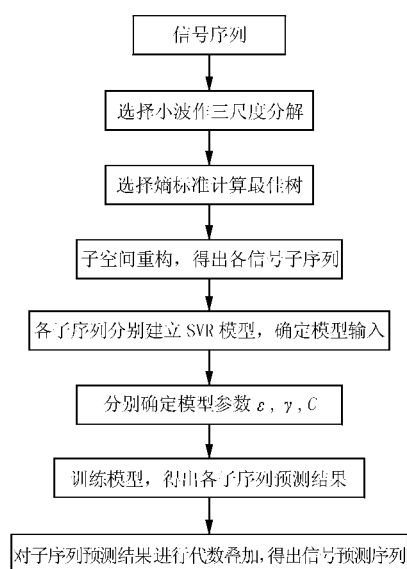


图 2 预测算法原理图

3.1 小波包分别对供水温度、回水温度、供热负荷分解

对全部数据进行归一化处理,使之在 $[-1, 1]$ 之间。通过对常用小波系数的对比,选择 coif5 小波分别对供水温度、回水温度、热负荷进行三尺度金字塔小波包分解。按照 Shannon 熵标准计算最佳树,根据在各尺度上得到的小波包系数,进行各小波包子空间信号重构,得出各子序列。3 个预测量的重构子序列均为 (3,0), (3,1), (2,1), (1,1) 序列。图 3 为供水温度序列及其小波包分解的子序列。

3.2 SVR 预测

由 $(\tau-1), (\tau-2), \dots$ 时刻的历史数据预测 τ 时刻的值。核函数参数 γ 影响数据在特征空间分布的复杂程度,误差惩罚系数 C 通过调整经验误差提高模型的推广能力。结合 LIBSVM 工具箱进

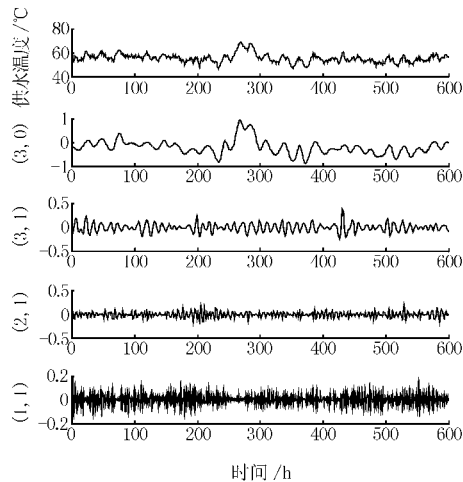


图 3 供水温度序列及其小波包分解的子序列

表 1 供水温度各子序列预测模型

	(3,0)	(3,1)	(2,1)	(1,1)
输入量数目	14	17	17	17
输入参数	$t_g(\tau-1), t_g(\tau-2), t_g(\tau-3), t_g(\tau-25), t_g(\tau-26),$ $t_g(\tau-27), t_g(\tau-49), t_g(\tau-50), t_g(\tau-51), t_w(\tau-1), t_h(\tau-1), Q(\tau-1), m_v(\tau-1), h_0(\tau-1)$	$t_g(\tau-1), t_g(\tau-2), t_g(\tau-3), t_g(\tau-4), t_g(\tau-5), t_g(\tau-6),$ $t_g(\tau-7), t_g(\tau-8), t_g(\tau-9), t_g(\tau-10), t_g(\tau-11), t_g(\tau-12), t_w(\tau-1), t_h(\tau-1), Q(\tau-1), m_v(\tau-1), h_0(\tau-1)$		

模,即表 1 中的输入参数数目相同,但模型输入参数的取值范围需要作出调整。以表 1 为参考,预测回水温度时, t_g 均替换为 t_h , t_h 替换为 t_g ;同理,预测热负荷时, t_g 替换为 Q , Q 替换为 t_g 。供水温度、回水温度、热负荷的预测结果如图 4 所示。

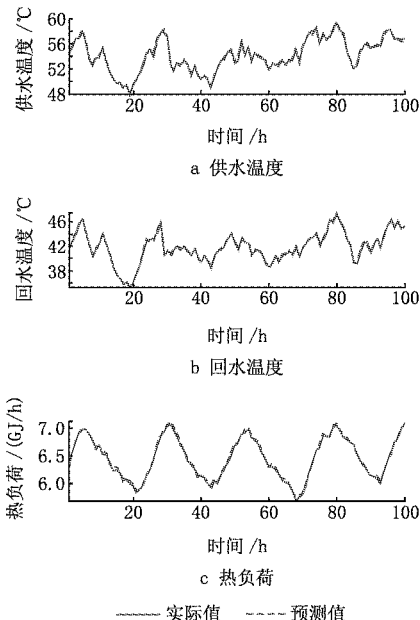


图 4 供水温度、回水温度、热负荷的预测结果与实际值的对比

为检验本文提出的方法(SVR 小波包),将其与未作小波包分解的 SVR 方法和 BP 神经网络进

行仿真,以径向基函数为 SVR 核函数。以供水温度为例,为简化模型,各子序列模型参数选择一致:损失函数参数 $\epsilon=0.001$,误差惩罚系数 $C=100$,径向基核函数参数 $\gamma=0.01$ 。其中,对子序列(3,0)单独建模,由于其变化反映了供水温度的变化趋势,有一定的规律性,波动小,以 3 d 内的数据为参考;其他子序列在 24 h 内波动大,缺乏规律性,为避免输入量过于复杂,以 $(\tau-12) \sim (\tau-1)$ 时刻的供水温度作输入,其他室外温度、回水温度、热负荷、供水流量的选择见表 1。

对各子序列分别建立 SVR 预测模型,预测结果进行代数叠加,作反归一化处理,得出供水温度。采用类似的 SVR 模型结构对回水温度、热负荷建

行了对比。SVR 模型输入与表 1 中(3,0)子序列相同,参数选择不改变。BP 模型为 14-8-1 结构,采用梯度下降/自适应学习率算法训练网络,学习率 0.01,学习率的增长比率 1.05,动量因数 0.9,训练误差 0.001,最大训练步数 5 000。以供水温度为例,比较结果见图 5,可以看出,采用 SVR 小波包有较好的拟合效果,在模型逼近和预测上都具有较高的精度。

表 2 给出了 3 种方法预测结果误差的对比。可以很明显看出,在小样本情况下,采用 SVR 后训练速度均显著提高。本文提出的方法,相比单独的 SVR,BP 神经网络,拟合精度高,预测跟踪能力强。神经网络以经验风险最小化为原则,网络训练结果随初始权值的不同具有不确定性,易收敛于局部极小点。

4 结论

结合小波包和支持向量回归对供热负荷进行预测,在小样本情况下,具有适应性强、训练耗时间短、泛化性能好等优点,克服了神经网络算法局部极小、对统计样本依赖较大的问题等。仿真实验表明本文提出的方法有良好的预测精度和较强的适应能力,可为供热系统进一步调节控制供热参数,使供热控制及时跟踪水力热力的变化,合理分配供热负荷提供必要依据。

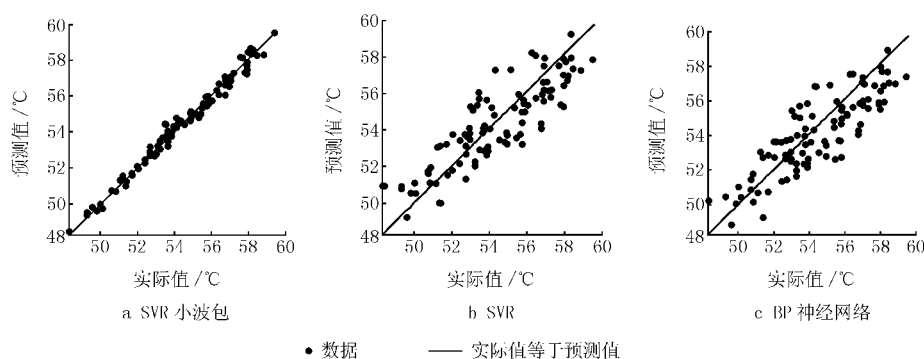


图 5 SVR 小波包、SVR 及 BP 神经网络的供水温度预测对比

表 2 误差分析对比

	供水温度		回水温度		热负荷		3 个量预测 总耗时/s
	相对平均误差/%	相关系数	相对平均误差/%	相关系数	相对平均误差/%	相关系数	
SVR 小波包	0.41	0.993 9	0.55	0.994 0	0.45	0.998 3	11.98
SVR	2.03	0.856 5	2.64	0.843 2	1.35	0.976 7	1.86
BP 神经网络	2.11	0.859 1	2.78	0.828 5	1.47	0.975 9	144.14

注:相对平均误差 e_M 计算式为 $e_M = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - P_i}{A_i} \right| \times 100\%$, 其中 n 为采样个数, A_i 为实际值, P_i 为预测值。

参考文献

- [1] 齐维贵,朱学莉. 时间序列预报法在供热控制中的应用研究[J]. 电子学报, 2003, 31(2):268-270
- [2] Kawashima M, Dorgan C E, Mitchell J W. Hourly thermal load prediction for the next 24 hours by Arima, Ewma, L R and an artificial neural network [G]//ASHRAE Trans, 1995, 101 (1):186-200
- [3] Dotzauer E. Simple model for prediction of load in district-heating system[J]. Applied Energy, 2002, 73 (3/4):277-284
- [4] 胡文斌,华贵,杨昌智,等. 灰色理论在城市供热负荷预测中的应用[J]. 煤气与热力, 2002, 22(1):28-31
- [5] 马涛,徐向东. 基于多尺度挖掘的区域供热系统负荷预测[J]. 暖通空调, 2005, 35(11):16-19
- [6] Hou Zhijian, Lian Zhiwei, Yao Ye, et al. Cooling load prediction based on the combination of rough set theory and support vector machine[J]. HVAC&R Research, 2006, 12(2):337-352
- [7] Vapnik V N. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1995
- [8] 廖剑利,张建,肖俊,等. 小波分析理论与 MATLAB7 实现[M]. 北京:电子工业出版社, 2005
- [9] 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
- [10] 涂光备. 供热计量技术[M]. 北京:中国建筑工业出版社, 2003
- [11] Sauer H J, Howell R H. Promise and potential of air-air energy recovery systems [G] // ASHRAE Trans, 1981, 87(1): 167-182
- [12] Philip B. Makeup air heat recovery saving energy in labs[J]. ASHRAE J, 2004, 46(2):35-40
- [13] Zhang Yinping, Jiang Yi, Zhang Lizhi, et al. Analysis of thermal performance and energy saving of membrane based heat recovery ventilator [J]. Energy, 2000, 25(6):515-527
- [14] Zhang L Z, Niu J L. Energy requirements for conditioning fresh air and the long term savings with a membrane-based energy recovery ventilator in Hong Kong[J]. Energy, 2001, 26(2):119-135
- [15] Renato M L, Andrea G. Technical and economical analysis of heat recovery in building ventilation systems[J]. Applied Thermal Engineering, 1998, 18 (1/2):47-67
- [16] 建设部工程质量安全监督与行业发展司,中国建筑标准设计研究所. 全国民用建筑工程设计技术措施 暖通空调·动力[M]. 北京:中国计划出版社, 2003
- [17] 中国有色金属工业总公司. GBJ 19-87 采暖通风与空气调节设计规范[S]. 2 版. 北京:中国计划出版社, 2001
- [18] 埃克尔特 E R G. 传热与传质[M]. 徐明德,译. 北京:科学出版社, 1963
- [19] 连之伟,张寅平,陈宝明,等. 热质交换原理与设备 [M]. 北京:中国建筑工业出版社, 2001
- [20] 孙一坚. 工业通风[M]. 3 版. 北京:中国建筑工业出版社, 1994
- [21] 日本冷冻空调工业会. JISB 8628:2003 全热交换器 [S]. 东京:日本规格协会, 2003

(上接第 67 页)