



基于 KPCA-WLSSVM 的建筑能耗预测模型

赵超,戴坤成,王贵评

(福州大学节能技术研究中心,福州 350108)

摘要:为降低建筑能耗影响因素间复杂相关性对模型性能的影响,建立了一种基于 KPCA-WLSSVM 的建筑能耗预测模型。利用核主元分析(KPCA)对输入变量进行数据压缩,消除变量之间的相关性,简化模型结构;进一步采用加权最小二乘支持向量机(WLSSVM)方法建立建筑能耗预测模型,同时结合一种新型混沌粒子群-模拟退火混合优化(CPSO-SA)算法对模型参数进行优化,以提高模型的预测性能及泛化能力。通过将 KPCA-WLSSVM 模型方法应用于某办公建筑能耗的预测中,并与 WLSSVM、LSSVM 及 RBFNN 模型相比,实验结果表明,KPCA-WLSSVM 模型方法能有效提高建筑能耗预测精度。

关键词:建筑能耗;预测;核主元分析;支持向量机

中图分类号:TU831 **文献标志码:**A **文章编号:**1674-4764(2015)05-0109-07

A prediction model for energy consumption of building based on KPCA-WLSSVM

Zhao Chao, Dai Kuncheng, Wang Guiping

(Research Center of Energy Saving Technology, Fuzhou University, Fuzhou 350108, P. R. China)

Abstract: The correlations among the building energy consumption factors can corrupt the prediction model's performance, and get undesirable results. A prediction model based on KPCA-WLSSVM is proposed to forecast building energy consumption. The kernel principal component analysis (KPCA) method could not only solve the linear correlation of the input and compress data but also simplify the model structure. A novel hybrid chaos particle swarm optimization simulated annealing (CPSO-SA) algorithm is applied to optimize WLSSVM parameters to improve learning performance and generalization ability of the model. Furthermore, the KPCA-WLSSVM model is applied to the energy consumption prediction for an office building, and the results show that the KPCA-WLSSVM has better accuracy compared with WLSSVM model, LSSVM model and RBF neural network model, and the KPCA-WLSSVM is effective for building energy consumption prediction.

Key words: energy consumption of building; forecasting; kernel principal component analysis; support vector machines

收稿日期:2015-06-16

基金项目:国家自然科学基金(6080402、61374133);高校博士点专项科研基金(20133314120004)

作者简介:赵超(1976-),男,副教授,主要从事最优控制研究,(E-mail) seasky76@163.com。

Received:2015-06-16

Foundation item: National Natural Science Foundation of China (No. 6080402, 61374133); Research Fund for the Doctoral Program of Higher Education of China (No. 20133314120004)

Author brief: Zhao Chao (1976-), associate professor, main research interest: optimal control, (E-mail) seasky76@163.com.

随着建筑能耗在能源消耗中所占比重的快速增长,建筑节能受到越来越多的关注,准确有效地预测建筑能耗是开展节能评估工作的重要基础,也是进行建筑节能优化设计的主要依据。由于建筑能耗系统涉及到机理复杂的热湿传递过程,是一个典型的多变量、非线性、强耦合和多扰动的复杂系统,采用机理方法建立建筑能耗预测模型十分困难。人工神经网络由于对非线性函数具有很强的拟合和逼近能力,作为一种有效的非线性建模方法,目前已有许多学者将其成功应用于建筑能耗预测领域^[1-4]。但由于神经网络训练基于经验风险最小化原则,易产生过学习导致泛化能力下降,另外还存在诸如网络结构难以确定和过分依赖大样本学习等缺陷,因而在建筑能耗预测领域还难以大范围推广。

支持向量机(Support Vector Machines, SVM)是 Vapnik 等^[5-6]基于统计学习理论提出的一种新的有监督机器学习方法。支持向量机利用结构最小化原理来提高其泛化能力,较好地解决了小样本、非线性、高维数、局部极小等实际问题,目前已成为建筑能耗预测领域的研究热点^[7-8]。最小二乘支持向量机(Least Squares Support Vector Machines, LS-SVM)是标准 SVM 的扩展,LS-SVM 模型在优化目标函数中选取了误差(松弛变量)二范数的形式,将 SVM 中需要求解的二次规划问题转化为线性方程组求解问题,从而降低了计算复杂度,有效提高了学习速度,较好地解决了大样本数据情况下 SVM 计算复杂的问题^[9]。文献^[10-12]基于 LSSVM 回归方法建立了建筑能耗预测模型,并获得了较好的预测结果。针对 LSSVM 算法鲁棒性能欠佳的问题,Suykens 等^[13]提出了加权最小二乘支持向量机算法(Weighted Least Squares Support Vector Machines, WLSSVM)。WLSSVM 在建模过程中依据训练样本的重要性分别赋予其不同权重,从而克服了奇异点对 LSSVM 鲁棒性能的影响。文献^[14]基于该算法建立了建筑能耗预测模型。尽管支持向量机回归方法能够有效处理高维数、非线性数据建模问题,但是建筑能耗受到气候条件、建筑特征参数、围护结构热工性能、室内热扰、新风负荷、室内环境参数、空调运行模式等众多因素的影响,且各影响因素之间严重相关,不可避免存在一定的噪声和冗余信息,在建模过程中会降低精度,增加计算复杂度^[15-16]。因此,有必要对能耗建模样本数据进行数据压缩和特征提取,通过降维处理来去除样本数据

集中的冗余信息以及消除噪声,以提高模型的预测精度和泛化能力。核主元分析法(KPCA)作为一种数据压缩和特征提取的非线性技术,已在数据处理中的到了广泛应用^[17-18]。KPCA 采用核函数的方法,把过程数据或变量组成的原始数据空间映射到高维特征空间,在高维特征空间使用线性 PCA 方法来寻找主成分。特征空间中的主元变量则保留原始变量的特征信息而消除冗余信息,它是解决数据相关问题,提高模型运算速度的重要工具。

为了提高建筑能耗预测模型的精度和可靠性,笔者提出了一种基于 KPCA 和 WLSSVM 的建筑能耗预测方法。首先采用 KPCA 方法对输入变量进行特征提取,有效消除变量之间的相关性,降低输入变量的维数;将通过特征提取后的数据集作为输入变量建立 WLSSVM 建筑能耗模型;考虑到 LSSVM 的正则化参数和核宽参数对模型拟合精度和泛化能力都有较大影响,利用混沌粒子群-模拟退火混合优化(CPSO-SA)算法对 LSSVM 的参数进行优化选择,从而建立起具有较高预测精度和泛化能力的建筑能耗预测模型。基于 DEST 仿真平台的模拟数据,对南方地区某办公建筑的能耗进行预测,并与 WLSSVM、LSSVM 和 RBF 神经网络预测模型的性能进行比较,结果表明,KPCA-WLSSVM 预测模型的性能优于另外 3 种预测模型。

1 KPCA-WLSSVM 模型算法

1.1 核主元分析

KPCA 方法利用非线性核函数将原空间的样本数据映射到线性高维特征空间,即通过非线性映射,将原空间 $x_i (i = 1, 2, \dots, N)$ 映射到某个高维特征空间,并在这个特征空间实现 PCA,高维特征空间的协方差矩阵 \mathbf{R} 为

$$\mathbf{R} = \frac{1}{N} \sum_{i=1}^N \varphi(x_i) \cdot \varphi(x_i)^T \quad (1)$$

求协方差矩阵 \mathbf{R} 的特征值 λ_i 和特征向量 \mathbf{u}_i 有

$$\mathbf{R} \mathbf{u}_i = \lambda_i \mathbf{u}_i \quad (2)$$

特征向量 \mathbf{u}_i 可以表示为 $\varphi(x)$ 的线性组合

$$\mathbf{u}_i = \sum_{j=1}^N a_j \varphi(x_j) \quad (3)$$

式(1)两边左乘 $\varphi(x_k)$, 整理得

$$\frac{1}{N} \sum_{i=1}^N \left\{ [\varphi(x_k)^T \cdot \varphi(x_i)] \cdot \sum_{j=1}^N a_j [\varphi(x_i)^T \cdot \varphi(x_j)] \right\} = \lambda_i \sum_{j=1}^N a_j [\varphi(x_k)^T \cdot \varphi(x_j)] \quad (4)$$

式中: $k = 1, 2, \dots, N$ 。定义一个 $N \times N$ 维的核函数矩阵 \mathbf{K}

$$\mathbf{K}_{ij} = \mathbf{K}(x_i, x_j) = \varphi(x_i)^\top \cdot \varphi(x_j) \quad (5)$$

式(4)进一步化简为

$$N\lambda_i a = \mathbf{K}a \quad (6)$$

式中: $a = (a_1, a_2, \dots, a_N)^\top$; 提取样本数据 x 映射后的第 t 非线性主元 p_t

$$p_t = \mathbf{u}_t^\top \cdot \varphi(x) = \sum_{j=1}^N a_j \varphi(x_j)^\top \cdot \varphi(x) = \sum_{j=1}^N a_j \mathbf{K}(x_j, x) \quad (7)$$

1.2 加权最小二乘支持向量机

由于标准 SVM 模型需要求解二次规划问题, 在大样本数据的情况下, 其运算量大, 所需支持向量的个数较多, 因此模型比较复杂, 从而导致模型的推广能力较差。Suykens 等在标准 SVM 基础上提出了最小二乘支持向量机模型。相对于支持向量机模型, LS-SVM 模型中的约束是等式约束, 在优化目标函数中选取了误差平方和的形式, 将原来 SVM 中需要求解的二次规划问题通过最优性条件转化为线性方程组求解问题, 从而降低了计算复杂度, 在非线性和系统建模方面更具有优势。其优化问题为^[19]

$$\min \frac{1}{2} \boldsymbol{\omega}^\top \boldsymbol{\omega} + \frac{1}{2} C \sum_{i=1}^N \xi_i^2 \quad (8)$$

$$\text{s. t.} \quad y_i = \boldsymbol{\omega}^\top \cdot \varphi(x_i) + b + \xi_i \quad (9)$$

式中: $\boldsymbol{\omega}$ 为权系数向量; $\varphi(\cdot)$ 为输入空间到高维空间的映射; C 为正则化参数; b 为阈值。

尽管 LS-SVM 模型在解决了 SVM 计算复杂度方面体现了优势, 但是由于在 LS-SVM 模型中目标函数采用了误差均方和的形式, 其对应的 Lagrange 乘子与误差成比例从而导致了该模型丢失了稀疏性, 从而降低了模型的鲁棒性。Suykens 等在 LS-SVM 算法的基础上提出了加权最小二乘支持向量机算法(WLS-SVM), 该方法对式(8)中误差的平方 ξ_i^2 进行加权处理, 设 ξ_i^2 的权值为 v_i , 则对应的优化问题可以描述为

$$\min \frac{1}{2} \boldsymbol{\omega}^\top \boldsymbol{\omega} + \frac{1}{2} C \sum_{i=1}^N v_i \xi_i^2 \quad (10)$$

$$\text{s. t.} \quad y_i = \boldsymbol{\omega}^\top \cdot \varphi(x_i) + b + \xi_i \quad (11)$$

引入 Lagrange 函数

$$L(\boldsymbol{\omega}, b, \boldsymbol{\xi}, \boldsymbol{\alpha}^*) = \frac{1}{2} \boldsymbol{\omega}^\top \boldsymbol{\omega} + \frac{1}{2} C \sum_{i=1}^N v_i \xi_i^2 - \sum_{i=1}^N \alpha_i^* [\boldsymbol{\omega}^\top \varphi(x_i) + b + \xi_i - y_i] \quad (12)$$

式中: α_i^* ($i = 1, 2, \dots, N$) 为 Lagrange 乘子, 根据优化条件, 引入核函数

$$\mathbf{K}(x_i, x_j) = \varphi(x_i)^\top \cdot \varphi(x_j) \quad (13)$$

则式(10)的优化问题变为式(14)线性方程组的求解问题

$$\begin{bmatrix} 0 & \mathbf{I}_{1 \times N} \\ \mathbf{I}_{N \times 1} & \mathbf{R} + \frac{1}{C} \mathbf{V} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha}^* \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \quad (14)$$

式中: $\mathbf{I}_{1 \times N}$ 是 $1 \times N$ 的单位行向量; $\mathbf{I}_{N \times 1}$ 是 $N \times 1$ 的单位列向量; $\mathbf{V} = \text{diag}(v_1^{-1}, v_2^{-1}, \dots, v_N^{-1})$; $\mathbf{R} = \{\mathbf{K}(x_i, x_j) \mid i, j = 1, 2, \dots, N\}$; $\boldsymbol{\alpha}^* = [\alpha_1^*, \alpha_2^*, \dots, \alpha_N^*]^\top$; $\mathbf{y} = [y_1, y_2, \dots, y_N]^\top$; 由此可得回归函数的形式

$$\mathbf{y} = \sum_{i=1}^N \alpha_i^* \mathbf{K}(x_i, x) + b \quad (15)$$

1.3 混沌粒子群-模拟退火混合优化算法

标准 PSO 算法以其简单易行及快速收敛的特性在工程优化领域得到了快速发展, 但也存在容易陷入局部最优的问题, 且搜索速度有待进一步提高。笔者提出一种混沌粒子群-模拟退火(CPSO-SA)混合算法用于搜索最佳模型参数, 该算法结合了 CPSO 算法的群体多样性和 SA 算法的渐进收敛性的优点, 并保持全局搜索和局部搜索的平衡, 可以有效提高算法的搜索效率, 从而加快算法的收敛速度。

基于 CPSO-SA 的参数优化算法流程为:

1) 输入训练样本集 $\{(x_i, y_i), i = 1, 2, \dots, N\}$; 设置 CPSO-SA 算法的控制参数: 种群规模 m 、最大迭代次数 t_{\max} 、初始温度 $T = T_0$; 随机产生 m 个粒子 $\{p_1, p_2, \dots, p_m\}$, 初始化粒子的位置值和速度值, 当前迭代次数设置为 $t = 0$;

2) 计算每个粒子适应度值 $f(p_i)$;

3) 比较当前粒子个体的适应度值, 找出当前最优个体 p_{best} , 按式(16)的 Logistic 方程对 p_{best} 进行 M 次迭代, 得到混沌变量可行解序列 $\{p_{\text{best}}^{(1)}, p_{\text{best}}^{(2)}, \dots, p_{\text{best}}^{(M)}\}$;

$$p^{(n+1)} = \lambda p^{(n)} (1 - p^{(n)}) \quad (16)$$

式中: λ 为控制参数; n 为迭代次数;

4) 找出可行解序列中的最优个体, 记作 \hat{p}_{best} , 计算 Δf

$$\Delta f = f(\hat{p}_{\text{best}}) - f(p_{\text{BEST}}) \quad (17)$$

式中: p_{BEST} 为全局最优个体。若 $\Delta f < 0$, 则更新 p_{BEST} ; 否则令 $s = \exp(-\Delta f/T)$, 若 $s > \text{rand}(0, 1)$, 则更新 p_{BEST} , 否则 p_{BEST} 不变;

5)对粒子的速度和位置进行进化,得到下一代粒子个体, $t = t + 1$,根据温度下降函数 $T = T(t)$,更新温度 T ;

6)若粒子适应度变化量小于设定阈值,或 $t > t_{\max}$,则算法终止;否则返回 2),进行下一步。

2 基于 KPCA-WLSSVM 建筑能耗预测

2.1 建筑能耗影响因素分析

建筑能耗系统是一个复杂的非线性系统,影响建筑能耗的因素众多,主要分为 4 大类^[12,15-16]:外部气候条件、围护结构、设备因素以及建筑运行管理模式。外部气候条件是影响建筑能耗的“自然”因素,包括温度、湿度、日照及风速等参数,它们能够直接影响热量通过围护结构的传递,进而对建筑能耗产生影响;围护结构可以看作是影响建筑能耗的“人为”因素,包括了反映建筑热工性能的围护结构传热系数、遮阳率等参数和反映建筑体形特征的体形系数、建筑面积、朝向、窗墙比等参数,这类因素可以通过改善建筑本身的热工性能及相关参数对建筑能耗起到积极作用;设备因素包括照度标准、灯具类型、设备功率等参数,这些参数取决于建筑的使用功能;建筑运行管理模式也是影响建筑能耗的重要因素之一,它包括了通风次数、人员密度、设备运行时间、新风指标等参数。

2.2 建模步骤

由于建筑能耗与众多能耗影响因素之间存在复杂的非线性关系,在建模过程中会增加计算复杂度,降低建模精度。因此,有必要通过降维处理去除样本数据集中的冗余信息以及消除噪声,以提高模型的预测精度和泛化能力。笔者提出的 KPCA-WLSSVM 的建筑能耗预测模型首先利用 KPCA 提取能耗因子的非线性特征,消除数据冗余并降低模型输入维数,然后采用 WLSSVM 方法建立能耗预测模型。模型的结构如图 1 所示。

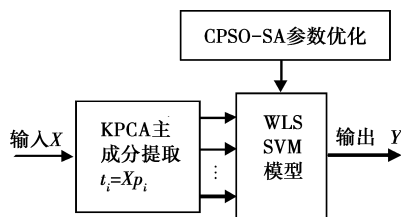


图 1 KPCA-WLSSVM 建模流程图

Fig 1 Sketch of KPCA-WLSSVM

考虑到 LSSVM 算法的正则化参数和核宽参数

对模型拟合精度和泛化能力的有较大影响^[20],笔者采用 CPSO-SA 算法在正则化参数 C 及核宽参数 σ 的定义域内进行搜索,寻找 C 和 σ 的组合最优值。参数优化问题定义为

$$\min_{C, \sigma} J = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (18)$$

$$\text{s. t.} \quad C_{\min} \leq C \leq C_{\max}$$

$$\sigma_{\min} \leq \sigma \leq \sigma_{\max}$$

式中: y_i 为第 i 个已知样本的值; \hat{y}_i 为第 i 个样本的模型输出值,由预测模型计算获得; J 是正则化参数 C 及核宽参数 σ 的函数。

综上所述,基于 CPSO-SA 优化的 KPCA-WLSSVM 建筑能耗预测模型算法关键步骤总结如下:

1)样本数据归一化处理。为了避免计算出现饱和和现象,按式(19)对建筑能耗训练样本进行归一化处理。

$$T = T_{\min} + \frac{T_{\max} - T_{\min}}{X_{\max} - X_{\min}}(X - X_{\min}) \quad (19)$$

式中: X 是原样本数据; X_{\max}, X_{\min} 分别代表样本数据中最大值和最小值; T 是目标数据; T_{\max} 是目标数据的最大值,取 $T_{\max} = 0.9$; T_{\min} 是目标数据的最小值,取 $T_{\min} = 0.1$ 。

2)KPCA 特征提取。按式(7)提取主成分,降低样本维数,消除变量之间的相关性。

3)最优模型参数确定。初始化 CPSO-SA 粒子种群,每个粒子由 C 和 σ 组成;基于 KPCA 提取的数据集和初始粒子个体参数,建立 LSSVM 模型,进而得到预测输出 \hat{y}_i ;根据式(18)计算每个粒子相应的适应度值(这里适应度函数定义为优化目标函数);根据 2.3 节步骤 3)~6)获得 LSSVM 模型的最优参数 C^* 和 σ^* 。

4)基于最优参数 C^* 和 σ^* 由式(14)求解 b, α^* ,建立 WLSSVM 建筑能耗预测模型。

5)输入建筑能耗测试样本,得到输出 \hat{y} 。

6)反归一化。按式(20)对模型输出进行反归一化处理,

$$X = X_{\min} + \frac{X_{\max} - X_{\min}}{T_{\max} - T_{\min}}(T - T_{\min}) \quad (20)$$

得到建筑能耗对应的预测值 \hat{y} 。

7)模型预测性能评价。采用相对误差绝对值最大值 E_{\max} 、相对误差绝对值平均值 E_{ave} 和均方根误差 RMSE 三个性能指标对模型进行评价。

$$E_{\max} = \max \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (21)$$

$$E_{\text{ave}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (22)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (23)$$

式中: N 为样本数。

2.3 实例分析

2.3.1 建筑物简介 用于验证文中所建立的 KPCA-WLSSVM 建筑能耗预测模型的建筑物是南方地区一幢 16 层的办公建筑,其外形如图 2 所示。

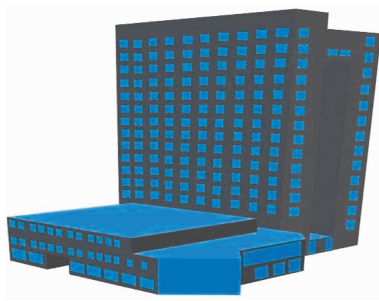


图 2 建筑外形示意图

Fig 2 Shape of office building

该建筑总高度为 57.4 m,主楼地上 16 层,地下 1 层,附楼(裙房)3 层。总建筑面积为 19 700.5 m²。地上建筑面积 17 315.62 m²,地下建筑 2 384.88 m²,标准层层高 3.5 m,外窗尺寸为 2.5 m×2.0 m,窗台离地高度 0.8 m。空调运行时间为 8:00—17:00,总计每天运行 9 h,空调季为 6 月 1 日到 8 月 30 日。

2.3.2 构造样本集 根据能耗影响因素分析,并结合《公共建筑节能设计标准》相关规定,在对办公建筑能耗建模过程中,主要考察以下能耗因子:室外温度(x_1)、室外湿度(x_2)、太阳辐射(x_3)、空调系统运行模式(x_4)、最低照度(x_5)、灯光热扰(x_6)、室内温度(x_7)、室内湿度(x_8)、设备热扰(x_9)、电热转换率(x_{10})、人均新风指标(x_{11})、家具系数(x_{12})、外墙传热系数(x_{13})、内墙传热系数(x_{14})、玻璃遮阳系数(x_{15})、玻璃传热系数(x_{16})、通风次数(x_{17})^[21]。

同时,为减少试验次数,并保证试验的全面性,采用正交试验方法获取数据样本集。DEST 中影响因素的取值,采用事先设定好的水平值,每个因素选取 4 个水平值,选用 L₆₄(4²¹) 正交表进行试验计算,建筑物能耗的基准值采用动态负荷计算软件 DEST-C 的计算值。为获得较为理想的训练数据,

选用 4 组正交试验,并且每组正交试验中每个因素选取不同的水平值,经过 DEST-C 动态计算,获得建筑物全年能耗结果。经 DEST-C 计算获 256 组实验数据,随机取其中的 3/4 作为训练样本,余下的 1/4 作为测试样本验证模型的预测性能。

2.3.3 模型训练 通过 KPCA 计算,取主元个数为 8 个,此时累计贡献率已经超过 90%。碎石图如图 3 所示。另外,核参数和正则化参数是 WLSSVM 性能的重要影响因素,不同的参数组合将得到不同的学习性能和泛化性能。为避免主观经验选取的盲目性,采用 CPSO-SA 算法对模型的两个参数进行寻优,优化后的参数分别为 $C = 326.7$ 、 $\sigma = 0.86$,满足预测模型对精度的要求,建筑能耗模型训练结果如图 4 所示,拟合值和期望值基本吻合。

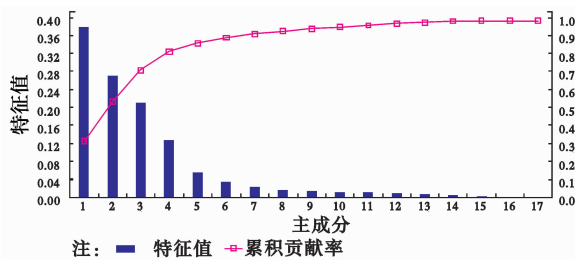


图 3 核主成分碎石图

Fig 3 The scree plot of kernel principal component

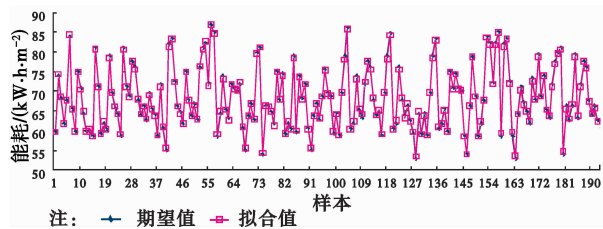


图 4 KPCA-WLSSVM 模型训练结果

Fig 4 The training results of KPCA-WLSSVM

2.3.4 结果与分析 为比较分析组合模型的预测效果,同时建立了 LSSVM、WLSSVM 和 RBFNN 神经网络模型对建筑能耗进行预测。通过 MATLAB 平台实现上述 4 种预测模型对建筑能耗进行预测,预测值和实际值的对比如图 5 所示。从图中可以知道,4 种模型的预测结果大都能够反映能耗的变化,但有个别样本点偏差较大,组合模型的预测值和期望值基本重合,相比于单一的 LSSVM、WLSSVM 和 RBFNN 神经网络模型,组合预测模型的预测效果较好。

为了从整体上评价 4 种预测模型的预测精度和稳定性,按式(21)、(22)及(23)对上述 4 种模型进行

评价,计算结果如表 1 所示。

表 1 种模型评价指标计算结果

Table 1 The predictive results of the four models

模 型	训练集			测试集		
	相对误差绝对 值最大值 $E_{max}/\%$	相对误差绝对 值平均值 $E_{ave}/\%$	均方根误 差 RMSE	相对误差绝对 值最大值 $E_{max}/\%$	相对误差绝对 值平均值 $E_{ave}/\%$	均方根误 差 RMSE
RBFNN	0.039	0.021	0.026	14.72	3.11	9.86
LSSVM	0.33	0.26	0.034	6.88	2.96	4.01
WLSSVM	0.49	0.31	0.039	7.81	1.13	2.89
KPCA-WLSSVM	0.48	0.29	0.026	2.81	0.89	0.84

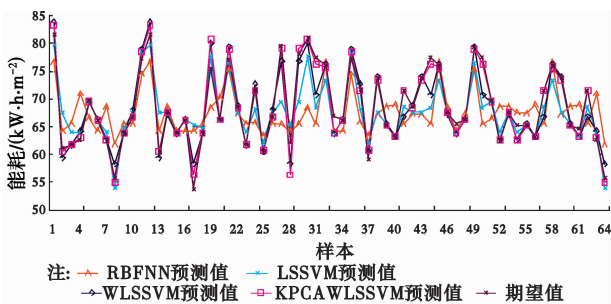


图 5 4 种模型的预测结果

Fig 5 Prediction of the four models

从表 1 的计算结果可以看出,采用 RBF 神经网络和 LSSVM 方法预测虽然具有很高的训练精度,但是当其用于预测时,误差较大,泛化性能差,预测效果不理想;采用 WLSSVM 进行预测,由于该方法对每个样本进行了加权处理,有效消除了显著误差的影响,具有较好的预测精度,其平均预测误差和均方根误差比 LSSVM 有所降低,但预测效果仍有待提高;KPCA-WLSSVM 方法首先采用 KPCA 对能耗样本数据进行数据压缩和特征提取,去除样本数据中的冗余信息以及消除噪声,使得 WLSSVM 模型的预测效果得到了很大的改善,最大相对误差绝对值从 7.81% 降到 2.81%,相对误差绝对值平均值从 1.13% 降低到 0.89%,这表明组合模型具有更高的预测精度和预测稳定性,泛化能力得到显著提高。通过对比分析模型的性能指标,结果表明 KPCA-WLSSVM 模型具有更好的预测效果,是一种有效、可行的建筑能耗预测方法。

3 结 论

通过 4 个模型的预测结果及性能评价指标的比较可以看出,基于 KPCA-WLSSVM 能耗预测模型的预测效果要优于其他 3 种模型,表明了 KPCA-

WLSSVM 模型在建筑能耗预测中的有效性和优越性。主要结论为:

1)核主元分析能够消除建筑能耗影响因素间的冗余性和共线性,正确提取输入样本的特征信息,简化模型的输入样本。

2)WLSSVM 充分考虑了各样本对建模的影响,有效地赋予各样本合适的权值,优化了 LSSVM 模型,使得模型具有良好的非线性处理能力。

3)采用 CPSO-SA 优化算法对模型的参数进行优化选择,避免了经验选择的缺陷,在此基础上建立的建筑能耗预测模型具有更好的学习精度和推广应用能力。

参考文献:

[1] Wong S L, Wan K K W, Lam T N T. Artificial neural networks for energy analysis of office buildings with daylighting [J]. Applied Energy, 2010, 87(2): 551-557.

[2] Azadeh A, Ghaderi S F, Tarverdian S, et al. Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption [J]. Applied Mathematics and Computation, 2007, 186(2): 1731-1741.

[3] 喻伟, 李百战, 杨明宇, 等. 基于人工神经网络的建筑多目标预测模型[J]. 中南大学学报: 自然科学版, 2012, 43(12): 4950-4955.

Yu W, Li B Z, Yang M Y, et al. Building multi-objective predicting model based on artificial network [J]. Journal of Central South University: Science and Technology, 2012, 43(12): 4950-4955. (in Chinese)

[4] Mathieu J L, Price P N, Kiliccote S, et al. Quantifying changes in building electricity use, with application to demand response [J]. IEEE Transactions on Smart

- Grid, 2011, 2(3): 507-518.
- [5] Vapnik V. The nature of statistical learning theory [M]. Springer, 2000.
- [6] Cortes C, Vapnik V. Support vector networks [J]. Machine Learning, 1995, 20(3): 273-297.
- [7] Dong B, Cao C, Lee S E. Applying support vector machines to predict building energy consumption in tropical region [J]. Energy and Buildings, 2005, 37(5): 545-553.
- [8] Li Q, Meng Q L, Cai J J, et al. Applying support vector machine to predict hourly cooling load in the building [J]. Applied Energy, 2009, 86(10): 2249-2256.
- [9] Suykens J A K, Van G T, De B J, et al. Least squares support vector machines [M]. Singapore: World Scientific, 2002.
- [10] Li X M, Shao M, Ding L X, et al. Particle swarm optimization-based LS-SVM for building cooling load prediction [J]. Journal of Computers, 2010, 5(4): 614-621.
- [11] Jung H C, Kim J S, Heo H. Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach [J]. Energy and Buildings, 2015, 90: 76-84.
- [12] Zhao H X, Magoules F. Parallel support vector machines applied to the prediction of multiple building energy consumption [J]. Energy, 2009, 4(2): 231-249.
- [13] Suykens J A K, de Brabanter J, Lukas L, et al. Weighted least squares support vector machines: robustness and sparse approximation [J]. Neurocomputing, 2002, 48: 85-105.
- [14] Guo Y, Nazarian E, Ko J, et al. Hourly cooling load forecasting using time-indexed arx models with two-stage weighted least squares regression [J]. Energy Conversion and Management, 2014, 80:46-53.
- [15] Li Q, Meng Q L, Cai J J, et al. Predicting hourly cooling load in the building: a comparison of support vector machine and different artificial neural networks [J]. Energy Conversion and Management, 2009, 50(1): 90-96.
- [16] Crawley D B, Hand J W, Kummert M, et al. Contrasting the capabilities of building energy performance simulation programs [J]. Building and Environment, 2008, 43(4): 661-673.
- [17] Fan Z Z, Wang J H, Xu B G, et al. An efficient KPCA algorithm based on feature correlation evaluation [J]. Neural Computing and Applications, 2014, 24(7/8): 1795-1806.
- [18] Jia P F, Tian F C, He Q H, et al. Feature extraction of wound infection data for electronic nose based on a novel weighted KPCA [J]. Sensors and Actuators, B: Chemical, 2014, 201:555-566.
- [19] Suykens J A K. Least squares support vector machines [J]. Neural Processing Letters, 1999, 9(3):293-300.
- [20] 樊晓雪. 最小二乘支持向量机的参数选择[D]. 河北保定: 河北大学, 2012.
- Fan X X. Least squares support vector machine parameter selection [D]. Baoding: Hebei University, 2012. (in Chinese)
- [21] 康一亭, 于丹, 苏华, 等. 办公建筑能耗基准中影响因素对建筑能耗的影响度[J]. 中南大学学报: 自然科学版, 2012, 43(Sup1): 146-150.
- Kang Y T, Yu D, Su H, et al. Impact factors of energy consumption benchmarking in office buildings [J]. Journal of Central South University: Science and Technology, 2012, 43(Sup1): 146-150. (in Chinese)

(编辑 胡英奎)