

特征空间闭操作驱动的短期电力负荷预测

刘辉, 钟俊

(安徽职业技术学院机电工程学院, 安徽 合肥 230011)

摘要:为了获得更高的短期负荷预测精度,有必要充分考虑负荷变化趋势与区域整体用电行为模式之间的关联,提出一种特征空间闭操作驱动的短期电力负荷预测方法。在综合模型的基础上,首先利用特征提取模型将历史用电负荷分解成多个分量作为刻画区域用电行为的特征;然后使用特征选择模型对用电行为特征进行选择,减少冗余或无效特征的干扰,优化预测模型;最后将选择的特征子集作为预测模型的输入特征从而进一步估计出各时段负荷的分布。结果表明采用本方法预测精度更高。

关键词:用电行为特征;特征提取;特征选择;短期负荷预测

中图分类号: TM715 **文献标志码:** A **文章编号:** 1673-1891(2020)01-0049-05

Short-term Power Load Forecasting Driven by Feature Space Closed Operation

LIU Hui, ZHONG Jun

(Department of Mechanical and Electrical Engineering, Anhui Vocational and Technical College, Hefei, Anhui 230011, China)

Abstract: To obtain higher short-term load forecasting accuracy, it is necessary to fully consider the relationship between load variation trend and regional overall power consumption behavior pattern, so in this paper a short-term power load forecasting method driven by feature space closed operation is proposed. With this method we first divide the historical load data into multiple components to describe the characteristics of the electricity consumption behavior of the delineated area; then, according to the regional load change trend, the characteristics of the power consumption behavior are selected, the redundant input features are eliminated, and the prediction model is optimized; finally, the remaining input features are brought into the prediction model to further estimate the distribution of the load in each period. The results show that the prediction accuracy of the proposed method is higher.

Keywords: power consumption behavior characteristics; feature extraction; feature selection; short-term load forecasting

0 引言

短期电力负荷预测是电网调度和运营的重要环节,主要用于预报未来一天到数天各时段的系统负荷^[1]。做好电力负荷预测能够提高电力系统的运行柔性,为电力系统运营商提供可靠的决策支持服务。

短期电力负荷预测的方法主要集中在统计学方法、机器学习方法和综合模型方法等。时间序列模型是统计学中常用预测方法,构造简单,考虑了时间影响因素,保证了负荷预测的连续性,但是随着时间的推移,其预测精度会减小。机器学习模型主要有支持向量回归(Support Vector Regression,

SVR)和人工神经网络(Artificial Neural Network, ANN)模型。机器学习模型设计简单,计算机易实现,对于动态变化的负荷数据可以进行高精度的非线性拟合,这些研究在过去取得了巨大的进步。但是单一的机器学习模型不能充分考虑负荷变化的影响因素,模型预测能力受到限制。将多种模型相结合的综合模型方法已经在电力负荷预测领域使用,并且得到了更多研究者的深入关注^[2-4]。综合模型从多方面分析负荷的波动趋势,利用各模型自身的优势对负荷数据进行分析提取,充分考虑负荷变化与影响因素之间的关联,为进一步提升预测效果提供可能。

基于此,本文提出一种特征空间闭操作驱动的

短期电力负荷预测方法。该方法首先对历史负荷数据进行分解得到刻画区域用电行为的特征;然后依据区域负荷变化趋势对用电行为特征进行选择,剔除冗余输入特征,优化预测模型;最后将剩余输入特征带入预测模型从而进一步估计出各时段负荷的分布。结果表明采用本文所提方法预测精度更高。

1 用电行为特征提取与选择

1.1 特征提取

电力负荷是由经济状况、温度、人类的生产活动等诸多因素叠加影响而形成的非线性、非平稳的具有周期性的信号。如果直接对整个负荷信号进行分析处理,则影响信号变化的潜在因素无法得到充分利用,负荷预测的效果也将受到影响。因此通过对历史负荷信号进行分解,将影响负荷信号变化的因素分离出来作为分量信号,从细节中了解区域用电情况和负荷信号的波动趋势,提升短期负荷预测效果。

信号分解的方法主要有以下 3 种。包括基于傅立叶变换的信号分解、小波分解和经验模态分解^[5](Empirical Mode Decomposition, EMD)。基于傅立

叶变换的信号分解,该方法仅适用于平稳和线性数据,无法准确分解非线性变化的负荷信号。小波分解是一种常用的信号分解方法,但是小波分解的效果容易受小波基函数和分解层数的影响。EMD 法可通过自适应地分解非平稳负荷序列,提高预测精度,因此在处理非平稳及非线性信号上具有非常明显的优势。针对此,本文通过 EMD 方法对负荷信号进行分解提取用电行为特征。

假设给定某区域一周的用电负荷信号为 $G(n)$, 每 1 h 记录一次,一周共有 168 个负荷数据。EMD 将 $G(n)$ 分解为各个不同时间尺度的 IMF (Intrinsic Mode Function) 分量和 一个 RSE (Residual StandardError) 分量。如式(1)所示:

$$G(n) = \sum_{l=1}^L I_l(n) + r(n), \tag{1}$$

其中 $I_l(n)(l=1,2,\dots,L)$ 表示本征模函数信号, $r(n)$ 表示残差函数信号。它们是与原信号 $G(n)$ 具有相同大小的平稳信号,包含了 $G(n)$ 不同时间尺度的局部特征信息,而这些局部特征信息中包含影响负荷变化的因素特征。信号 $G(n)$ 与各分解分量如图 1 所示。

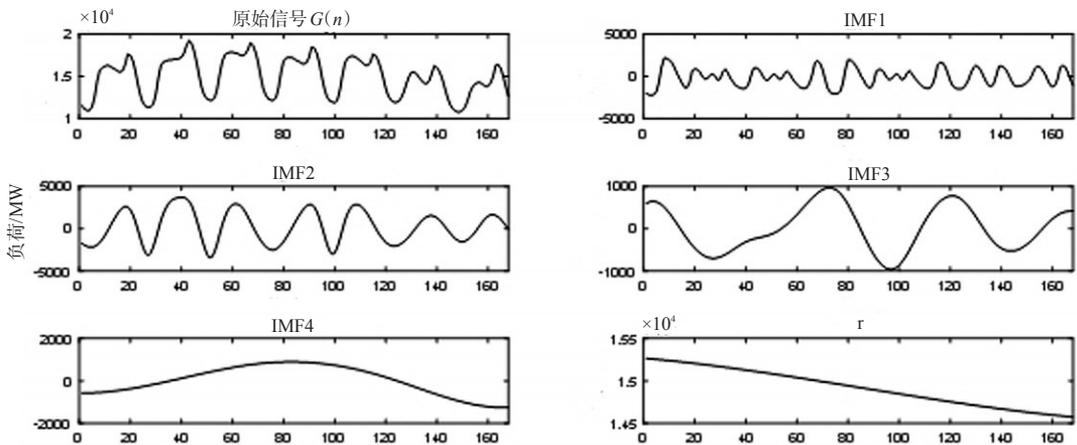


图 1 原始信号 $G(n)$ 与 EMD 分解分量

$G(n)$ 被分解成 5 个信号分量,其中图 IMF1-IMF4 表示本征模函数信号分量,图 r 表示残差函数信号分量。根据区域一周的分解分量可以得出区域一天的信号分量为式(2)。

$$W = \begin{pmatrix} IMF1_{t_1} & IMF1_{t_2} & \dots & IMF1_{t_{24}} \\ IMF2_{t_1} & IMF2_{t_2} & \dots & IMF2_{t_{24}} \\ IMF3_{t_1} & IMF3_{t_2} & \dots & IMF3_{t_{24}} \\ IMF4_{t_1} & IMF4_{t_2} & \dots & IMF4_{t_{24}} \\ r_{t_1} & r_{t_2} & \dots & r_{t_{24}} \end{pmatrix} \tag{2}$$

W 中每.1 行代表 1 个分量信号 1 天 24 个时刻点

的数值。每个分量信号在各时刻点上具有各自的波动趋势,将每个信号分量各时刻点的波动趋势进行整合,得出刻画该时刻点的用电行为特征,依此得出该天的用电行为特征如式(3)所示。

$$f = \begin{pmatrix} IMF1_{t_1} & IMF2_{t_1} & IMF3_{t_1} & IMF4_{t_1} & r_{t_1} \\ IMF1_{t_2} & IMF2_{t_2} & IMF3_{t_2} & IMF4_{t_2} & r_{t_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ IMF1_{t_{24}} & IMF2_{t_{24}} & IMF3_{t_{24}} & IMF4_{t_{24}} & r_{t_{24}} \end{pmatrix} \tag{3}$$

1.2 特征选择

在用电行为特征提取过程中,可能会提取到一些对负荷信号分析无作用甚至是有偏差的特征,为

了得到更准确的负荷预测结果,需要对用电行为特征进行选择。

特征选择可以看作是一个寻优过程。依据特征或者特征子集评价准则的不同,特征选择方法可以分为封装法、过滤法和嵌入法。在封装法中,特征选择方法使用优化后的分类算法来选择特征子集,不包含学习的过程,计算要求很高。过滤法通常在计算上更高效,但是这种方法不包含关于分类或回归函数的特定结构的知识,因此可能无法找到特征选择的最佳子集解决方案^[6]。与封装法和过滤法不同,在嵌入法中,学习部分和特征选择部分紧密相联,通过学习判断特征对模型的贡献从而实现特征的自动选择,提高特征选择的精度和效率。

最常见的嵌入式特征选择方法是正则化方法,如 LASSO (Least Absolute Shrinkage and Selection Operato) 回归及岭回归。LASSO 回归使用 $L1$ 正则化方法将特征矩阵中特征系数置为 0, 获得特征系数的稀疏解, 完成特征的选择。对比岭回归使用 $L2$ 正则化将特征系数压缩但不会置为 0, LASSO 回归能够更好的实现特征选择。因此通过 LASSO 回归对用电行为特征集合进行学习判断, 建立依据负荷波动趋势的学习模型, 完成特征的选择。LASSO 方法应用如下:

假设区域历史用电负荷被分为两个数据集, 训练集和测试集, 其中训练集用 $C1$ 表示。将 $C1$ 作为分析对象提取区域的用电行为特征集合, 则依据式(3)以及负荷变化的周期规律可得出 $C1$ 一天的特征集合为式(4)。

$$F_i(x) = \begin{pmatrix} x_{T_1 1} & x_{T_1} & \cdots & x_{T_1 50} \\ x_{T_2 1} & x_{T_2 2} & \cdots & x_{T_2 50} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T_{24} 2} & x_{T_{24} 2} & \cdots & x_{T_{24} 50} \end{pmatrix}, \quad (4)$$

其中 $F_i(x)$ 表示 $C1$ 第 i 天的特征集合, $T_j(j=1, 2, \dots, 24)$ 表示该天所对应的 24 个时刻点。每一行特征是以第 i 天为基准所选用的历史 10 d(除双休日)同时刻点的用电行为特征所构成。用电行为特征选择步骤中将 $C1$ n 天的特征集合 $X(x) = \{F_1(x), F_2(x), \dots, F_n(x)\}$ 作为 LASSO 回归的输入, 将 n 天各时刻点所对应的实际用电负荷作为输出进行回归训练, 得出选择后的用电行为特征。

2 短期用电负荷预测

2.1 用电负荷预测

神经网络、岭回归和支持向量回归等预测算法

在负荷预测领域应用都很广。由于短期负荷预测数据量相对较小, 特征维度低, 所以需要将特征映射到高维空间进行负荷拟合和预测。在回归算法中, SVR 的核函数可以将低维数据映射到高维空间, 实现小样本数据的回归预测。本文使用 SVR 作为回归预测算法进行短期负荷预测。

SVR 算法是支持向量机 (support vector machine, SVM) 在回归问题的推广^[7]。SVR 模型从数学角度研究给定输入特征 x 和输出特征 y 之间的最佳映射关系。具体的算法定义如下:

给定训练集的输入输出特征 $\{(x_1, y_1), \dots, (x_n, y_n)\}$, SVR 算法推导出如式(5)的回归函数, b 代表模型的截距向量。

$$f(x) = \sum_{i=1}^n w_i \phi_i(x_i) + b = w^T \phi(x) + b, \quad (5)$$

其中 $\phi(x) = [\phi_1(x_1), \dots, \phi_n(x_n)]^T$, $W = [w_1, w_2, \dots, w_n]^T$ 是模型的权重向量。

输入特征 x 由函数 $\phi(x)$ 映射到高维空间, 式(5)可通过最小正则化风险函数计算得式(6):

$$R(w) = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n |y_i - f(x_i)|_{\varepsilon}, \quad (6)$$

其中

$$|y_i - f(x_i)|_{\varepsilon} = \begin{cases} 0 & (|y_i - f(x_i)| < \varepsilon) \\ |y_i - f(x_i)| - \varepsilon & (|y_i - f(x_i)| \geq \varepsilon) \end{cases}. \quad (7)$$

式(6)中的 C 是惩罚因子, 它通过对式(6)施加足够的惩罚来最小化错误, $|y_i - f(x_i)|_{\varepsilon}$ 是不敏感损失函数 ε 。式(7)中的不等式表示如果训练集的预测值 $f(x_i)$ 与实际值 y_i 的误差水平在 ε 内, 则 $R(w)$ 不会受到惩罚。针对此, 将式(6)的拉格朗日函数用式(8)表示。

$$R(w, \xi_i, \xi_i^*) = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n (\xi_i + \xi_i^*)$$

$$\text{s.t.} \begin{cases} y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, n \end{cases}. \quad (8)$$

通过不等式约束来解决这个优化问题, 其中 ξ_i 和 ξ_i^* 是引入的松弛变量, 它能够对 SVR 回归算法的输出边界进行约束^[8]。通过应用拉格朗日乘子和标准二次规划技术^[9], 可以得到式(8)中优化问题的双重形式, 其中 $\lambda^* = (\lambda_1, \lambda_1^*, \dots, \lambda_n, \lambda_n^*)^T$ 表示拉格朗日乘子:

$$R(\lambda_i, \lambda_i^*, \lambda_j, \lambda_j^*)$$

$$\lambda^* \in R^{2n}$$

$$= \frac{1}{2} \sum_{i,j=1}^n (\lambda_i^* - \lambda_j) (\lambda_j^* - \lambda_i) \kappa(x_i, x_j) + \varepsilon \sum_{i=1}^n (\lambda_i^* + \lambda_i) - \sum_{i=1}^n y_i (\lambda_i^* - \lambda_i)$$

$$\text{s.t. } \sum_{i=1}^n (\lambda_i^* - \lambda_i) = 0, 0 \leq \lambda_i^* \leq c, i = 1, \dots, n. \quad (9)$$

上式中 $\kappa(x_i, x_j)$ 是满足 Mercer 定理的核函数,使用高斯径向基核函数进行回归预测。通过式(9)可以计算出 $\hat{\lambda}^* = (\hat{\lambda}_1, \hat{\lambda}_1^*, \dots, \hat{\lambda}_n, \hat{\lambda}_n^*)^T$, 则截距 $\hat{b} = y_i - \sum_{i=1}^n (\hat{\lambda}_i^* - \hat{\lambda}_i) \kappa(x_i, x_j) + \varepsilon$, 最终得出 SVR 回归函数如式(10)所示。

$$f(x) = \sum_{i=1}^n (\hat{\lambda}_i^* - \hat{\lambda}_i) \kappa(x_i, x_j) + \hat{b}. \quad (10)$$

2.2 建立短期负荷预测模型

本文所构造的短期负荷预测模型共包含三个组成部分,分别为用电行为特征提取、用电行为特征选择和用电负荷预测。负荷预测模型的流程如图2所示。

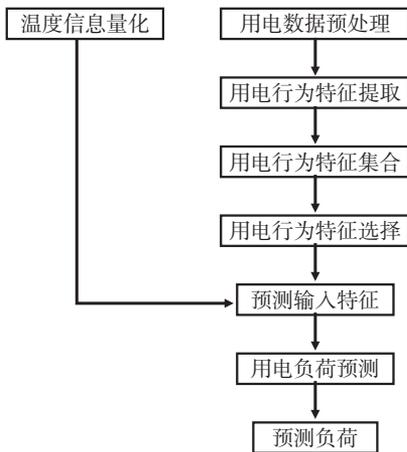


图2 短期负荷预测模型流程图

3 算例分析

3.1 算例数据

为验证本文所提方法的有效性,实验采用了我国某市区的真实用电数据。用电数据每1h记录一次,每d共24个记录值,实验使用前10周作为训练集。

3.2 算例设计

本实验提取该地区连续1a的用电负荷数据,实验使用后1周的数据作为测试集,对测试集的日负荷进行预测。

本文共设计3组算例实验,实验内容包括:

1)将本文所提模型(EMD-LASSO-SVR)与一般拟合模型(SVR和GRNN)预测目标日负荷的结果进行对比。

2)使用SVR作为回归算法,在用电行为特征提取中使用EMD方法分解提取用电行为特征,在对比实验中选用小波方法分解用电负荷提取特征。

3)使用SVR作为回归算法,在用电行为特征选择中使用LASSO方法对特征进行筛选,在对比实验中使用前向选择、关联度选择方法进行对比。

在实验中,将使用2个量化指标来评价负荷预测的精度。平均绝对百分误差(Mean Absolute Percentage Error, MAPE)值越小,模型预测效果越好;均方根误差(Root Mean Squard Error, RMSE)能够验证模型有效程度和评价模型精密程度。2个量化指标表示如式(11)(12)。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%, \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (12)$$

其中 y_i 为第 i 时刻的实际负荷值, \hat{y}_i 为该时刻的预测值, \bar{y} 表示均值, n 表示时刻点个数。

3.3 实验结果分析

表1对比了本文所提模型(EMD-LASSO-SVR)与一般模型(SVR和GRNN)预测目标日负荷的精度。图3展示了实验1预测目标日负荷的结果。

表1 EMD-LASSO-SVR与一般模型预测目标日负荷精度

模型类别	MAPE/%	RMSE/MW
GRNN	1.51	287.30
SVR	1.42	227.30
EMD-LASSO-SVR	1.12	102.24

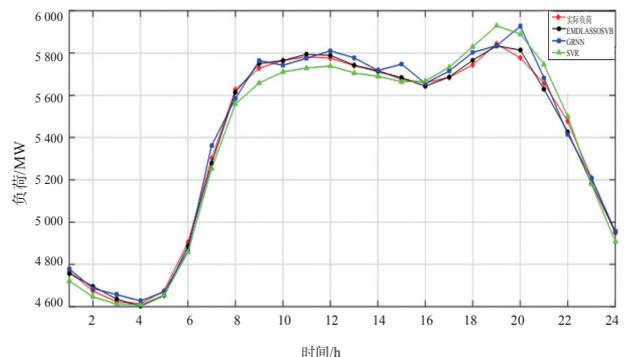


图3 EMD-LASSO-SVR与一般模型预测目标日负荷结果

从表1和图3中可以看出在预测精度上,GRNN要劣于SVR,2种一般拟合模型与本文相比,本文所

提模型的预测精度和拟合效果更佳。

表 2 对比了在使用 SVR 作为回归算法的基础上,选用不同特征提取方法预测目标日负荷的精度。图 4 展示了实验 2 预测目标日负荷的结果。表 2 和图 4 展示了实验 2 预测目标日负荷的精度和结果。从表 2 和图 4 可以看出,本文所提模型所使用的 EMD 方法在特征提取过程中表现出更好的效果,目标日负荷预测的两个指标均优于小波方法,且目标日负荷预测的结果更能拟合实际负荷值。

表 2 选用不同特征提取方法预测目标日负荷精度

预测模型	MAPE/%	RMSE/MW
小波-LASSO-SVR	1.44	254.29
EMD-LASSO-SVR	1.09	98.41

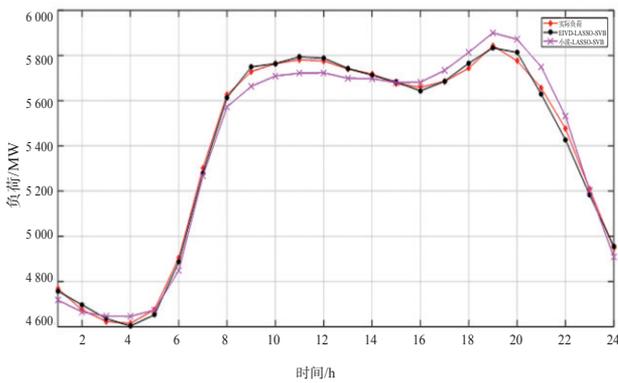


图 4 选用不同特征提取方法预测目标日负荷结果

表 3 和图 5 展示了实验 3 预测目标日负荷的结果。通过表 3 和图 5 可以分析得出,选用 LASSO 回归对特征进行选择的方法在目标日负荷预测上的精度高于其余对比选择方法,且目标日负荷预测的结果更符合实际负荷的波动趋势。

参考文献:

- [1] 李若晨,朱帆,朱永利,等.结合受限玻尔兹曼机的递归神经网络电力系统短期负荷预测[J].电力系统保护与控制,2018,46(17):83-88.
- [2] 陈王瑞,周晨曦,逯静.基于一种组合模型的短期电力负荷预测[J].计算机仿真,2018,35(12):396-400.
- [3] 程津,黎燕,夏向阳,等.基于双组合预测的经济—电力负荷预测模型[J].电力科学与技术学报,2018,33(3):18-22.
- [4] 陈勇,李鹏,张忠军,等.基于 PCA-GA-LSSVM 的输电线路覆冰负荷在线预测模型[J].电力系统保护与控制,2019(10):110-119.
- [5] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings Mathematical Physical & Engineering Sciences,1998,454(1971): 903-995.
- [6] TIBSHIRANI R. Regression shrinkage and selection via the lasso[J]. Journal of the Royal Statistical Society Series B (Methodological), 1996, 58(1): 267-288.
- [7] BALABIN R M, LOMAKINA E I.Support vector machine regression—an alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data[J]. Analyst, 2011,136(8):1703-1712.
- [8] HUANG N E, WU M L, QU W D. Applications of Hilber-Huang transform to non-stationary financial time series analysis[J]. Applied Stochastic Models In Busine-ss And Industry. 2003, 361(19): 245-268.
- [9] ZOU Hui,HASTIE T.R.regularization and variable selection via the elastic net[J]. Journal of the Royal Statistical Society, 2010,67(5): 768-768.

表 3 选用不同特征选择方法预测目标日负荷精度

预测模型	MAPE/%	RMSE/MW
EMD-关联度-SVR	1.56	274.17
EMD-前向选择-SVR	1.49	228.92
EMD-LASSO-SVR	1.01	99.07

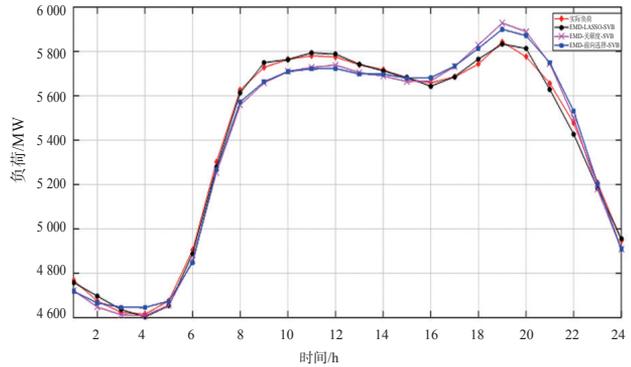


图 5 选用不同特征选择方法预测目标日负荷结果

4 结语

本文通过分析区域用电负荷的波动特征和负荷影响因素,提出一种特征空间闭操作驱动的短期电力负荷预测方法。该方法首先采用 EMD 模型对区域历史用电负荷信号进行分解,将分解分量作为刻画区域用电行为的特征;然后将待测日前 2 周工作日的用电行为特征构成特征集合,通过 LASSO 方法对特征集合中的特征进行选择,将有利于刻画区域用电行为的特征选择出来作为特征子集;最后将选择得到的特征子集作为待测日负荷预测的输入特征,待测日的实际负荷作为输出特征进行 SVR 模型拟合预测。通过使用我国某市区的真实用电数据进行对比实验表明,本文所提方法预测结果更加准确。