

# Short-Term Load Forecasting Based on Variable Weighted Synthesis of Different Kernel SVM

Dongfen Ma

School of Statistics and Data Science, Xinjiang University of Finance and Economics, Urumqi Xinjiang  
Email: 1224190081@qq.com

Received: Jan. 17<sup>th</sup>, 2020; accepted: Jan. 31<sup>st</sup>, 2020; published: Feb. 6<sup>th</sup>, 2020

---

## Abstract

To improve the accuracy and stability of short-term load forecasting, a method of short-term load forecasting based on different kernel support vector machine (SVM) variable weight synthesis is proposed. In this method, firstly, the load history data is expanded, the feature is selected by correlation analysis, and the historical data is mapped to the input-output relationship view to build the forecasting space. Then, support vector machines of Gaussian kernel, Laplace kernel and Polynomial kernel function are used to study in the forecasting space respectively, and the performance of the model is tested by the 10 fold cross validation. Finally, the variable weight is constructed by using the accuracy and standard deviation of performance test, and the power load forecasting is realized by variable weight synthesis of multiple model. Example analysis shows that compared with methods such as Gaussian kernel support vector machine, partial least squares, decision tree and Bagging, the new method improves accuracy by 0.382%, 3.079%, 3.188% and 2.6%, and stability by 0.383%, 2.452%, 1.781% and 1.43%, respectively.

## Keywords

Short-Term Load Forecasting, SVM, Kernel Function, Variable Weight Synthesis of Multiple Models

---

# 基于不同核支持向量机变权综合的短期负荷预测

马冬芬

新疆财经大学统计与数据科学学院, 新疆 乌鲁木齐  
Email: 1224190081@qq.com

收稿日期: 2020年1月17日; 录用日期: 2020年1月31日; 发布日期: 2020年2月6日

## 摘要

针对短期负荷预测准确度提升及预测稳定性改进问题,给出了基于不同核支持向量机变权综合的短期负荷预测方法。该方法首先将负荷历史数据进行特征展开,利用相关分析进行特征选择,映射历史数据为输入输出关系视图,构建预测空间。随后,分别采用高斯径向基核函数、拉普拉斯核函数以及多项式核函数的支持向量机在预测空间进行训练学习,使用十折交叉验证进行模型性能测试。最后,利用性能测试的准确率及其标准差构造变权,借助多模型变权综合实现电力负荷预测。实例分析表明,与高斯核支持向量机、偏最小二乘、决策树及Bagging等常用方法相比,新方法将准确度分别提升了0.382%、3.079%、3.188%以及2.6%,将稳定性分别改进了0.383%、2.452%、1.781%以及1.43%。

## 关键词

短期负荷预测, 支持向量, 核函数, 多模型变权综合

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

电力负荷预测是指通过研究历史数据变化规律,在满足一定准确度条件下,运用数学方法或建立统计模型的方式对未来几小时、一天到几天的负荷变化做出估计与推断[1][2]。准确的电力负荷预测对电网合理规划与电网安全运营具有重要意义[3]。

电力负荷预测方法可以被归为三类,一类以传统预测方法为主进行预测[4][5][6],例如,Chen J. F.等采用自适应移动平均进行短期负荷预测[4]。张伏生与汪鸿等使用偏最小二乘方法建立短期负荷预测模型[5]。万坤等采用ARMA方法构建短期负荷预测模型[6]。第二类以包含改进型在内的智能预测方法进行预测[7]-[11],例如,魏立兵与赵峰等采用人群搜索算法优化的支持向量机构建短期负荷预测模型[7]。Zhang X.等采用布谷鸟优化的支持向量机构建短期负荷预测模型[8]。梁智等采用PSO优化的深度信念网络进行负荷预测[9]。吴润泽等采用深度学习进行短期电力负荷预测[10]。第三类以组合型预测方法为主进行预测[12][13][14][15],例如,Sudheer G.等构建了组合小波变换、三重指数平滑和加权最近邻的负荷预测模型[12]。Li Song与Goel Lalit等集成小波变换、极值学习机与偏最小二乘回归进行短期负荷预测[13]。Wang L.与Lee E. W. M.等结合人工神经网络与集成方法进行负荷预测[14]。陆继翔等采用混合卷积神经网络与长短期记忆模型进行负荷预测[15]。

本文沿用组合预测方法,通过变权综合将高斯核函数的支持向量机、多项式核函数的支持向量机以及拉普拉斯核函数的支持向量机进行融合,构建基于多种核支持向量机变权综合的短期负荷预测模型。

## 2. 支持向量机算法原理

支持向量机是一种常用的统计建模方法,在解决高维度以及非线性问题方面具有优势。设函数表达式为: $f(x) = \omega^T \phi(x) + b$ ,其中, $\omega$ 为权重向量, $\phi(x)$ 为关于 $x$ 的映射函数, $b$ 为常量。要使 $\|\omega\|^2/2$ 达到最小,求解最优化问题:

$$\begin{aligned} & \min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i, \xi_i^*) \\ & \begin{cases} y_i - \langle \omega, \varphi(x_i) \rangle - b \leq \varepsilon + \xi_i, & i = 1, \dots, n; \\ \langle \omega, \varphi(x_i) \rangle + b - y_i \leq \varepsilon + \xi_i^*, & i = 1, \dots, n. \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \quad (1)$$

其中,  $\langle \cdot, \cdot \rangle$  表示内积,  $\xi_i, \xi_i^*$  为松弛变量,  $C, \varepsilon$  均为参数。使用 Lagrange 乘数法求解式(1),

$$\begin{aligned} L = & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i, \xi_i^*) - \sum_{i=1}^n (\eta_i \xi_i, \eta_i^* \xi_i^*) \\ & - \sum_{i=1}^n \alpha_i (\varepsilon + \xi_i - y_i + \langle \omega, \varphi(x_i) \rangle + b) \\ & - \sum_{i=1}^n \alpha_i^* (\varepsilon + \xi_i^* - y_i + \langle \omega, \varphi(x_i) \rangle + b) \end{aligned} \quad (2)$$

其中,  $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$  为 Lagrange 乘数。求解 Lagrange 乘数, 记  $\hat{\alpha}, \hat{\alpha}^*, \hat{\omega}, \hat{b}$  是对应  $\alpha, \alpha^*, \omega, b$  的解, 引入核函数  $k(x, x_i) = \phi(x)^T \phi(x_i)$ , 可解得

$$f(x) = \sum_{i=1}^n (\hat{\alpha}_i - \hat{\alpha}_i^*) k(x, x_i) + \hat{b}。$$

### 3. 基于不同核支持向量机变权综合的短期负荷预测方法

#### 3.1. 主要思路

基于不同核支持向量机变权综合的短期负荷预测方法主要有以下几个环节: 首先, 将负荷历史数据进行特征展开, 借助相关分析进行特征选取, 组建特征集, 通过对历史数据进行输入输出关系化展开, 构建预测空间。随后, 采用高斯径向基核函数(G-K)、拉普拉斯核函数(La-K)、多项式核函数(P-K)支持向量机在预测空间上进行建模, 使用交叉验证进行模型性能测试, 根据性能测试结果设计不同模型的权重。最后, 对三个不同核函数的支持向量机进行变权综合, 实现多种核函数支持向量机变权综合的负荷预测。

#### 3.2. 基于相关分析的负荷特征选择

设长度为  $N$  的负荷序列  $S = \{s_t, t = 1, 2, \dots, N\}$  的最大与最小的可能负荷取值为  $S_{\max}$  与  $S_{\min}$ , 取值论域为  $X(f) = [S_{\min}, S_{\max}]$ 。记待预测时刻  $t$  对应的负荷为特征  $f_0$ ,  $t$  前  $i$  个时刻对应的负荷记为特征  $f_i$ , 则对时刻  $t$  的负荷进行预测的候选特征集可被记为  $F_t = \{f_i | s_{t-i} \in S\}$ 。

记待预测时刻点对应的负荷特征  $f_0$  的前  $m$  个相邻的向量为:  $X_t = (s_{t-1}, \dots, s_{t-m})$ , 其中,  $t = 1, \dots, M$ ,  $M = N - m$ 。借助相关度量  $f_0$  与  $f_i$  之间的相关程度, 则  $f_0$  与  $f_i$  之间的相关度可表示为:

$$r_i = r(f_0, f_i) = \frac{\left| \sum_{k=1}^m \left( s_{t,k} - \frac{1}{m} \sum_{k=1}^m s_{t,k} \right) \left( s_{t-i,k} - \frac{1}{m} \sum_{k=1}^m s_{t-i,k} \right) \right|}{\sqrt{\sum_{k=1}^m \left( s_{t,k} - \frac{1}{m} \sum_{k=1}^m s_{t,k} \right)^2} \sqrt{\sum_{k=1}^m \left( s_{t-i,k} - \frac{1}{m} \sum_{k=1}^m s_{t-i,k} \right)^2}} \quad (3)$$

$r_i$  的取值为  $r_i \in [0, 1]$ ,  $r_i$  越接近于 1,  $f_0$  与  $f_i$  之间的相关度越强。设定阈值  $\alpha$ , 将满足条件的特征所组成的特征集合表示为  $F_t^\alpha = \{f_i | r_i \geq \alpha\}$ 。

### 3.3. 预测空间构造

以  $f_0$  为输出特征, 以特征集  $F_t^\alpha$  内的特征为输入特征, 建立预测空间  $\Omega = \prod_{i=1}^n X(f_i) \times X(f_0)$ 。

对给定的历史负荷数据集, 在将历史负荷数据进行输入输出关系化展开后, 得到关系化数据集为  $D = \{(X_t, s_t), t = 1, \dots, N\}$ , 其中  $X_t = (s_{t-1}, s_{t-2}, \dots, s_{t-n})$ ,  $N$  为关系化数据记录条数。

### 3.4. 核函数选择及模型建立

支持向量机中常见的核函数有高斯径向基核函数(Gaussian Basis Kernel, G-K)、拉普拉斯核函数(Laplacian Kernel, La-K)、线性核函数(Linear Kernel, Ln-K)、多项式核函数(Polynomial kernel, P-K)及 Bessel 核函数等。本文选取常用的高斯径向基核、拉普拉斯核及多项式核用以在关系化数据集上建立不同支持向量机模型。其中, 核函数的表达式为:

$$k_{G-K}(u, v) = \exp\left(-\gamma \|u - v^T\|^2\right) \quad (4)$$

$$k_{P-K}(u, v) = (\gamma \langle u, v \rangle + c)^p \quad (5)$$

$$k_{La-K}(u, v) = \exp\left(-\gamma \|u - v^T\|\right) \quad (6)$$

### 3.5. 模型性能测试及变权构造

借助  $P$  折交叉验证对不同核函数支持向量机进行性能测试, 记第  $i$  折的训练数据记录条数为  $H$ , 测试得到的准确率及其标准差分别为  $A_i$  与  $\sigma_i$ ,

$$A_i = 1 - E_i = 1 - \frac{1}{H} \sum_{h=1}^H \varepsilon_{ih} = 1 - \frac{1}{H} \sum_{h=1}^H \left( \frac{|y_{ih} - \hat{y}_{ih}|}{y_{ih}} \right) \quad (7)$$

$$\sigma_i = \sqrt{\frac{1}{H} \sum_{h=1}^H [(1 - \varepsilon_{ih}) - A_i]^2} \quad (8)$$

依据准确率及其标准差设计每个支持向量机的权重为:

$$w_k = \frac{\sum_{i=1}^P A_{ki} (1 - \sigma_{ki})}{\sum_{k=1}^3 \sum_{i=1}^P A_{ki} (1 - \sigma_{ki})} \quad (9)$$

记不同核函数支持向量机为  $\hat{y}_k$ , 由  $\hat{y}_k$  变权综合得到的综合模型可表示为  $\hat{y}^*$ ,

$$\hat{y}^* = \sum_{k=1}^3 w_k \hat{y}_k \circ$$

### 3.6. 基于不同核函数 SVM 变权综合的负荷预测

基于不同核函数核支持向量机变权综合的短期负荷预测方法的实现可总结为以下几个步骤:

Step 1: 确定给定的历史数据集的特征展开参数  $m$ , 将给定时刻  $t$  前  $m$  个负荷数据组成电力负荷特征向量  $X_t$ ,  $t = 1, 2, \dots, N - m$ 。

Step 2: 设待预测时刻  $t = t_z$ ,  $z = 1, \dots, Z$ , 确定负荷预测备选特征集  $F_t$ , 采用相关度分析进行特征选取, 组建特征集  $F_t^\alpha$ , 进而对历史数据集进行关系化展开, 构建待预测时刻的预测空间  $\Omega$ 。

Step 3: 选用高斯径向基核函数、拉普拉斯核函数及多项式核函数的支持向量机在预测空间上建立预测模型, 借助  $P$  折交叉验证进行模型测试。

Step 4: 依据性能测试结果的准确度  $A_i$  及其标准差  $\sigma_i$  构造变权, 通过变权综合不同核函数的 SVM 实现  $t_z$  的负荷预测。令  $z = z + 1$ , 当  $z > Z$ , 则预测结束; 否则, 转 Step 2。

#### 4. 算例分析

选用东肯塔基州电力合作社的 2018 年采样间隔为 1 小时的电力负荷数据进行实例分析, R 语言作为算例分析的实现工具。取 1 月 1 日~4 月 23 日的电力负荷时间序列作为训练数据, 取值论域  $X(f) = [800, 3500]$ , 单位: MW。分别采用高斯核函数、多项式核函数以及拉普拉斯核函数的支持向量机进行建模, 预测 4 月 24 日~30 日(记为  $T_1, T_2, \dots, T_7$ ) 的电力负荷值。以预测  $T_1$  第一时刻  $t_1$  的电力负荷过程为例, 其余时刻点负荷以相同步骤滚动预测获得。

计算  $t_1$  的负荷特征  $f_0$  与历史负荷特征  $f_1, f_2, \dots, f_k, \dots$  之间的加权相关度, 结果如图 1 所示。

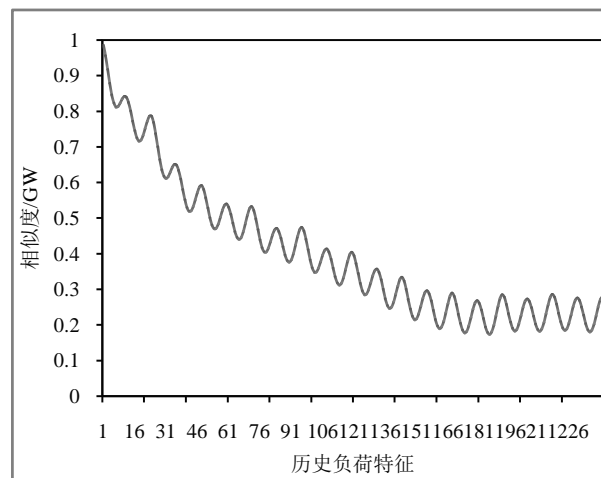


Figure 1. Correlation between load feature to be forecast and historical load features

图 1. 待预测负荷特征与历史负荷特征的相关度

上图所示, 待预测负荷特征与历史负荷特征之间的相似度呈衰减趋势。确定阈值  $\alpha = 0.8$ , 则选出的特征组成的相关特征集  $F_t^\alpha(t) = \{f_i, i = 1, \dots, 14\}$ 。以  $f_0$  为输出, 特征集中的特征为输入, 构建关于待预测时刻  $t_1$  的预测相空间  $\Omega$ 。通过将历史数据输入输出关系化, 得到关系化数据集。

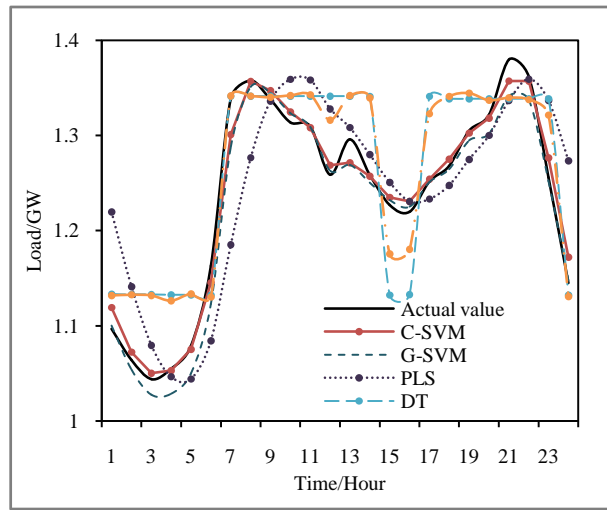
分别采用不同核函数的支持向量机在训练集上建立模模型, 并借助十折交叉验证对支持向量机进行测试, 依据测试得到的准确度及其标准差计算权重, 综合不同核函数支持向量机的预测结果, 得到预测值 1119.452 MW。

更新训练集与测试集以滚动预测其它时刻的电力负荷。采用预测模型的相平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)以及相对误差的标准差(SRE)作为评价指标进行评价, 其中

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (10)$$

$$SRE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% - MAPE \right)^2} \quad (11)$$

对变权综合模型(C-SVM)与常用的高斯核函数 SVM (G-SVM)、偏最小二乘(PLS)、决策树(DT)以及 Bagging 等预测方法进行对比分析。T<sub>1</sub> 的预测结果如图 2 所示。



**Figure 2.** Comparison of power load forecasting values  
**图 2.** 电力负荷预测值对比

如图 2 所示, 对比几种方法预测值的曲线与真实值曲线的差异, C-SVM 预测值的曲线与真实真实值曲线最为贴近, 其次是 G-SVM, 但在谷处的预测不及 C-SVM, 偏离了真实值。DT 与 Bagging 预测值曲线围绕较真实值曲线上下波动, 但与真实值曲线偏离最为明显。

对比以上模型预测 24~30 日负荷得到的 MAPE 值, 如表 1 所示。

**Table 1.** MAPE of forecast models  
**表 1.** 模型预测所得的 MAPE

	C-SVM	G-SVM	PLS	DT	Bagging
T <sub>1</sub>	0.900	1.196	4.014	4.162	3.656
T <sub>2</sub>	0.877	1.239	3.800	4.991	4.489
T <sub>3</sub>	1.376	1.623	4.729	4.625	3.830
T <sub>4</sub>	1.373	1.820	4.285	4.743	4.124
T <sub>5</sub>	0.988	1.549	3.552	5.556	4.641
T <sub>6</sub>	1.435	1.859	4.693	3.567	3.242
T <sub>7</sub>	2.171	2.505	5.602	3.793	3.340
Avg	1.303	1.685	4.382	4.491	3.903
改进	—	0.382	3.079	3.188	2.600
Max	2.171	2.505	5.602	5.556	4.641
改进	—	0.334	3.431	3.385	2.470

注: 表中数据均为百分数。

如表 1 所示, C-SVM 所得 MAPE 普遍小于其它方法的 MAPE, MAPE 的均值为 1.303%, 最大值为 2.171%。对比 G-SVM、PLS、DT 以及 Bagging, C-SVM 将预测准确度分别改进了 0.382%、3.079%、3.188%、2.6%; 并将最大 MAPE 值分别改进了 0.334%、3.431%、3.385% 以及 2.47%。

对比以上模型预测 24~30 日负荷所得的 MRE 值, 结果如表 2 所示。

**Table 2.** *SRE* of forecast models**表 2.** 模型预测所得的 *SRE*

	C-SVM	G-SVM	PLS	DT	Bagging
T <sub>1</sub>	0.791	1.127	3.497	2.665	2.223
T <sub>2</sub>	0.680	1.063	3.273	2.867	2.665
T <sub>3</sub>	0.777	1.233	3.547	2.545	1.865
T <sub>4</sub>	0.921	1.261	3.896	2.096	1.924
T <sub>5</sub>	0.742	1.234	2.155	3.667	3.375
T <sub>6</sub>	0.837	1.064	2.167	2.439	2.095
T <sub>7</sub>	1.074	1.520	4.454	2.008	1.685
Avg	0.832	1.215	3.284	2.612	2.261
改进	—	0.383	2.452	1.781	1.430

注：表中数据均为百分数。

如表 2 所示，C-SVM 的 *SRE* 值大部分小于 1%，可知 C-SVM 在预测过程波动幅度不大，预测相对稳定。与 G-SVM、PLS、DT 以及 Bagging 的 *SRE* 均值相比，C-SVM 将预测稳定性分别改进了 0.383%、2.452%、1.781% 以及 1.43%。

## 5. 总结

本文针对短期负荷预测准确度提升及预测稳定性改进问题，建立了基于不同核函数支持向量机变权综合的预测模型。利用相关分析实现了特征选取，采用交叉验证进行了模型性能测试，结合测试结果构造了变权，通过变权融合实现了基于综合模型的负荷预测。东肯塔基州电力合作社的实例分析结果证明，与高斯核函数的 SVM、偏最小二乘、决策树以及 Bagging 等单一模型相比，新方法将体现准确度的 *MAPE* 均值分别改进了 0.382%、3.079%、3.188%、2.6%；将最大 *MAPE* 值分别改进了 1.52%、4.62%、41.34%、3.3% 以及 4.43%；将最大相对误差的均值改进了 0.334%、3.431%、3.385% 以及 2.47%。将体现稳定性的 *SRE* 均值分别改进了 0.383%、2.452%、1.781% 以及 1.43%。可见，多模型变权综合是一种提升负荷预测准确度与稳定性的有效方式，今后，可对变权设计及综合方式进行深入研究。

## 基金项目

新疆财经大学研究生科研创新项目(XJUFE2018K043)。

## 参考文献

- [1] 康重庆, 夏清, 刘梅. 电力系统负荷预测[M]. 北京: 中国电力出版社, 2017, 1-17.
- [2] Khan, A.R., Mahmood, A., Safdar, A., *et al.* (2016) Load Forecasting, Dynamic Pricing and DSM in Smart Grid: A Review. *Renewable & Sustainable Energy Reviews*, **54**, 1311-1322. <https://doi.org/10.1016/j.rser.2015.10.117>
- [3] 顾熹, 廖志伟. 基于相空间重构和高斯过程回归的短期负荷预测[J]. 电力系统保护与控制, 2017, 45(5): 73-79.
- [4] Chen, J.F., Wang, W.M. and Huang, C.M. (1995) Analysis of an Adaptive Time-Series Autoregressive Moving-Average (ARMA) Model for Short-Term Load Forecasting. *Electric Power Systems Research*, **34**, 187-196. [https://doi.org/10.1016/0378-7796\(95\)00977-1](https://doi.org/10.1016/0378-7796(95)00977-1)
- [5] 张伏生, 汪鸿, 韩悌, 等. 基于偏最小二乘回归分析的短期负荷预测[J]. 电网技术, 2003, 27(3): 36-40.
- [6] 万坤, 柳瑞禹. 区间时间序列向量自回归模型在短期电力负荷预测中的应用[J]. 电网技术, 2012, 36(11): 77-81.
- [7] 魏立兵, 赵峰, 王思华. 基于人群搜索算法优化参数的支持向量机短期电力负荷预测[J]. 电测与仪表, 2015, 53(8): 45-49.
- [8] Zhang, X.B., Wang, J.Z. and Zhang, K.Q. (2017) Short-Term Electric Load Forecasting Based on Singular Spectrum

---

Analysis and Support Vector Machine Optimized by Cuckoo Search Algorithm. *Electric Power Systems Research*, **146**, 270-285. <https://doi.org/10.1016/j.epsr.2017.01.035>

- [9] 梁智, 孙国强, 等. 基于 VMD 与 PSO 优化深度信念网络的短期负荷预测[J]. 电网技术, 2018, 42(2): 598-606.
- [10] 吴润泽, 包正睿, 宋雪莹, 等. 基于深度学习的电网短期负荷预测方法研究[J]. 现代电力, 2018, 35(2): 43-48.
- [11] 杨芳君, 王耀力, 王力波, 等. 基于改进 CS 算法优化 Elman-IOC 神经网络的短期负荷预测[J]. 电测与仪表, 2019, 56(9): 39-44.
- [12] Sudheer, G. and Suseelatha, A. (2015) Short Term Load Forecasting Using Wavelet Transform Combined with Holt-Winters and Weighted Nearest Neighbor Models. *International Journal of Electrical Power & Energy Systems*, **64**, 340-346. <https://doi.org/10.1016/j.ijepes.2014.07.043>
- [13] Li, S., Goel, L., and Wang, P. (2016) An Ensemble Approach for Short-Term Load Forecasting by Extreme Learning Machine. *Applied Energy*, **170**, 22-29. <https://doi.org/10.1016/j.apenergy.2016.02.114>
- [14] Wang, L., Lee, E.W.M. and Yuen, R.K.K. (2019) Novel Dynamic Forecasting Model for Building Cooling Loads Combining an Artificial Neural Network and an Ensemble Approach. *Applied Energy*, **228**, 1740-1753. <https://doi.org/10.1016/j.apenergy.2018.07.085>
- [15] 陆继翔, 张琪培, 等. 基于 CNN-LSTM 混合神经网络模型的短期负荷预测方法[J]. 电力系统自动化, 2019, 43(8): 191-197.