

# 基于MM-LSTM的非工楼宇中央空调负荷预测

杨斌<sup>1</sup>, 陈振宇<sup>1</sup>, 霍尧<sup>1</sup>, 阮文骏<sup>1</sup>, 罗玉孙<sup>2</sup>

<sup>1</sup>国网江苏电力有限公司, 江苏 南京

<sup>2</sup>北京慧和仕科技有限责任公司, 北京

Email: yangbin@js.sgcc.com.cn

收稿日期: 2021年5月28日; 录用日期: 2021年6月12日; 发布日期: 2021年6月24日

## 摘要

空调负荷已经成为电网负荷中非常重要的组成部分, 特别是在非工楼宇空调, 对非工楼宇空调负荷进行准确负荷预测, 可以进行更加合理的需求响应, 提高区域用电质量。MM-LSTM (多路多层LSTM网络) 预测方法对非工楼宇中央空调的用电负荷进行预测精度较高。通过数据实验表明, 该负荷预测方法在适当的数据特征工程后, 预测普遍优于一般LSTM网络的空调负荷预测方法, 具有良好的预测精度。该方法可有效满足非工楼宇空调需求响应的需要, 具有良好的实用价值。

## 关键词

中央空调负荷预测, 多路多层LSTM, 神经网络, 非工楼宇

# Central Air Conditioning Load Forecasting for Non-Industrial Buildings Based on MM-LSTM

Bin Yang<sup>1</sup>, Zhenyu Chen<sup>1</sup>, Yao Huo<sup>1</sup>, Wenjun Ruan<sup>1</sup>, Yusun Luo<sup>2</sup>

<sup>1</sup>State Grid Jiangsu Electric Power Company, Nanjing Jiangsu

<sup>2</sup>Wellsteeh Limited Company, Beijing

Email: yangbin@js.sgcc.com.cn

Received: May 28<sup>th</sup>, 2021; accepted: Jun. 12<sup>th</sup>, 2021; published: Jun. 24<sup>th</sup>, 2021

## Abstract

Air-conditioning load has become a very important part of the grid load, especially in non-industrial buildings. Accurate load forecasting of non-industrial building air-conditioning loads can perform

文章引用: 杨斌, 陈振宇, 霍尧, 阮文骏, 罗玉孙. 基于 MM-LSTM 的非工楼宇中央空调负荷预测[J]. 智能电网, 2021, 11(3): 288-295. DOI: 10.12677/sg.2021.113027

more reasonable demand response and improve regional power quality. MM-LSTM (multi-channel & multi-layer LSTM network) prediction method is used to predict the load of non-industrial building air conditioners in a short time, and the high-precision building air conditioning load prediction results are obtained. The experimental data shows that the proposed method is superior to the general LSTM network method under the premise of suitable feature engineering, and has good prediction accuracy. The method can effectively meet the needs of non-industrial building air conditioning demand response and has good practical value.

## Keywords

Central Air Conditioning Load Forecasting, MM-LSTM, Neural Network, Non-Industrial Buildings

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

## 1. 引言

在经济高速发展的大趋势下，人们对高质量生活水平的追求越来越高，空调用电负荷逐渐成为电网负荷中非常重要的组成部分，特别是非工楼宇空调。2017年江苏省的最高用电负荷达到10,219万千瓦，其中空调负荷占比超过40%。随着社会经济的发展以及人们生活水平的不断提高，空调负荷占比还将继续上升。对非工楼宇空调负荷进行准确负荷预测，可以进行更加合理的需求响应，提高区域用电质量[1][2]。

国内外学者对电力负荷预测方法及空调用电负荷预测方法做了大量的研究，也得了许多研究成果。当前，学者们对空调负荷预测通常采用的方法包括：多元分析法、时序分析法、灰色理论法以及人工神经网络法。在空调负荷预测方面，Forrester等在1984年采用多元线性回归预测方法，对某大型商业建筑的短期用电负荷进行预测，取得了不错的效果[3]。李爱旗，2006年对重庆市城镇住宅建筑室内热环境进行分析和负荷预测，但把未考虑因素当成固定因素，有一点局限性[4]。谢艳群，2007年采用问卷调查对长沙市住宅建筑的冬、夏季室内环境和家庭能耗进行统计分析，使用多元线性分析法建立影响能耗的回归模型，并提出了一些住宅建筑节能措施，但存在数据分析方面的不足[5]。Macarthur等人于1989年提出在线回归估计负载剖面预测的方法，指出温度和辐射的预测算法是简单和有效的，但存在预测精度不够准确的问题[6]。何大四、张旭在2004年采用季节性指数平滑预测模型，根据空调负荷和温度之间的关系，对上海市某医学中心周内的空调用电负荷进行预测，实证表明季节性指数平滑预测模型不仅结构简单，同时在空调用电负荷方面有一定的效果[7]。孙靖、程大章，2004年采用时间序列法对上海市某医学中心2002年的空调负荷进行预测，结果表明季节性时间序列模型结构不仅简单，预测精度也比较理想[8]。孙光伟等人，1999年采用BP网络对某办公楼的空调负荷预测的建模方法进行研究，并对神经网络的输入进行了比较研究，结果表明，将前几天的负荷数据、日气象参数和预测目的特性作为网络输入时，预测精度最高[9]。石磊等人在2003年采用人工神经网络模型，对西安市楼宇用电负荷进行预测，结果表明神经网络预测模型对楼宇冷负荷预测具有较强的适应性[10]。王蕾等在2014年构建了基于粒子群与控制误差回馈的BP神经网络预测技术，其预测精度较BP神经网络和粒子群BP神经网络分别提高5.94%和0.82% [11]。在电力负荷预测方面，Vermaak和Botha在2002年将RNN用于短期负荷预测，实证表明RNN模型可以较好地捕获输入数据的特性，同时保留历史信息的影响[12]。Marino等在2016年采用标准长短期记忆网络(LSTM)和S2S-LSTM对个人住宅的负荷进行预测，实证结果表明，相对于标准LSTM模型，S2S-LSTM在长期和短期预测方面均

表现良好[13]。

综合以上，不同学者在负荷预测方法的研究都有其各自的优势，但是其模型的应用也具有一定局限性。在对空调系统进行负荷预测时，选择合适的预测方法，不仅需要考虑预测的精准性，同时也要照顾到对象所在的背景条件。不同空调负荷预测模型所纳入的影响因子不同，决定空调运行采集端的测量点的部署，过多的测量点造成前期工程成本过大，同时其后期的软硬件的维护成本也会较高，给用户造成一定负担，给模型的应用性造成比较大的困难。

据此，本文提出一种基于 MM-LSTM (多路多层 LSTM 网络)的非工楼宇中央空调负荷预测模型。在模型的输入端，采用多路 LSTM 对不同类型的影响因子进行处理，对多路输入进行合并后采用多层 LSTM 进行非工楼宇中央空调负荷预测，解决了单序列 LSTM 网络存在模型过拟合严重、泛化能力差等问题。

## 2. 长短期记忆网络(LSTM)原理

长短期记忆网络(LSTM)是一种改进的时间循环神经网络(RNN)，LSTM 由 Hochreiter & Schmidhuber (1997)提出[14]，并由 Alex Graves 进行了改良和推广[15][16][17]。LSTM 可以学习时间序列长短期依赖信息，由于 LSTM 网络中包含时间记忆单元，因此适用于处理和预测时间序列。

### 2.1. 循环神经网络

循环神经网络(RNN)是一种改进的多层感知器网络(见图 1)，包括输入层、隐藏层、输出层，在隐藏层中有输入到下一时间隐藏层的连接。按时间展开后，可以发现当前时刻隐藏层的输入包括两部分，即当前时刻的输入层输入和上一时刻的隐藏层输入。

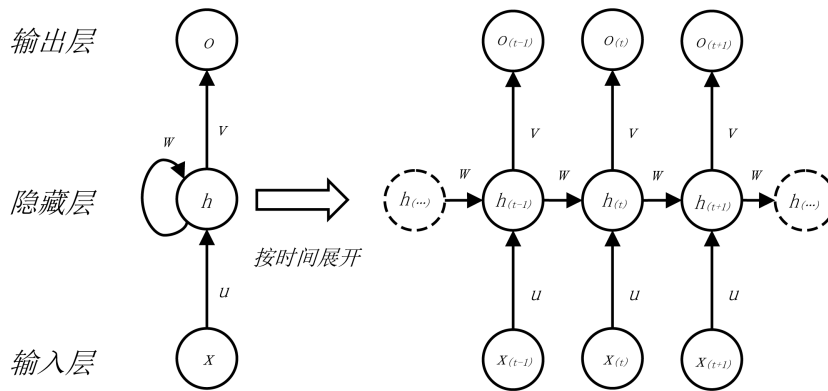


Figure 1. Structure of RNN  
图 1. 循环神经网络结构图

$x$  是输入层， $h$  是隐藏层， $o$  为输出层； $V$ 、 $W$ 、 $U$  是权值，同一类型的权连接权值相同。在前向传播过程中，当前层按照网络连接及权值对输入数据进行计算并输出到下一层，计算过程为：

$$h(t) = f(x(t) \cdot u + h(t-1) \cdot w + b) \tag{1}$$

$$o(t) = h(t) \cdot v + c \tag{2}$$

式中  $x(t)$  是在时刻  $t$  的输入层输入， $h(t-1)$  是在时刻  $t$  的隐藏层输入， $b$  为隐藏层偏置， $f(\cdot)$  为激活函数， $h(t)$  为  $t$  时刻的隐藏层输出和输出层输入， $c$  为输出层偏置， $o(t)$  为  $t$  时刻输出层输出。循环神经网络计算过程中虽然加入了上一时刻的输入，但是随着时间的推移，后面节点对前面时间节点的感知能力下降，即存在 RNN 的梯度消失问题。

## 2.2. 长短期记忆网络

鉴于 RNN 只能对较短的时间序列数据进行记忆，引起梯度消失而导致预测模型失效。LSTM 可以提供一种有效的解决方案，该网络通过传递转移函数来处理输入的时间序列，LSTM 在保持 RNN 网络结构的基础上重新设计了隐藏层，相比于 RNN，LSTM 增加了细胞状态 C (Cell State)，输入门，遗忘门和输出门，细胞状态包含了历史信息，输入门决定隐藏层信息是否进行更新，遗忘门决定更新后的信息是否包含上一时刻的信息，输出门决定输出更新后信息的哪个部分，LSTM 的隐藏层如图 2 所示。

LSTM 隐藏层“门”由神经网络层(图中灰色部分，如 sigmoid 层和 tan 层)和运算操作(图中圆形和椭圆部分，如加法、乘法和 tanh 函数)构成；细胞状态保存了历史信息，代表了长期记忆；通过控制各种“门”删除或增加信息来更新节点状态。

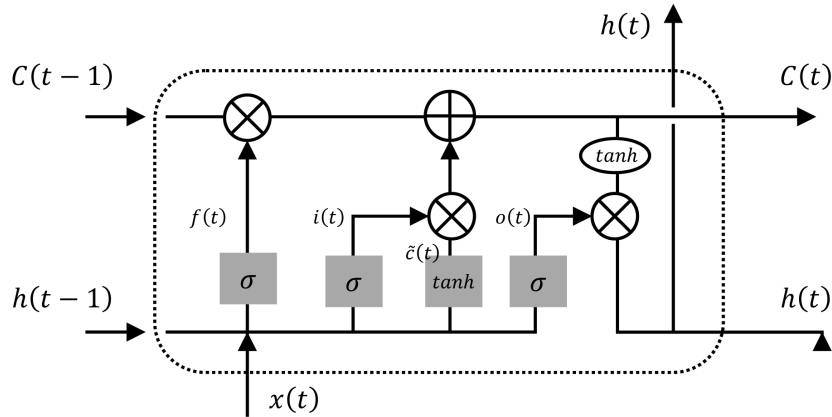


Figure 2. Structure of LSTM  
图 2. 长短期记忆网络隐藏节点结构图

### 2.2.1. 遗忘门

LSTM 中的第一步是决定我们从细胞状态中舍弃什么信息，这个决定通过一个称为遗忘门完成，遗忘门计算过程为：

$$f(t) = \sigma(W_f \cdot (h(t-1), x(t)) + b_f) \quad (3)$$

公式中  $h(t-1)$  是在时刻  $t$  时的隐藏层输入(下同)， $x(t)$  是在时刻  $t$  的输入层输入(下同)， $b_f$  是遗忘门偏置， $\sigma(\cdot)$  是遗忘门激活函数， $f(t)$  是遗忘门输出；遗忘门通过乘法运算，决定网络记住细胞状态的历史信息比例，激活函数  $\sigma(\cdot)$  使输出结果在 0 到 1 之间，1 表示“完全保留”，0 表示“完全舍弃”。

### 2.2.2. 输入门

LSTM 中的第二步是确定新信息被存放在细胞状态中，通过一个输入门来完成，这里包含两个部分，一是 sigmoid 层决定什么值我们将要更新，二是 tanh 层创建一个新的细胞状态候选值加入到状态中。遗忘门计算过程为：

$$i(t) = \sigma(W_i \cdot (h(t-1), x(t)) + b_i) \quad (4)$$

$$\tilde{C}(t) = \tanh(W_c \cdot (h(t-1), x(t)) + b_c) \quad (5)$$

公式中  $b_i$  为输入门 sigmoid 层偏置， $i(t)$  是输入门 sigmoid 层输出， $b_c$  输入门 tanh 层偏置， $\tanh(\cdot)$  是输入门 tanh 层激活函数， $\tilde{C}(t)$  是输入门 tanh 层输出；输入门通过乘法运算得到  $i(t) \cdot \tilde{C}(t)$ ，即为新的细胞状态候选值。通过遗忘门舍弃部分历史信息和输入门新增新的信息，我们可以更新细胞状态，计算过程为：

$$C(t) = f(t) \cdot C(t-1) + i(t) \cdot \tilde{C}(t) \tag{6}$$

### 2.2.3. 输出门

LSTM 中的最后一步是确定输出信息，通过一个输出门来完成，首先要确定输出哪些内容，通过一个 sigmoid 层来完成，接着用 tanh 函数处理状态信息并将它和 sigmoid 层的输出相乘来确定输出的内容，输出门计算过程为：

$$o(t) = \sigma(W_o \cdot (h(t-1), x(t)) + b_o) \tag{7}$$

$$h(t) = o(t) \cdot \tanh(C(t)) \tag{8}$$

公式中  $b_o$  为输出门 sigmoid 层偏置， $o(t)$  为输出门 sigmoid 层输出， $h(t)$  为输出门输出，也是时刻  $t$  时的隐藏层输出。

## 2.3. 多路多层 LSTM

本文提出的 MM-LSTM (多路多层 LSTM 网络)结构如下图 3：

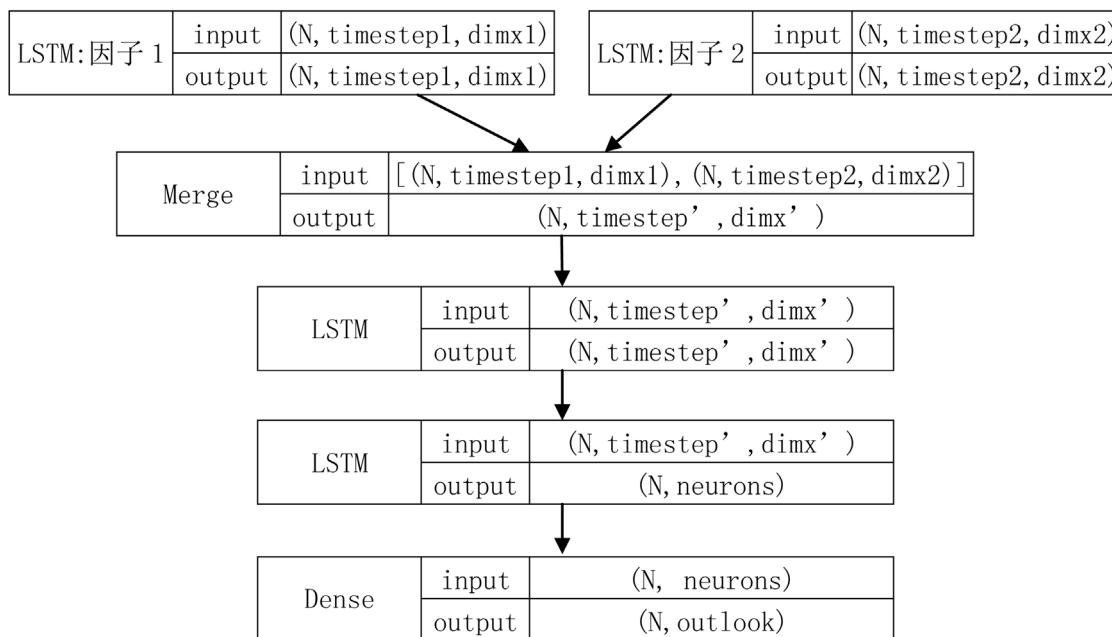


Figure 3. Structure of MM-LSTM

图 3. MM-LSTM 网络结构图

在本模型中，不同的影响因子组构成的两路输入序列，通过两个 LSTM 被编码为特征向量两路特征向量被 Merge 层组合在一起，输入到下一层的 LSTM 模型；在该模型中，我们将两路三层 LSTM 堆叠在一起，是该模型能够学习更高层次的时域特征表示。开始的两层 LSTM 返回其全部输出序列，而第三层 LSTM 只返回其输出序列的最后一步结果，从而其时域维度降低(即将输入序列转换为单个向量)。

## 3. 多路多层 LSTM 模型的空调负荷预测

根据使用空调负荷单序列数据训练 LSTM 网络预测空调负荷时，存在模型过拟合严重、泛化能力差等问题。为了解决此类问题，本文提出了利用多类型外部因子序列与多层 LSTM 相结合的方案，基于 MM-LSTM 模型的空调负荷预测模型。

### 3.1. 数据特征工程

#### 3.1.1. 数据标准化

各指标值在标准化后都处于同一个数量级别上, 让神经网络模型不同维度之间的特征在数值上有比较性, 可以提高神经网络预测的准确性; 另外, 标准化还可以提升模型的收敛速度以及防止模型梯度爆炸。

##### 1) 线性标准化

常用的线性标准化方法有 Min-max 标准化(极差标准化)和 Z-score 标准化(标准分数)。

##### a) Min-max 标准化

Min-max 标准化是对原始数据的线性变换, 使结果落到[0, 1]区间, 转换函数为:

$\tilde{x} = (x - x_{\min}) / (x_{\max} - x_{\min})$ ; 其中  $x_{\max}$  为样本数据的最大值,  $x_{\min}$  为样本数据的最小值; 如果想要将数据映射到[-1, 1], 则通过  $2 \cdot \tilde{x} - 1$  进行转换; 或近似值  $2 \cdot (x - x_{\text{mean}}) / (x_{\max} - x_{\min})$ , 其中  $x_{\text{mean}}$  为样本均值。这种方法有一个缺陷就是当有新数据加入时, 可能导致  $x_{\max}$  和  $x_{\min}$  的变化, 需要重新定义, 否则映射值会超出设定区间。

##### b) Z-score 标准化

Z-score 标准化也叫标准差标准化, 这种方法通过原始数据的均值和标准差进行数据的标准化。经过处理的数据符合标准正态分布, 即均值为 0, 标准差为 1, 转换函数为  $\tilde{x} = (x - \mu) / \sigma$ , 其中  $\mu$  为样本均值,  $\sigma$  为样本标准差; Z-score 标准化方法适用于样本的最大值和最小值未知的情况, 或有超出取值范围的离群数据的情况。该种归一化方式要求原始数据的分布可以近似为高斯分布, 否则归一化的效果会变得很差。

##### 2) 非线性标准化

在处理非均匀分布的样本时, 尤其是样本头尾局部集中数据的处理存在一定局限性, 无法提升数据辨识度。刘学之等 2008 年提出了利用 Logistic 曲线函数的特性构建 S 型曲线模型, 对数据进行非线性标准化处理, 在不改变数据序列及整体分布的前提下对各数据点的取值进行非线性放缩, 在将数据稠密的区间进行扩大的同时将数据稀疏的区间进行压缩[18]。转换函数为  $\tilde{x} = 1 / (1 + e^{-\alpha(x-\beta)})$ , 经过 Logistic 函数将数据映射到[0, 1], 其中  $\alpha$  和  $\beta$  为待定系数, 由样本中实际数据确定, 确定方法如下:

a)  $\beta = \text{media}(X)$ , 即样本的中位数;

b)  $\alpha = (\ln 9) / t$ , 其中  $t = \min(|X_{90\%} - \beta|, |X_{10\%} - \beta|)$ , 即为 90% 分位点到中位点的距离和 10% 分位点到中位点的距离的最小值。

#### 3.1.2. One-Hot 编码

神经网络模型是基于向量空间中的度量来进行计算的, 为了使非偏序关系的变量取值不具有偏序性, 并且到圆点是等距的。使用 one-hot 编码, 将离散特征的取值扩展到了欧式空间, 离散特征的某个取值就对应欧式空间的某个点。将离散型特征使用 one-hot 编码, 会让特征之间的距离计算更加合理。如工作日、节假日、风向等变量。

### 3.2. 数据集构建

模型前, 将特征工程后的样本数据划分为训练样本集、验证样本集和测试样本集 3 个部分。其中, 验证样本集的作用对训练后的网络进行评估。为了避免验证样本集过小或偶然性影响参数的调节和对模型的可靠评估, 采取了 K 折交叉验证法进行验证。具体做法为: 将数据划分为 K 个分区; 然后, 实例化 K 个相同的模型, 对于每个模型, 将 K-1 个分区作为其训练集, 余下的 1 个分区作为验证集; 最终, 将 K 次结果的均值作为对算法误差的验证结果, 本文 K 取 4。

## 4. 算例分析

本文通过基于 tensorflow 的 Keras 深度学习库搭建神经网络算法平台进行模型训练和预测。以南京市某楼宇 2017 年 5 月~9 月的数据作为研究样本,采用多路序列 LSTM 建立空调负荷预测模型,并进行负荷预测。将 2017 年 5 月~2018 年 9 月 23 日作为模型训练和验证样本,2017 年 9 月 24 日~30 日样本模拟模型预测,查看模型效果。

为了验证模型的效果,将模型的结果与普通 LSTM 进行对比,输入变量的选择如表 1 中所述:

**Table 1.** Input characteristics of air conditioning load forecasting  
**表 1.** 空调负荷预测模型输入特征

因子组	特征	特征描述	特征步长(天)
历史信息	Power	空调负荷	-31~-1
	temp_max	最高温度(历史温度)	-31~-1
	temp_min	最低温度(历史温度)	-31~-1
预报信息	temp_max	最高温度(预报温度)	1~7
	temp_min	最低温度(预报温度)	1~7
可获得信息	work_type	工作日类型(工作、节假、双休日)	1~7
	Month	月	1~7
	Day	年积日	1~7
	Week	星期	1~7
	Week_num	周数	1~7

本模型预测未来 7 天的空调负荷,将上表中的因子根据步长分成两组,一组为历史数据步长为历史的 31 期,一组为预测同步数据,为所需预测的时间同期数据 7 期;神经网络模型中第一层 LSTM 隐藏节点数即输出数据维度设定为两路输入数据数据维度的孰大值,四三层 LSTM 隐藏节点设定为 100。为了减少过拟合对预测模型的影响,添加 dropout 正则化,dropout 取值为 0.2,激活函数采用 tanh 函数,损失函数选择平均绝对误差(mean absolute deviation, MAE)。

模型预测精度比较如下表 2:

**Table 2.** Results comparison of MM-LSTM and LSTM  
**表 2.** MM-LSTM 与 LSTM 所得结果的比较

模型	误差(MAPE)
双路三层 LSTM	6.82%
LSTM	9.43%

由表中可知双路三层 LSTM 方法预测结果的预测精度高于 LSTM 方法。

## 5. 结语

本文通过对 LSTM 模型的剖析,得出其改进模型 MM-LSTM 模型。并且,使用 MM-LSTM 预测模型对非工楼宇空调用电负荷进行短期预测,得到了精度较高的预测结果。实测数据实验表明,MM-LSTM 模型在适当的特征工程后,预测效果普遍优于一般 LSTM 网络的空调负荷预测方法,具有良好的预测精度。该方法可有效满足非工楼宇空调需求响应的需要,具有良好的实用价值。

## 参考文献

- [1] 杨永标, 颜庆国, 徐石明, 肖文举, 薛铭枫. 公共楼宇空调负荷参与电网虚拟调峰的思考[J]. 电力系统自动化, 2015, 39(17): 103-107.
- [2] 王怡岚, 童亦斌, 黄梅, 杨利, 赵洪刚. 基于需求侧响应的空调负荷虚拟储能模型研究[J]. 电网技术, 2017, 41(2): 394-401.
- [3] Forrester, R.J. and Wepfer, W.J. (1984) Formulation of Prediction Algorithms for Management of Commercial and Industrial Energy Loads. *ASHRAE Transactions*, **85**, 234-239.
- [4] 李爱旗. 重庆市小城镇居住建筑热环境分析和建筑冷热负荷预测研究[D]: [硕士学位论文]. 重庆: 重庆大学, 2006.
- [5] 谢艳群. 长沙市住宅建筑能耗调查实测及其影响因素统计分析[D]: [硕士学位论文]. 长沙: 湖南大学, 2007.
- [6] Macarthur, J.W., Mathur, A. and Zhao, J. (1989) On-Line Recursive Estimation for Load Profile Prediction. *ASHRAE Transactions*, **95**, 621-628.
- [7] 何大四, 张旭. 改进的季节性指数平滑法预测空调负荷分析[J]. 同济大学学报, 2004, 33(12): 1672-1676.
- [8] 孙靖, 程大章. 基于季节性时间序列模型的空调负荷预测[J]. 电工技术学报, 2004, 19(3): 88-93.
- [9] 孙光伟, 索胜军, 陶永纯, 等. 基于 BP 网络的空调负荷预报建模方法的研究[J]. 哈尔滨建筑大学学报, 1999, 32(6): 79-82.
- [10] 石磊, 赵蕾, 王军, 等. 应用人工神经网络预测建筑物空调负荷[J]. 暖通空调, 2003, 33(1): 103-104, 113.
- [11] 王蕾, 张九根, 李腾, 等. 基于改进 BP 神经网络的中央空调冷负荷预测研究[J]. 计算机测量与控制, 2014, 22(6): 1690-1692.
- [12] Vermaak, J. and Botha, E.C. (1998) Recurrent Neural Networks for Short-Term Load Forecasting. *IEEE Transactions on Power Systems*, **13**, 126-132. <https://doi.org/10.1109/59.651623>
- [13] Marino, D.L., Amarasinghe, K. and Manic, M. (2016) Building Energy Load Forecasting Using Deep Neural. *IECON 2016: 42nd Annual Conference of the IEEE Industrial Electronics Society*, Florence, 23-26 October 2016 7046-7051. <https://doi.org/10.1109/IECON.2016.7793413>
- [14] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [15] Gers, F.A., Schmidhuber, J. and Cummins, F. (2000) Learning to Forget: Continual Prediction with LSTM. *Neural computation*, **12**, 2451-2471. <https://doi.org/10.1162/089976600300015015>
- [16] Gers, F.A., Schraudolph, N. and Schmidhuber, J. (2003) Learning Precise Timing with LSTM Recurrent Networks. *The Journal of Machine Learning Research*, **3**, 115-143.
- [17] Graves, A. (2012) Supervised Sequence Labelling with Recurrent Neural Networks. Vol. 385, Springer, Berlin. <https://doi.org/10.1007/978-3-642-24797-2>
- [18] 刘学之, 杨泽宇, 沈凤武, 尚胡佟, 刘嘉. 基于 S 型曲线的指标非线性标准化研究[J]. 统计与信息论坛, 2018, 33(2): 17-21.