

# 基于分布式优化思想的配电网 用电负荷多层协同预测方法

谭 嘉<sup>1</sup>, 李知艺<sup>1</sup>, 杨 欢<sup>1</sup>, 赵荣祥<sup>1</sup>, 鞠 平<sup>1, 2</sup>

(1. 浙江大学 电气工程学院, 杭州 310007;

2. 河海大学 能源与电气学院, 南京 211100)

**摘 要:** 当前, 分布式新能源、电动汽车等新元素在配电网中涌现, 改变了负荷的构成, 丰富了负荷的内涵, 给负荷预测带来了严峻挑战. 事实上, 用电负荷在配电网的多个电压层级以自下而上的方式聚合, 但现有预测研究鲜少考虑此类层级化特征. 为保障负荷自下而上的聚合一致性并且联合提升各层级的负荷预测性能, 提出了一种基于分布式优化算法的用电负荷多层协同预测方法. 首先, 采用基于交替方向乘子法的分布式优化理念, 构建了适配配电网层级特征、数据交互量少的多层协同预测框架. 随后, 提出了基于长短期记忆神经网络和联邦学习的具体预测方法, 通过将底层负荷预测结果逐级聚合, 能实现自下而上的配电网负荷一体化预测. 算例结果表明, 所提方法得到的用电负荷预测准确度高, 应用前景好.

**关键词:** 多层负荷预测; 交替方向乘子法; 长短期记忆神经网络; 联邦学习

**中图分类号:** TM 715

**文献标志码:** A

## A Multi-Level Collaborative Load Forecasting Method for Distribution Networks Based on Distributed Optimization

TAN Jia<sup>1</sup>, LI Zhiyi<sup>1</sup>, YANG Huan<sup>1</sup>, ZHAO Rongxiang<sup>1</sup>, JU Ping<sup>1, 2</sup>

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310007, China;

2. College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China)

**Abstract:** At present, new elements such as distributed new energy and electric vehicles have emerged in the distribution network, which changes the composition of loads, enriches the connotation of loads, and poses severe challenges to load forecasting. In fact, loads are aggregated in a bottom-up manner in multiple voltage levels of the distribution network, but such hierarchical characteristics are rarely considered in current load forecasting researches. Therefore, a multi-level load collaborative forecasting method based on the distributed optimization algorithm is proposed aimed at ensuring the bottom-up aggregation consistency of loads and jointly improving the performance of load forecasting at all levels. First, the distributed optimization concept based on the alternating direction method of multipliers is adopted to construct a multi-level load collaborative forecasting framework which adapts to the hierarchical characteristics of distribution network and has less data interaction. Then, a specific forecasting method based on the long short term-memory neural network and federated learning is proposed. By aggregating

收稿日期: 2021-07-30

基金项目: 国家自然科学基金(U2066601, 52007164)

作者简介: 谭 嘉(1997-), 女, 四川省达州市人, 硕士生, 从事人工智能在电力系统中的应用研究.

通信作者: 李知艺, 男, 研究员, 博士生导师, 电话(Tel.): 15658076963; E-mail: zhiyi@zju.edu.cn.

the bottom load forecasting results step by step, the bottom-up integrated load forecasting of distribution network can be realized. The results of calculation examples show that the proposed method has a high accuracy and a great application prospect.

**Key words:** multi-level load forecasting; alternating direction method of multipliers; long short-term memory neural network; federated learning

为加快推进我国“碳达峰、碳中和”的目标进程,构建以新能源为主体的新型电力系统是必由之路.随着新能源装机容量跨越式增长<sup>[1]</sup>,分布式新能源、储能、电动汽车等新元素在配电网层面涌现,丰富了负荷内涵,改变了负荷特性<sup>[2]</sup>.相比传统负荷,集成了这些新元素的配电网负荷具有更大的波动性和随机性,这将为配电网功率实时平衡带来巨大的挑战.

用电负荷预测(不包含网损)是保障电网安全和经济运行的基本前提之一<sup>[3]</sup>.智能电表、传感器等设备的广泛使用使得配电网积累了海量的、来源多样的历史负荷数据,为数据驱动的负荷预测技术奠定了必要的数据基础<sup>[4]</sup>.目前,负荷预测的方法主要有数学统计方法和机器学习方法<sup>[5]</sup>.数学统计方法主要包括回归分析法<sup>[6]</sup>、卡尔曼滤波法<sup>[7]</sup>、时间序列分析法<sup>[8]</sup>等.这类方法主要采用线性分析的思路,因此在处理负荷与其他因素(比如天气)非线性关系时具有一定的局限性,预测效果一般.从 20 世纪 80 年代起,支持向量机<sup>[9]</sup>、人工神经网络<sup>[10]</sup>等机器学习方法取得了长足发展,但这些方法在数据样本较多时普遍存在无法收敛的问题<sup>[11]</sup>.近年来,以长短期记忆(Long Short-Term Memory, LSTM)神经网络为代表的深度学习算法解决了时间序列的长期依赖问题,在负荷预测应用上取得了不错的效果.除了研究传统的单一对象负荷预测以外,部分学者同样致力于研究如何提高包含多个对象的负荷预测准确度<sup>[12-16]</sup>.尤其在电压层级的待预测对象具有层级化结构的情形下.例如,220 kV 高压输电变电站采集负荷与 110 kV 次级输电变电站采集负荷.针对层级化负荷结构,通常需要在不同层级分别进行负荷预测,从而形成完整的多层负荷预测结果.然而,该研究面临的主要挑战是如何保证产生的预测结果既在数学上满足单负荷预测准确度,又在物理上符合不同层级负荷之间的关联规律.具体来说,呈现层级化结构的上层负荷与下层负荷具有聚合一致性,即下层所有相关节点的负荷预测值之和应等于上层节点的负荷预测值,这样才能保证电力调控、负荷控制等决策的实现效果.而由于各层负荷功率在峰值大小、测量精度等方面存在差异,现阶段各层独

立预测的结果难以达到聚合一致性.传统的方法(如自下而上和自上而下方法)由于忽略了负荷预测偏差随层级结构的叠加性,难以协同提升各层级负荷的预测准确度.文献[17]将跨层级负荷预测建模为最优广义最小二乘法回归问题,然后通过梯度下降法求解,以改善底层负荷的预测结果,进而自下而上按层级结构累加汇总得到上层的负荷预测.文献[18]采用机器学习中的极端梯度提升机(XGBoost)和随机森林算法先对所有层级负荷分别进行预测,然后建立新的机器学习模型,将先前得到的底层和上层预测结果作为输入,并把修正后的底层负荷预测结果作为输出.

理论上,多层负荷协同预测可以看作一个以各层负荷预测总误差最小为目标,以聚合一致性为约束条件的分布式优化问题.以机器学习问题为例,分布式优化算法主要有分布式梯度下降算法、二阶优化算法、邻近梯度算法、坐标下降算法和交替方向乘子算法(Alternating Direction Method of Multipliers, ADMM)等,每种算法各有优劣<sup>[19]</sup>.其中,交替方向乘子法<sup>[20]</sup>因融合对偶分解和拉格朗日乘子法优点,在解决分布式优化问题时收敛性好,鲁棒性强,受到研究学者的关注.因此,本文选用 ADMM 算法作为求解多层负荷协同预测问题的算法框架.另外,由于底层负荷数据来源多,同步性差,难以在不同供电区域之间实时共享,即负荷量测和预测数据以孤岛的形式呈现,亟需一种新的框架来打破数据孤岛.联邦学习因为能解决数据孤岛问题而受到研究学者的广泛关注<sup>[21]</sup>,其于 2016 年由谷歌首次提出,并实现了数万台移动设备在终端数据不共享的情况下通过本地更新模型来完成学习任务<sup>[22]</sup>.由于联邦学习解决的问题与本文面临的供电区域数据孤岛问题相似,故考虑采用联邦学习的架构实现多层负荷协同预测模型的分布式训练.

鉴于上述分析,本文提出一种基于分布式优化思想的配电网负荷多层协同预测算法,旨在通过多个深度学习模型的分布式联邦训练,提升所有层级负荷的整体预测准确度.算例结果表明,与传统方法相比,本文方法在整体预测准确度上有明显的提升.

# 1 基于 ADMM 算法的多层负荷协同预测模型

## 1.1 模型构建方法

配电网多层级负荷预测以 2 级和 3 级负荷为主<sup>[23]</sup>. 例如,“220 kV 高压输电变电站-110 kV 次级输电变电站-10 kV 中压配变”属于 3 级负荷预测问题,“220 kV 高压输电变电站采集负荷-110 kV 次级输电变电站采集负荷”等问题属于 2 级负荷预测问题. 由于 3 级负荷层级结构具有现实普遍性且兼具理论研究难度,故本文以 3 级负荷层级结构为例,介绍基于 ADMM 算法的多层负荷协同预测模型.

典型的 3 级负荷层级结构如图 1 所示. 图中 L2 共有  $h$  个节点,第  $j$  个节点下属的子节点有  $m_j$  个,底层节点按顺序编号. 具体地, $M_1$  的子节点为供电区域  $1 \sim m_1$ , $M_2$  的子节点为供电区域  $m_1 + 1 \sim m_1 + m_2$ ,以此类推. 定义  $n$  为底层 L1 供电区域的总数,即  $n = \sum_{j=1}^h m_j$ . 从图 1 可以看出,负荷逐层累加,上层负荷包含了与之关联的下层负荷,父节点的负荷与子节点的负荷存在聚合一致性. 本文的研究目标即是利用负荷聚合一致性的条件,优化调整各层负荷预测结果,使所有层级负荷预测的总误差最小. 由于底层负荷曲线往往具有更高的信噪比,且相比上层负荷类型更丰富、数量更多,故底层负荷预测更容易受到分布式新能源、储能等配电网新元素的影响,预测误差比上层的预测误差更大<sup>[24]</sup>,成为影响总误差的主要原因. 为减小总误差,首先考虑优化调整底层负荷的预测结果. 为同时保证负荷预测结果的聚合一致性,将优化调整后的底层预测结果按层级结构聚合,进而实现自下而上的负荷一体化预测.

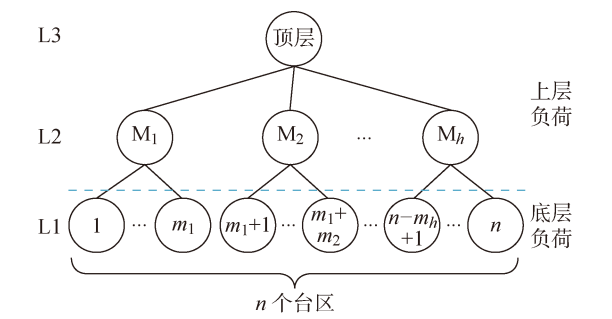


图 1 3 级负荷层级结构  
Fig. 1 Three-level load hierarchical structure

基于上述分析,搭建  $n$  个负荷预测模型,分别对应  $n$  个底层供电区域. 首先考虑底层区域的预测准确度,则整体的负荷预测目标函数定义为

$$\min \sum_{i=1}^n \| f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_i \|_2^2 \quad (1)$$

式中:  $f_i(\cdot)$  表示 L1 供电区域  $i$  的负荷预测结果;  $\boldsymbol{\theta}_i$  和  $\mathbf{x}_i$  分别表示模型参数和模型输入特征;  $\hat{\mathbf{y}}_i$  表示 L1 供电区域  $i$  的负荷实际值.

将底层供电区域的负荷预测结果逐级聚合,得到除 L1 以外的全部节点预测值. 需注意的是,要使得所有上层节点负荷预测误差最小,仅考虑 L2 各节点负荷预测约束条件是不够的. 例如,当得到的 L2 节点预测值均偏小,在聚合计算顶层负荷预测结果时,误差的叠加会导致顶层负荷预测误差更大. 因此,同时考虑顶层节点的预测约束条件也很有必要. 综上,结合负荷实际值,构造  $h+1$  个等式约束:

$$\left. \begin{aligned} \sum_{i=1}^{m_1} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) &= \hat{\mathbf{y}}_{M_1} \\ \sum_{i=m_1+1}^{m_1+m_2} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) &= \hat{\mathbf{y}}_{M_2} \\ &\vdots \\ \sum_{i=n-m_h+1}^n f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) &= \hat{\mathbf{y}}_{M_h} \\ \sum_{i=1}^n f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) &= \hat{\mathbf{y}}_{\text{top}} \end{aligned} \right\} \quad (2)$$

式中:  $\hat{\mathbf{y}}_{M_1}, \hat{\mathbf{y}}_{M_2}, \dots, \hat{\mathbf{y}}_{M_h}$  分别表示  $h$  个 L2 节点的负荷实际值;  $\hat{\mathbf{y}}_{\text{top}}$  表示顶层节点的负荷实际值. 需要注意的是,父节点(上层)负荷与子节点(下层)负荷在同一时间点才具有聚合一致性,因此所有的负荷实际值与模型输入特征需对应相同待预测日或时刻.

经过分析,本文把一个 3 级负荷预测问题描述成具有  $h+1$  个等式约束的  $n$  分块优化问题,然后采用 ADMM 算法求解.

## 1.2 分解协调方法

ADMM 是一种求解分布式优化问题的算法,其通过分解协调过程,将大的全局问题分解为多个较小的局部子问题,并通过协调子问题的解来得到全局问题的解. 在处理复杂的优化问题时,ADMM 算法具有较快的收敛速度,能得到更高质量的解,且在解决具有加和约束的优化问题上具有独特优势,适用于求解本文的约束优化问题<sup>[25]</sup>. 因此,对于本文 1.1 节所构建的具有等式约束的  $n$  分块优化问题,采用 ADMM 算法求解.

具体地,对于如式(1)、(2)所示的优化问题,采用 ADMM 算法引入对偶变量,根据拉格朗日乘子法,得到增广拉格朗日函数:

$$L_\rho(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \sum_{i=1}^n \| f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_i \|_2^2 +$$

$$\begin{aligned}
& \lambda_1 \psi \left( \sum_{i=1}^{m_1} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_{M_1} \right) + \\
& \frac{\rho}{2} \left\| \sum_{i=1}^{m_1} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_{M_1} \right\|_2^2 + \\
& \lambda_2 \psi \left( \sum_{i=m_1+1}^{m_1+m_2} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_{M_2} \right) + \\
& \frac{\rho}{2} \left\| \sum_{i=m_1+1}^{m_1+m_2} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_{M_2} \right\|_2^2 + \cdots + \\
& \lambda_{h+1} \psi \left( \sum_{i=1}^n f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_{\text{top}} \right) + \\
& \frac{\rho}{2} \left\| \sum_{i=1}^n f_i(\boldsymbol{\theta}_i; \mathbf{x}_i) - \hat{\mathbf{y}}_{\text{top}} \right\|_2^2 \quad (3)
\end{aligned}$$

式中:  $[\boldsymbol{\theta}_1 \ \boldsymbol{\theta}_2 \ \cdots \ \boldsymbol{\theta}_n]$  是  $n$  个模型的待优化参数;  $\lambda$  为拉格朗日乘子;  $\psi$  为求取向量所有元素均值的函数;  $\rho$  为惩罚因子. 对每个等式约束, 都引入一个拉格朗日乘子  $\lambda_j$ , 因此式(3)中  $\lambda$  共含有  $h+1$  项.

由于 ADMM 算法在更新单个模型参数时需固定其余模型, 故对于正在更新的模型来说,  $L_\rho(\boldsymbol{\theta}, \lambda)$  损失函数中其余模型预测的均方误差项可视为常数项. 为简化求解过程, 当更新某个模型参数时, 删去损失函数中与该模型不存在耦合关系的“常数项”, 均方误差项只含该模型预测的单个供电区域误差. 同理, 聚合负荷的误差项也只包含与该模型耦合的上层节点误差.

基于上述分析, 得到求解式(3)的 ADMM 迭代过程如下, 按顺序交替执行每个步骤:

$$\begin{aligned}
\boldsymbol{\theta}_1^{k+1} &= \arg \min_{\boldsymbol{\theta}_1} L_\rho(\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_{m_1}^k, \boldsymbol{\theta}_{m_1+1}^k, \cdots, \\
& \quad \boldsymbol{\theta}_n^k, \lambda_1^k, \lambda_{h+1}^k) \cdots \\
\boldsymbol{\theta}_{m_1}^{k+1} &= \arg \min_{\boldsymbol{\theta}_{m_1}} L_\rho(\boldsymbol{\theta}_1^{k+1}, \cdots, \boldsymbol{\theta}_{m_1}, \boldsymbol{\theta}_{m_1+1}^k, \cdots, \\
& \quad \boldsymbol{\theta}_n^k, \lambda_1^k, \lambda_{h+1}^k) \\
\boldsymbol{\theta}_{m_1+1}^{k+1} &= \arg \min_{\boldsymbol{\theta}_{m_1+1}} L_\rho(\boldsymbol{\theta}_1^{k+1}, \cdots, \boldsymbol{\theta}_{m_1}^{k+1}, \boldsymbol{\theta}_{m_1+1}, \cdots, \\
& \quad \boldsymbol{\theta}_n^k, \lambda_2^k, \lambda_{h+1}^k) \cdots \\
\boldsymbol{\theta}_n^{k+1} &= \arg \min_{\boldsymbol{\theta}_n} L_\rho(\boldsymbol{\theta}_1^{k+1}, \cdots, \boldsymbol{\theta}_{m_1}^{k+1}, \boldsymbol{\theta}_{m_1+1}^{k+1}, \cdots, \\
& \quad \boldsymbol{\theta}_n, \lambda_h^k, \lambda_{h+1}^k) \\
\lambda_1^{k+1} &= \lambda_1^k + \rho \psi \left( \sum_{i=1}^{m_1} f_i(\boldsymbol{\theta}_i^{k+1}; \mathbf{x}_i) - \hat{\mathbf{y}}_{M_1} \right) \\
\lambda_2^{k+1} &= \lambda_2^k + \rho \psi \left( \sum_{i=m_1+1}^{m_1+m_2} f_i(\boldsymbol{\theta}_i^{k+1}; \mathbf{x}_i) - \hat{\mathbf{y}}_{M_2} \right) \cdots \\
\lambda_{h+1}^{k+1} &= \lambda_{h+1}^k + \rho \psi \left( \sum_{i=1}^n f_i(\boldsymbol{\theta}_i^{k+1}; \mathbf{x}_i) - \hat{\mathbf{y}}_{\text{top}} \right)
\end{aligned}$$

式中:  $k$  为迭代次数. 可以看出, ADMM 迭代过程与对偶上升法相似, 变量  $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_n$  交替更新, 在更新一组变量时, 需固定其他变量和拉格朗日乘子,

这样在每次更新时只需求解单个变量, 缩小了问题规模. 在每轮训练过程中,  $n$  个模型各自进行分布式训练(即应用深度学习技术进行单点负荷预测), 但模型之间需要频繁地通信, 以传输对应的模型输出值. 算法迭代结束将会得到收敛的目标函数值、残差以及最优解  $\mathbf{p}^* = \{\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*, \cdots, \boldsymbol{\theta}_n^*\}$ .

## 2 LSTM-ADMM 多层负荷协同预测模型

### 2.1 基于联邦学习架构的模型训练流程

考虑到底层负荷的数据孤岛问题, 采用联邦学习的架构进行模型的交替训练和通信.

联邦学习目标是在保证数据隐私和不共享数据的情况下完成用户客户端的联合建模. 联邦学习根据是否需要第三方服务器分为两种, 分别是客户端-服务器架构<sup>[26]</sup>和对等网络架构<sup>[27]</sup>. 前者是让多个客户端利用本地数据训练本地模型, 通过中央服务器聚合各个客户端的模型参数并更新全局模型, 然后将全局模型参数下发给各个客户端, 继续用本地数据更新模型, 直到全局模型收敛为止. 后者在训练过程中, 不需要第三方服务器, 客户端之间可以直接通信. 为了降低通信开销, 选择对等网络架构来对  $n$  个底层供电区域进行联邦训练.

多层负荷协同预测模型的联邦学习架构如图 2 所示. 在第  $k+1$  次迭代过程中, 对  $n$  个模型按顺序依次进行更新. 当更新第  $i$  个模型时, 前  $i-1$  个模型已完成第  $k+1$  次更新, 因此输出值固定为第  $k+1$  次迭代的值; 第  $i+1 \sim n$  个模型由于还未进行第  $k+1$  次更新, 输出值固定为第  $k$  次迭代的值. 图 2 中的 ①、②、③ 表示更新第  $i$  个模型的一系列执行步骤, 其中步骤 ① 表示各模型根据输入特征得到输出值, 步骤 ② 表示模型  $i$  同步接收这  $n-1$  个模型的输出

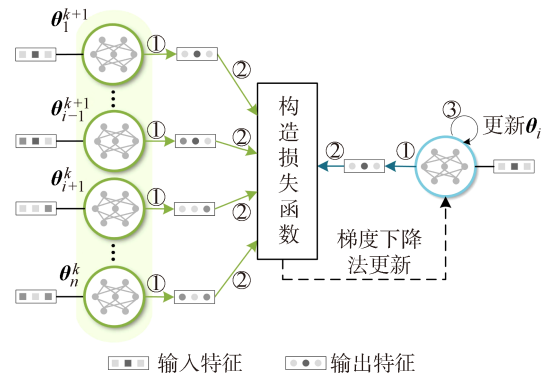


图 2 多层负荷协同预测模型的联邦学习架构

Fig. 2 Federated learning structure of multi-level load collaborative forecasting model

值并构造损失函数,步骤③则指采用梯度下降法更新本地模型参数.紧接着采用同样方法更新第  $i+1$ ,  $i+2, \dots, n$  个模型的参数,完成后进入第  $k+2$  次迭代,继续联邦训练,直到迭代次数达到设定值,算法迭代结束.

## 2.2 LSTM 网络与输入/输出特征构建

多层负荷预测模型的性能不仅与整体模型架构有关,还与单个供电区域负荷模型的预测性能息息相关.目前,应用于负荷预测研究且取得不错效果的深度学习算法有门控循环网络 (GRU)、深度残差网络 (ResNet)、卷积神经网络 (CNN)、LSTM 以及几种算法的组合模型,这些模型都可应用于本文提出的基于 ADMM 算法的多层负荷协同预测方法中.而且在具体的研究过程中,  $n$  个供电区域负荷模型可选取相同模型,也可以是异构模型.为使本文所提方法极具说服力,采用最具代表性的 LSTM 神经网络作为供电区域负荷预测的基础模型.

LSTM 网络是在循环神经网络基础上,通过引入遗忘门、输入门和输出门,并通过遗忘机制和保存机制剔除某些不重要信息,保留有用信息并存入神经元信息中<sup>[28]</sup>. LSTM 结构如图 3 所示,其特殊结构使得它有利于解决长时间序列对历史数据长期依赖的问题,并且可以显著改善训练过程中梯度消失及梯度爆炸问题<sup>[29]</sup>. LSTM 内部计算过程如下:

遗忘门:

$$f_t = \sigma(W_{hf}h_{t-1} + W_{xf}x_t + b_f) \quad (4)$$

输入门:

$$i_t = \sigma(W_{hi}h_{t-1} + W_{xi}x_t + b_i) \quad (5)$$

$$\tilde{c}_t = \tanh(W_{hg}h_{t-1} + W_{xg}x_t + b_g) \quad (6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (7)$$

输出门:

$$o_t = \sigma(W_{ho}h_{t-1} + W_{xo}x_t + b_o) \quad (8)$$

$$h_t = o_t \odot \tanh c_t \quad (9)$$

式中:  $f_t, i_t, o_t$  分别是遗忘门、输入门和输出门的输出值;  $[W_{hf} \ W_{xf}]$ 、 $[W_{hi} \ W_{xi}]$ 、 $[W_{hg} \ W_{xg}]$ 、 $[W_{ho} \ W_{xo}]$ 、 $b_f$ 、 $b_i$ 、 $b_g$ 、 $b_o$  分别是 4 个神经网络层的网络权重和网络偏置,这些权重和偏置构成了单个供电区域模型参数  $\theta_i$ ;  $\sigma$ 、 $\tanh$  分别是 Sigmoid 激活函数和  $\tanh$  激活函数;  $\odot$  是矩阵的哈达玛积运算符;  $h_{t-1}$  是  $t-1$  时刻的输出值;  $x_t$  是  $t$  时刻的输入值;  $c_{t-1}$  是  $t-1$  时刻的神经元状态;  $\tilde{c}_t$  是  $t$  时刻尚未经过处理的输入信息.通过遗忘门保存部分  $t-1$  时刻的神经元信息,输入门添加  $t$  时刻的输入信息,实现神经元状态的更新,  $c_t$  即为更新后  $t$  时刻的神经元

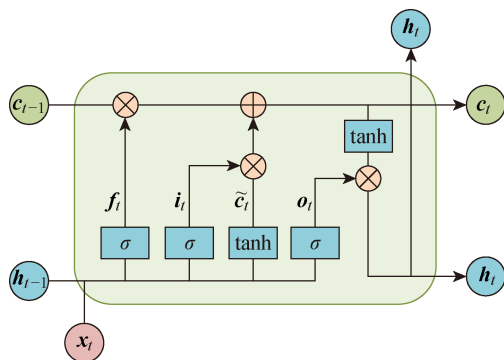


图 3 LSTM 结构

Fig. 3 Structure of LSTM

状态.最终,输出门控制更新后的神经元状态,从而得到输出值  $h_t$ .

确定 LSTM 网络的输入和输出特征是使模型获得优异性能的关键.其中输入特征通常是影响负荷变化的属性信息,例如气象因素、季节信息、日期信息(星期)等<sup>[30]</sup>.同时考虑到负荷数据具有周期性,待预测日当天的负荷与前一周的负荷以及 14 日前同一时刻的负荷有很大相关性.因此本文 LSTM 网络的输入特征包括历史负荷、温度、月份和星期,且所有的输入特征串联成向量作为模型的单个特征,输出特征即为待预测日当天的实际负荷数据.另外, LSTM 网络采用典型结构,即 1 个输入层、多个隐藏层和 1 个输出层的结构<sup>[31]</sup>.前一层的输出作为后一层的输入,最后通过全连接层输出得到负荷预测值.

## 2.3 多层负荷协同预测算法流程

基于分布式优化思想的多层负荷协同预测算法训练步骤如下.

**步骤 1** 搭建  $n$  个 LSTM 模型,各模型在本地独立预训练  $L$  次.

**步骤 2** 将 1.2 节构造的增广拉格朗日函数作为模型损失函数,采用梯度下降法并结合联邦学习架构,按照 ADMM 迭代顺序交替训练各模型,并更新拉格朗日乘子.

**步骤 3** 通过不断迭代训练各模型参数,最终模型学习到相应供电区域的负荷变化规律以及上下级负荷的聚合一致性.训练结束,保存各模型参数.

**步骤 4** 将测试集样本中底层区域的输入特征分别输入训练好的模型中,输出得到 L1 各区域的负荷预测结果:

$$L1: f_1(\theta_1; x_1), f_2(\theta_2; x_2), \dots, f_n(\theta_n; x_n)$$

**步骤 5** 将 L1 负荷预测结果逐级聚合,得到 L2 与 L3 全部节点预测结果:



$$\text{L2: } \sum_{i=1}^{m_1} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i), \sum_{i=m_1+1}^{m_1+m_2} f_i(\boldsymbol{\theta}_i; \mathbf{x}_i), \dots,$$
$$\sum_{i=n-m_h+1}^n f_i(\boldsymbol{\theta}_i; \mathbf{x}_i)$$
$$\text{L3: } \sum_{i=1}^n f_i(\boldsymbol{\theta}_i; \mathbf{x}_i)$$

其中第 1 个训练步骤,各 LSTM 模型在本地独立预训练  $L$  次,是为了使各模型达到预测准确度较高的初始状态,在此基础上进行联邦训练可以减少迭代次数.不仅如此,在准确的初始状态下,基于 ADMM 算法优化调整各模型能进一步提高整个层级结构的负荷预测准确度.

3 算例分析

3.1 实验数据与平台

本文所用数据来源于 2017 年全球能源预测竞赛<sup>[32]</sup>,由独立系统运营商新英格兰 (ISONE) 发布,该数据集包含 3 层负荷,底层包含 8 个区域.选择马萨诸塞州 (MASS) 的东南部 (Southeastern MASS, SEMA)、中西部 (Western and Central MASS, WCMA)、东北部 (Northeastern MASS, NEMA) 和佛蒙特州 (VT) 这 4 个底层区域的负荷数据等比缩小,构造配电网 3 级负荷层级结构.其中 L2 的 MASS 负荷数据是 SEMA、WCMA、NEMA 3 个区域负荷数据的总和,L2 的 VT 负荷数据与 L1 对应的 VT 区域负荷数据相同,构造的 3 级负荷层级结构如图 4 所示.使用数据包括从 2013 年 1 月 1 日至 2016 年 12 月 31 日的负荷数据、干球温度、露点温度和对应的月份与星期,数据分辨率都为 1 h.在构造数据集时,对于底层区域,选取待预测日前 1,2,3,4,5,6,7,14 日的负荷数据以及当天干球温度、露点温度、当天对应的月份、星期作为输入特征 (共 242 个数据),待预测日当天的实际负荷数据作为模型输出特征.对于 L2 和 L3 的区域,将待预测日当天的实际负荷数据作为标签.数据预处理时,输入特

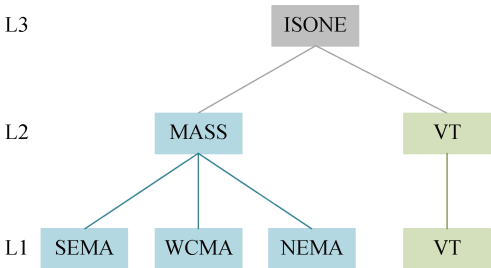


图 4 ISONE 负荷数据层级结构

Fig. 4 Hierarchical structure of ISONE load data

征、输出特征和标签数据按照区域的不同,分别进行归一化.以上构造的全部输入、输出特征和标签数据共同构成联邦预测模型的数据样本,按照 8 : 2 的比例划分训练集和测试集.本文的模型构建及训练均在 Python 编程环境下的 PyTorch 深度学习框架下进行,使用硬件平台为 Intel Core i5 CPU 和 NVIDIA GTX 1660 GPU.

3.2 模型超参数设置

为实现多客户端的联邦训练,共搭建 4 个 LSTM 模型,每个模型结构和超参数的选择都相同.经过多次试验,最终确定 LSTM 模型由 1 个输入层、4 个隐藏层和 1 个输出层构成,其中隐藏层神经元个数都为 100.输出层是一个线性的全连接层,将输出数据特征限制为 24,即待预测日当天 24 个时刻的负荷值.

此外,采用 Adam 优化算法更新模型参数,并将学习率设置为  $10^{-3}$ .多次试验表明,模型预训练迭代次数为 200 时,底层负荷的预测误差已小于 3%,达到了较准确的初始状态,因此确定预训练迭代次数为 200,联邦训练迭代次数设为 500,批处理数为 128,拉格朗日乘子  $\lambda$  初始值设为 0.1,惩罚因子  $\rho$  设为 0.1.

3.3 实验评价指标

选取平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE)、均方根误差 (Root Mean Square Error, RMSE) 和平均精度<sup>[33]</sup> (Mean Accuracy, MA) 作为预测结果的评价指标:

$$\text{MAPE} = \frac{1}{N} \sum_{t=1}^N \frac{|y(t) - \hat{y}(t)|}{\hat{y}(t)}$$

(10)

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t))^2}$$

(11)

$$\text{MA} = 1 - \text{MAPE}$$

(12)

式中:  $y(t)$  和  $\hat{y}(t)$  分别表示  $t$  时刻的负荷预测值和实际值;  $N$  为预测样本的采样点总数.

3.4 实验结果分析

在对比实验中,将本文方法与各层独立预测以及自下而上的传统方法进行比较,对测试集的数据进行 24 h 的短期负荷预测,得到 MAPE、RMSE、MA 这 3 种指标的实验结果分别如表 1~3 所示.各层独立预测方法分别对每个区域负荷采用单个 LSTM 模型进行预测,忽略了上下级负荷聚合一致性的约束,该方法作为实验的基础对照方案.自下而上的传统方法用 LSTM 模型得到底层区域的负荷预测,再按层级结构累加汇总成上层的负荷预测,以下简称传统方法.同时,为了体现本文方法选取

LSTM 作为基础模型的优越性,将 LSTM 替换为人工神经网络(ANN)和线性回归方法,对比 3 种模型在本文方法下的预测效果.表 1~3 中,将各列指标最优的数据加粗,以突出预测性能最佳的方法.另外,表中分别用“√”、“×”表示是否考虑聚合一致性.从表中可以看出,结合 LSTM 的本文方法获得的负荷预测结果不仅满足了上下级负荷聚合一致性的约束条件,而且显著提高了预测精度.

表 1 不同方法的负荷预测 MAPE 比较  
Tab. 1 Comparison of load forecasting MAPE by different methods

方法	MAPE/%				聚合 一致性
	L1	L2	L3	平均值	
独立预测	2.54	2.39	2.17	2.36	×
传统方法	2.54	2.58	3.14	2.75	√
本文方法(LSTM)	<b>1.61</b>	<b>1.36</b>	<b>1.03</b>	<b>1.33</b>	√
本文方法(ANN)	2.07	1.76	1.71	1.84	√
本文方法(线性回归)	3.27	2.92	2.78	2.99	√

表 2 不同方法的负荷预测 RMSE 比较  
Tab. 2 Comparison of load forecasting RMSE by different methods

方法	RMSE/MW				聚合 一致性
	L1	L2	L3	平均值	
独立预测	70.06	127.54	234.77	144.12	×
传统方法	70.06	154.41	337.45	187.30	√
本文方法(LSTM)	<b>41.35</b>	<b>61.91</b>	<b>112.02</b>	<b>71.76</b>	√
本文方法(ANN)	56.31	97.11	184.57	112.66	√
本文方法(线性回归)	103.14	196.01	361.39	220.18	√

表 3 不同方法的负荷预测 MA 比较  
Tab. 3 Comparison of load forecasting MA by different methods

方法	MA/%				聚合 一致性
	L1	L2	L3	平均值	
独立预测	97.46	97.60	97.83	97.55	×
传统方法	97.46	97.42	96.86	97.36	√
本文方法(LSTM)	<b>98.40</b>	<b>98.63</b>	<b>98.97</b>	<b>98.54</b>	√
本文方法(ANN)	97.93	98.24	98.29	98.06	√
本文方法(线性回归)	96.73	97.08	97.22	96.89	√

具体来说,首先对比均采用 LSTM 的独立预测方法、传统方法以及本文方法.与传统方法相比,本文方法对每层负荷预测的误差(MAPE, RMSE)都

更小、平均精度(MA)都更高.随着层级的增加,MAPE 逐层递减,L3 的 MAPE 低至 1.03%,是采用传统方法获得的同一层级 MAPE 的 1/3.对比均方根误差,本文方法获得的 RMSE 在绝对数值上显著小于传统方法,前者的 RMSE 平均值是后者平均值的 0.38,可见本文方法预测效果显著突出.对比两种方法在整个层级结构的平均精度,本文方法的 MA 高达 98.54%,同样优于传统方法.将本文方法与独立预测方法进行对比,不难发现,本文方法获得的预测结果不仅满足上下级负荷的聚合一致性,在预测精度上更是优于独立预测方法.这主要得益于 ADMM 算法在模型联邦训练过程中融入了难以显式建模的区域负荷间的相关性以及负荷的聚合一致性,在模型交替训练过程中不断修正原始负荷预测数据,因而得到精准的负荷预测结果.综上可以说明,本文的 ADMM 算法在预测效果上全面优于各层独立预测方法和传统方法,可作为电网统一各个层级的负荷预测结果的有效方法之一.此外,对比 LSTM、ANN 和线性回归方法,不难发现,在本文方法的框架下,采用 LSTM 基础模型获得了最高的预测精度.这说明本文方法的预测效果依赖于所选取的基础模型,而 LSTM 作为深度学习算法的代表性模型,相比机器学习方法和数学统计方法体现出了明显的优势.

为验证结合 LSTM 的本文方法所训练的模型已达到稳定,绘制训练过程中平均绝对百分比误差随迭代次数的变化曲线,如图 5 所示.图中分别展示了 3 个层级的 MAPE,其中 L1、L2、L3 分别选取 SEMA、MASS 和 ISONE 数据作为代表.可以看出,当迭代次数达到 500 次时,各层的 MAPE 已趋于最小值并在最小值附近轻微振荡,说明模型已达到收敛.

另外,绘制 5 种方法的负荷预测曲线,如图 6 所示,预测数据选取测试集中连续的 5 天,数据分辨率

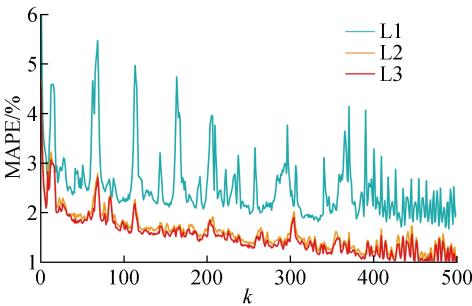


图 5 MAPE 变化曲线  
Fig. 5 Variation curves of MAPE

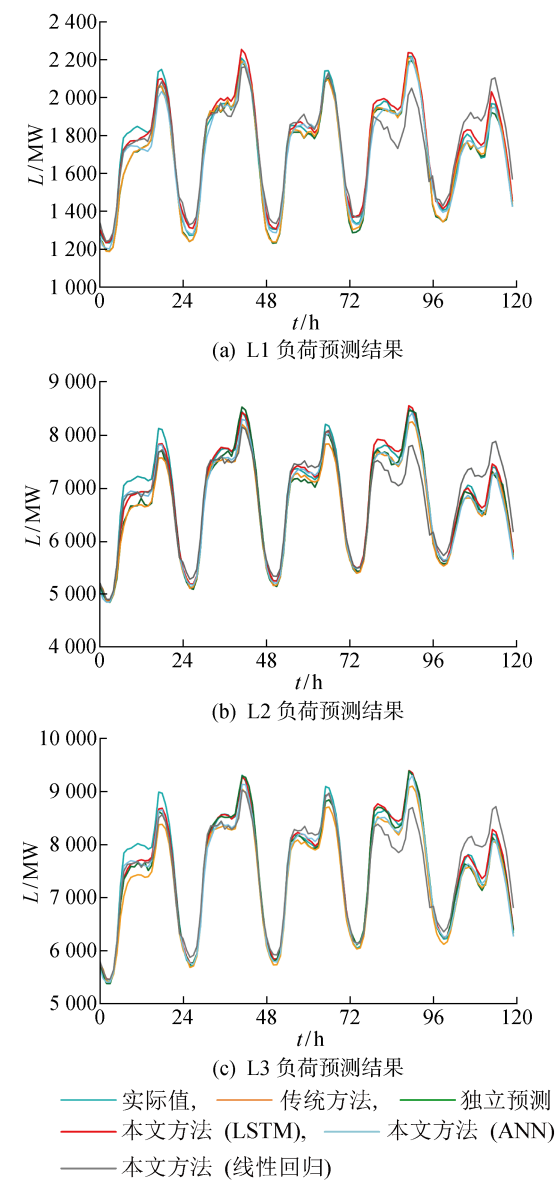


图 6 三级负荷层级结构的预测结果

Fig. 6 Forecast results of three-level load hierarchical structure

为 1 h,图中  $L$  为负荷.可以看出,结合 LSTM 的本文方法得到的负荷预测曲线误差最小,对于每一层级的数据都最好地拟合了实际负荷的变化曲线.而结合线性回归的本文方法预测误差最明显,这说明数学统计方法在预测效果上已逐渐被机器学习和深度学习方法所超越.当仅关注独立预测方法、传统方法和本文方法(LSTM)时,对于 L2 和 L3 的数据,传统方法的误差最大,这主要是由于随着层级结构的增加,下层负荷预测误差会在上层负荷叠加.对于各层数据,本文方法(LSTM)的负荷预测曲线都比各层独立预测得到的曲线更贴近实际值,这进一步佐证了本文方法预测准确度比独立预测方法还要高,大大提升了各层级负荷预测准确度.

4 结语

本文提出了一种融合 ADMM 与 LSTM 算法优势的多层负荷协同预测方法,其在满足负荷预测数据自下而上聚合一致性的同时,显著提升了所有层级负荷的整体预测准确度.采用的深度学习算法依靠纯数据驱动,无需任何显式的建模,就能学习负荷的变化规律.同时,在负荷聚合一致性的约束下使用 ADMM 算法,使得模型的训练过程融入了负荷数据间的相关性以及负荷聚合一致性,进而达到提升负荷预测准确度的目的.算例结果表明,该方法具有突出的预测效果,可以帮助电网更好地协调各个层级的负荷预测结果.然而,本文未考虑节假日等假期因素对负荷预测的影响,在后续研究中,可在本文基础上进一步增加节假日的输入特征,也可以考虑电价、湿度、降水量等其它特征,通过关键影响因素关联分析确定最佳的输入特征集,进一步提高负荷预测准确度.

参考文献:

[1] 谢小荣,贺静波,毛航银,等.“双高”电力系统稳定性的新问题及分类探讨[J].中国电机工程学报,2021,41(2):461-475.  
XIE Xiaorong, HE Jingbo, MAO Hangyin, *et al.* New issues and classification of power system stability with high shares of renewables and power electronics[J]. *Proceedings of the CSEE*, 2021, 41(2): 461-475.

[2] 韦善阳,黎静华,黄乾,等.考虑多重因素耦合的广义负荷特征曲线的模式分析[J].电力系统自动化,2021,45(1):114-122.  
WEI Shanyang, LI Jinghua, HUANG Qian, *et al.* Pattern analysis of generalized load characteristic curve considering coupling of multiple factors[J]. *Automation of Electric Power Systems*, 2021, 45(1): 114-122.

[3] 陈海文,王守相,王绍敏,等.基于门控循环单元网络与模型融合的负荷聚合体预测方法[J].电力系统自动化,2019,43(1):65-72.  
CHEN Haiwen, WANG Shouxiang, WANG Shaomin, *et al.* Aggregated load forecasting method based on gated recurrent unit networks and model fusion[J]. *Automation of Electric Power Systems*, 2019, 43(1): 65-72.

[4] 孔祥玉,李闯,郑锋,等.基于经验模态分解与特征相关分析的短期负荷预测方法[J].电力系统自动化,2019,43(5):46-52.



- KONG Xiangyu, LI Chuang, ZHENG Feng, *et al.* Short-term load forecasting method based on empirical mode decomposition and feature correlation analysis[J]. **Automation of Electric Power Systems**, 2019, 43(5): 46-52.
- [5] 赵兵, 王增平, 纪维佳, 等. 基于注意力机制的 CNN-GRU 短期电力负荷预测方法[J]. **电网技术**, 2019, 43(12): 4370-4376.
- ZHAO Bing, WANG Zengping, JI Weijia, *et al.* A short-term power load forecasting method based on attention mechanism of CNN-GRU[J]. **Power System Technology**, 2019, 43(12): 4370-4376.
- [6] GOIA A, MAY C, FUSAI G. Functional clustering and linear regression for peak load forecasting[J]. **International Journal of Forecasting**, 2010, 26(4): 700-711.
- [7] TRUDNOWSKI D J, MCREYNOLDS W L, JOHNSON J M. Real-time very short-term load prediction for power-system automatic generation control[J]. **IEEE Transactions on Control Systems Technology**, 2001, 9(2): 254-260.
- [8] 张帆, 张峰, 张士文. 基于提升小波的时间序列分析法的电力负荷预测[J]. **电气自动化**, 2017, 39(3): 72-76.
- ZHANG Fan, ZHANG Feng, ZHANG Shiwen. Power load forecasting in the time series analysis method based on lifting wavelet[J]. **Electrical Automation**, 2017, 39(3): 72-76.
- [9] 肖白, 聂鹏, 穆钢, 等. 基于多级聚类分析和支持向量机的空间负荷预测方法[J]. **电力系统自动化**, 2015, 39(12): 56-61.
- XIAO Bai, NIE Peng, MU Gang, *et al.* A spatial load forecasting method based on multilevel clustering analysis and support vector machine[J]. **Automation of Electric Power Systems**, 2015, 39(12): 56-61.
- [10] SINGH P, DWIVEDI P. Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem[J]. **Applied Energy**, 2018, 217: 537-549.
- [11] 孔祥玉, 郑锋, 鄂志君, 等. 基于深度信念网络的短期负荷预测方法[J]. **电力系统自动化**, 2018, 42(5): 133-139.
- KONG Xiangyu, ZHENG Feng, E Zhijun, *et al.* Short-term load forecasting based on deep belief network[J]. **Automation of Electric Power Systems**, 2018, 42(5): 133-139.
- [12] 房龙江. 多尺度空间分辨率下的多级负荷预测方法研究[D]. 吉林: 东北电力大学, 2018.
- FANG Longjiang. Study on the prediction method of multi-stage load in multi-scale spatial resolution[D]. Jilin: Northeast Dianli University, 2018.
- [13] 顾洁, 孟璐, 郑睿程, 等. 考虑集群辨识的海量用户负荷分层概率预测[J]. **电力系统自动化**, 2021, 45(5): 71-78.
- GU Jie, MENG Lu, ZHENG Ruicheng, *et al.* Load-stratified probability forecasting for massive users considering cluster identification[J]. **Automation of Electric Power Systems**, 2021, 45(5): 71-78.
- [14] 肖峻, 张璇, 张婷, 等. 应用信息熵原理的多路径负荷预测协同方法[J]. **电力系统及其自动化学报**, 2013, 25(2): 42-47.
- XIAO Jun, ZHANG Xuan, ZHANG Ting, *et al.* Method of multi-path load forecasting collaboration using maximum entropy principle[J]. **Proceedings of the Chinese Society of Universities for Electric Power System and Its Automation**, 2013, 25(2): 42-47.
- [15] LLOYD J R. GEFCom2012 hierarchical load forecasting: Gradient boosting machines and Gaussian processes[J]. **International Journal of Forecasting**, 2014, 30(2): 369-374.
- [16] BEN TAIEB S, HYNDMAN R J. A gradient boosting approach to the Kaggle load forecasting competition[J]. **International Journal of Forecasting**, 2014, 30(2): 382-394.
- [17] PANG Y, YAO B, ZHOU X D, *et al.* Hierarchical electricity time series forecasting for integrating consumption patterns analysis and aggregation consistency[C] // **Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence**. California: International Joint Conferences on Artificial Intelligence Organization, 2018: 3506-3512.
- [18] SPILOTIS E, ABOLGHASEMI M, HYNDMAN R J, *et al.* Hierarchical forecast reconciliation with machine learning[J]. **Applied Soft Computing**, 2021, 112: 107756.
- [19] 亢良伊, 王建飞, 刘杰, 等. 可扩展机器学习的并行与分布式优化算法综述[J]. **软件学报**, 2018, 29(1): 109-130.
- KANG Liangyi, WANG Jianfei, LIU Jie, *et al.* Survey on parallel and distributed optimization algorithms for scalable machine learning[J]. **Journal of Software**, 2018, 29(1): 109-130.
- [20] BOYD S. Distributed optimization and statistical learning via the alternating direction method of multipliers[M]. San Francisco, CA, USA: Now Publishers Inc, 2010.
- [21] CHAMIKARA M A P, BERTOK P, KHALIL I, *et al.* Privacy preserving distributed machine learning

- with federated learning[J]. **Computer Communications**, 2021, 171: 112-125.
- [22] MCMAHAN H B, MOORE E, RAMAGE D, *et al.* Communication-efficient learning of deep networks from decentralized data[EB/OL]. [2021-06-20]. <https://arxiv.org/abs/1602.05629>.
- [23] 康重庆, 牟涛, 夏清. 电力系统多级负荷预测及其协调问题: (一)研究框架[J]. **电力系统自动化**, 2008, 32(7): 34-38.  
KANG Chongqing, MU Tao, XIA Qing. Power system multilevel load forecasting and coordinating: Part one. Research framework[J]. **Automation of Electric Power Systems**, 2008, 32(7): 34-38.
- [24] 沈泉江, 郭乃网, 郑作梁. 基于用电模式聚类的层级电力时序预测方法[J]. **计算机应用与软件**, 2020, 37(11): 73-78.  
SHEN Quanjiang, GUO Naiwang, ZHENG Zuoliang. Hierarchical electricity time series forecasting method based on clustering of electricity consumption patterns[J]. **Computer Applications and Software**, 2020, 37(11): 73-78.
- [25] 张世旭, 苗世洪, 杨炜晨, 等. 基于自适应步长ADMM的配电网分布式鲁棒优化调度策略[J]. **高电压技术**, 2021, 47(1): 81-93.  
ZHANG Shixu, MIAO Shihong, YANG Weichen, *et al.* Distributed robust optimal dispatch for active distribution networks based on alternative direction method of multipliers with dynamic step size[J]. **High Voltage Engineering**, 2021, 47(1): 81-93.
- [26] HU L, YAN H Y, LI L, *et al.* MHAT: An efficient model-heterogenous aggregation training scheme for federated learning [J]. **Information Sciences**, 2021, 560: 493-503.
- [27] HEGEDŰS I, DANNER G, JELASITY M. Decentralized learning works: An empirical comparison of gossip learning and federated learning[J]. **Journal of Parallel and Distributed Computing**, 2021, 148: 109-124.
- [28] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. **Neural Computation**, 1997, 9(8): 1735-1780.
- [29] 陈纬楠, 胡志坚, 岳菁鹏, 等. 基于长短期记忆网络和LightGBM组合模型的短期负荷预测[J]. **电力系统自动化**, 2021, 45(4): 91-97.  
CHEN Weinan, HU Zhijian, YUE Jingpeng, *et al.* Short-term load prediction based on combined model of long short-term memory network and light gradient boosting machine[J]. **Automation of Electric Power Systems**, 2021, 45(4): 91-97.
- [30] 陆继翔, 张琪培, 杨志宏, 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): 131-137.  
LU Jixiang, ZHANG Qipei, YANG Zhihong, *et al.* Short-term load forecasting method based on CNN-LSTM hybrid neural network model[J]. **Automation of Electric Power Systems**, 2019, 43(8): 131-137.
- [31] 姚子麟, 张亮, 邹斌, 等. 含高比例风电的电力市场电价预测[J]. **电力系统自动化**, 2020, 44(12): 49-55.  
YAO Zilin, ZHANG Liang, ZOU Bin, *et al.* Electricity price prediction for electricity market with high proportion of wind power[J]. **Automation of Electric Power Systems**, 2020, 44(12): 49-55.
- [32] HONG T, XIE J R, BLACK J. Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting[J]. **International Journal of Forecasting**, 2019, 35(4): 1389-1399.
- [33] 孙庆凯, 王小君, 张义志, 等. 基于LSTM和多任务学习的综合能源系统多元负荷预测[J]. **电力系统自动化**, 2021, 45(5): 63-70.  
SUN Qingkai, WANG Xiaojun, ZHANG Yizhi, *et al.* Multiple load prediction of integrated energy system based on long short-term memory and multi-task learning[J]. **Automation of Electric Power Systems**, 2021, 45(5): 63-70.

(本文编辑:陈晓燕)