第45卷 第22期 2021年11月25日
DOI：10.7500/AEPS20210423011
电力系统自动化
Automation of Electric Power Systems
Vol. 45 No. 22 Nov. 25，2021

基于虚拟相似日与DA-LSTPNet的地区电网短期负荷预测

李 滨，高 枫
（广西电力系统最优化与节能技术重点实验室（广西大学），广西壮族自治区南宁市 530004）

摘要：针对短期负荷预测精细化的需求，提出一种基于虚拟相似日与双阶段注意力机制的长短期时序神经网络(DA-LSTPNet)的地区级短期负荷预测方法。为获得与负荷相匹配的细粒度实时气象数据，首先基于粗粒度的气象数据，利用灰色关联度和关联度加权法获取含细粒度气象数据的气象虚拟相似日。然后，采用最大信息系数(MIC)对气象特征信息与负荷进行非线性关联性分析，构建MIC加权下的负荷虚拟相似日选取算法，解决选取历史日作为传统负荷相似日而导致的过分局部相似乃至非相似的问题。最后，针对过往预测模型缺乏考虑特征因素与局部负荷细粒度变化之间联系特性的问题，构建能够有效挖掘负荷特征数据长期宏观以及短期局部变化特性的DA-LSTPNet进行日前短期负荷预测。以中国南方某地区电网实际负荷数据为例，采用多种形式的仿真验证了所提预测方法具有更高的预测精度和普适性。

关键词：短期负荷预测；虚拟相似日；双分支神经网络；时间注意力机制

0 引言

随着经济的发展和大规模分布式能源的并网，地区电网气象敏感负荷基数不断上升，峰谷差持续拉大，负荷的局部变化情况变得越发随机与复杂[1]。如何科学有效地挖掘细粒度气象特征与局部负荷变化之间的联系，是提高负荷预测精度、实现负荷预测工作精细化管理的必然方向。

传统的地区电网短期负荷预测研究一般分为两个步骤：一是选取合适的相似日；二是基于相似日进行建模预测[2]。过往的研究仅针对负荷进行相似日选取，实际上相似日的选取亦可扩展至以气象等为对象的特征相似日。在实际的负荷预测中，采用粗粒度气象预测数据进行建模预测难以反映细粒度变化特性，对预测效果影响很大。因此，如何细粒度化原有粗粒度的气象预测数据是亟待解决的问题之一。

大部分负荷相似日选取方法依旧使用粗粒度日特征气象（如日最高温度等）[2-5]来构建每日特征向量，难以映射到局部变化。文献[6]引入实时气象因素构建日特征向量进行相似日选取，提升了选取的相似度，但局部相似问题还未解决。文献[7]为克服选取历史日作为相似日出现的局部相似问题，将日负荷序列分成多段并分段建立相似日选取模型，不仅复杂度较高且存在割裂日负荷序列变化连续性的可能。

对于建模部分，时间序列法[8-9]等传统统计方法对于负荷的非线性特征无法充分挖掘；支持向量机[10-11]、人工神经网络[12-13]等机器学习方法往往只从特征数据的角度出发，对时序依赖关系缺乏应有的考虑。深度学习中归属循环神经网络（RNN）类的长短期记忆（LSTM）网络[14-15]和门控循环单元（GRU）[16-17]网络可考虑负荷特征数据中蕴含的非线性时序联系，前景较好。但这类网络未能对时序特征进行有效抉择，当时序信息过多时易出现特征信息的弱化与遗忘，特别是关键信息[18]。日趋波动复杂的负荷预测更需关注其局部变化情况，单一结构的模型已难以有效应对，须构建组合模型以提升其特征挖掘和预测的能力。

针对上述问题，本文提出了一种基于虚拟相似日与双阶段注意力机制的长短期时序神经网络（dual-stage attention-based long and short time pattern network，DA-LSTPNet）的短期负荷预测方法。采用灰色关联度与关联度加权法从粗粒度气象预测数据中选取日负荷曲线相对应的气象虚拟相似日以细粒度化气象预测数据；利用加权灰色关联度与关联度加权法获得虚拟负荷相似日曲线；采用卷积神经网络（CNN）、GRU以及时间注意力机制

收稿日期：2021-04-23；修回日期：2021-08-01。
上网日期：2021-09-26。
国家自然科学基金资助项目（51767004）；广西创新驱动发展专项资助项目（AA19254034）。

(time attention mechanism, TAM)构建双分支结构下的日前短期负荷预测模型DA-LSTPNet,其中,CNN用来捕捉局部短期时间特征,GRU用来提取长期宏观时间信息,经过双阶段TAM的优化,充分学习特征因素与负荷序列变化间的时序波动联系。最后,仿真通过实际电网数据验证了本文方法的有效性。

1　虚拟相似日选取

本文参照机器学习理论将数据集划分为训练集、验证集和测试集。虚拟相似日选取分为两类:①测试集待预测日的气象虚拟相似日;②数据集中所有待预测日的负荷虚拟相似日。其中,训练集和验证集的待预测日利用气象实际数据直接进行虚拟负荷相似日选取;测试集待预测日先进行气象虚拟相似日选取后,再进行负荷虚拟相似日选取。

1.1　气象虚拟相似日选取

天气预报是一种概率预报,短期气象预测的数据包括气温、湿度、风向、风速、气压等,以最高值、最低值、平均值或小时级数值呈现,颗粒度较粗。本文引入气象虚拟相似日概念,采用灰色关联度与关联度加权法进行选取,以获得分钟级的细颗粒度气象预报值。

气象虚拟相似日是指以粗粒度气象数据为基础,采用相似度评价函数对待预测日选择特征相似度最高的一日作为气象相似日,并采取相应的手段对气象相似日曲线进行修正,以获得相似度更高的细粒度相似日虚拟气象曲线。具体做法如附录A图A1所示,步骤如下。

1)将日特征气象组建为气象相似日选取所需的日特征序列向量,该向量所含的日特征气象因素具体如附录A表A1和表A2所示。

2)采用灰色关联度对待预测日进行气象历史相似日选择,选取范围为每个待预测日对应往前一个月的日期范围以内,最终选取关联度最高的前3个历史日作为气象相似日,并确定三者对应的关联度大小。

3)以关联度数值所占的百分比作为该关联度所对应的权重,根据权重为3类气象相似日的实时气象进行加权组合,形成最终的虚拟气象相似日气象序列。

1.2　负荷虚拟相似日选取

为优化相似日负荷的相似特性、扩大局部相似特性的范围,乃至实现全局相似,本文提出基于最大信息系数(maximal information coefficient, MIC)赋权的加权灰色关联度算法选取历史相似日,再依据关联度加权法自行组合多类历史相似日形成相似度更高的虚拟相似日负荷序列。

1.2.1　负荷相似日选取所输入的因素集合

负荷相似日方法的思想为气象等特征因素相似的两日,负荷变化情况也相似[19]。考虑日特征气象在局部负荷映射方面的局限性,本文引入更为具体的每日96点细粒度实时气象(含综合气象)等特征来构建相似日选取所需要的日实时特征量序列,具体气象特征如附录A表A3所示。

1.2.2　MIC赋权的加权灰色关联度虚拟相似日选取

与气象虚拟相似日的构建方法类似,本文采用MIC赋权的加权灰色关联度方法选取获得负荷虚拟相似日,相关步骤如下。

1)步骤1:MIC赋权

由于传统灰色关联度方法在选取负荷相似日时对各影响特征采用平权处理,一定程度上掩盖了不同影响特征对负荷的影响,因此,在负荷相似日选取前需要对影响程度不同的诸多气象特征赋予相匹配的权重[20]。实时特征和负荷之间均为非线性变化的时序序列,为充分表征负荷与特征之间的关联程度,本文引入MIC[21]对气象特征与负荷进行关联特性分析。附录A表A4展示了负荷和各类实时气象的MIC关联特性情况,计算方式见附录B。

2)步骤2:基于加权灰色关联度的历史负荷相似日选取

本文设置只从与待预测日相同的日类型中选取历史负荷相似日。取各气象因素的MIC值所占百分比作为对应的权重,计算公式如附录B式(B3)所示。采用加权灰色关联度对每个待预测日进行负荷相似日选择,选取范围为每个待预测日对应的往前20个同类型日,最终选取出关联度最高的前3个历史日作为每个待预测日所需的历史负荷相似日,并记录下这3个负荷相似日的关联度大小。计算方式见附录B式(B4)和式(B5)。

3)步骤3:基于关联度权重组合的虚拟负荷相似日

设需要进行负荷相似日选取的待预测日有m个,对应选取得到的3类历史相似日负荷序列为:

$$\boldsymbol{L}_{\mathrm{ss}}=\begin{bmatrix} L_{\mathrm{fir},1}(d) & L_{\mathrm{sec},1}(d) & L_{\mathrm{thi},1}(d) \\ L_{\mathrm{fir},2}(d) & L_{\mathrm{sec},2}(d) & L_{\mathrm{thi},2}(d) \\ \vdots & \vdots & \vdots \\ L_{\mathrm{fir},m}(d) & L_{\mathrm{sec},m}(d) & L_{\mathrm{thi},m}(d) \end{bmatrix} \tag{1}$$

式中：$L_{\mathrm{fir},i}$、$L_{\mathrm{sec},i}$、$L_{\mathrm{thi},i}$分别为数据集第i个待预测日的第1最佳、第2最佳、第3最佳历史相似日负荷序列，其中$i=1,2,\cdots,m$；$d=0,1,\cdots,95$表示每日96个时间点。这3类负荷相似日分别对应的灰色关联度分别为$G_{\mathrm{fir},i}$、$G_{\mathrm{sec},i}$、$G_{\mathrm{thi},i}$。以灰色关联度数值所占的百分比表示该关联度对应的权重$w_{j,i}$。

$$w_{j,i}=\frac{G_{j,i}}{\sum_{j=\mathrm{fir,sec,thi}}G_{j,i}} \tag{2}$$

最终每个待预测日对应的虚拟相似日负荷$L_{\mathrm{vir},i}$计算公式为：

$$L_{\mathrm{vir},i}=\sum_{j=\mathrm{fir,sec,thi}}L_{j,i}w_{j,i} \tag{3}$$

4）步骤4：相似日负荷修正

第1最佳历史相似日与其余第2、第3最佳相似日进行权重组合形成的虚拟相似日负荷有可能出现相似度下降的问题，为此，本文对虚拟负荷相似日的日实时特征量序列进行提取，修正虚拟相似日的负荷。

设$\boldsymbol{f}_{\mathrm{ss},i}$为第$i$个虚拟相似日负荷的各类历史相似日负荷对应的日实时特征分量。

$$\boldsymbol{f}_{\mathrm{ss},i}=[f_{\mathrm{fir},i}(k)\quad f_{\mathrm{sec},i}(k)\quad f_{\mathrm{thi},i}(k)] \tag{4}$$

式中：$f_{\mathrm{fir},i}$、$f_{\mathrm{sec},i}$、$f_{\mathrm{thi},i}$分别为第i个虚拟相似日第1最佳、第2最佳以及第3最佳历史相似日下对应的日实时特征分量；$k=1,2,\cdots,672$为特征分量中各元素的位置。

调用步骤3中获得的组合权重$w_{j,i}$，与前述日实时特征分量$f_{j,i}$相乘，遂获得第i个虚拟负荷相似日对应的实时特征气象$f_{\mathrm{smi},i}$。

$$f_{\mathrm{smi},i}=\sum_{j=\mathrm{fir,sec,thi}}f_{j,i}w_{j,i} \tag{5}$$

运用灰色关联度算法分别计算负荷虚拟相似日和最佳负荷相似日实时气象与待预测日实时气象之间的关联度，判断两者关联度大小，若虚拟相似日实时气象的关联度大，则保留该日的相似日负荷为虚拟相似日负荷；反之，则替换为最佳历史相似日负荷。

2　DA-LSTPNet预测模型

本文引入了CNN、GRU网络以及TAM来构建日前短期负荷预测模型DA-LSTPNet，以充分挖掘负荷与特征数据之间的长期宏观、短期局部特征信息和准确描述时序依赖关系，其机理如下。

2.1　CNN

CNN的局部连接和权值共享特性可以使模型以较低的复杂度提取输入矩阵不同粒度的局部特征。CNN的卷积层内含的多个卷积核会沿输入信息的时间方向进行窗口滑动，最终输出一个表征局部短期时间依赖的特征矩阵，其中卷积核的大小决定了卷积层能够提取到的或窄或宽的时间段特征[22]。为了减少输出特征矩阵的复杂度并降低模型训练出现过拟合的风险，同时保证输出特征未出现遗漏和弱化，在卷积层后面还会设置最大池化层。

卷积层包含多个卷积核，本文将卷积核高度设置为4，旨在关注小时级的局部时间变化特性，宽度的大小与输入特征矩阵的特征元素数目相同，第i_{th}个卷积核在输入矩阵$\boldsymbol{X}$上进行计算的方式如式(6)所示[23]。

$$\boldsymbol{h}_{\mathrm{c}}=\mathrm{ReLU}(\boldsymbol{W}_{\mathrm{c}}*\boldsymbol{X}+\boldsymbol{b}_{\mathrm{c}}) \tag{6}$$

式中：$*$为卷积运算符；$\boldsymbol{h}_{\mathrm{c}}$为第$i_{\mathrm{th}}$个卷积核在输入特征矩阵进行卷积运算后得到的特征向量；$\boldsymbol{W}_{\mathrm{c}}$为该卷积核所要学习的权值参数；$\boldsymbol{b}_{\mathrm{c}}$为偏置向量；$\mathrm{ReLU}(\cdot)$为线性整流函数。纵向拼接不同的$\boldsymbol{h}_{\mathrm{c}}$，最终获得表征短期局部信息的特征矩阵$\boldsymbol{H}_{\mathrm{c}}$。

2.2　GRU

GRU通过其特有的更新门和重置门结构实现对“历史负荷长期特征”的记忆与遗忘功能。相较于RNN，其可以有效缓解“梯度爆炸”问题；相较于LSTM网络，其结构简化、训练时间短。GRU结构如附录C图C1所示，具体计算过程如下[17]。

$$\boldsymbol{r}_t=\sigma(\boldsymbol{W}_{\mathrm{r}}[\boldsymbol{h}_{t-1}\quad \boldsymbol{x}_t]) \tag{7}$$

$$\boldsymbol{z}_t=\sigma(\boldsymbol{W}_{\mathrm{s}}[\boldsymbol{h}_{t-1}\quad \boldsymbol{x}_t]) \tag{8}$$

$$\tilde{\boldsymbol{h}}_t=\mathrm{relu}(\boldsymbol{W}_{\mathrm{h}}[\boldsymbol{r}_t\boldsymbol{h}_{t-1}\quad \boldsymbol{x}_t]) \tag{9}$$

$$\boldsymbol{h}_t=(1-\boldsymbol{z}_t)\boldsymbol{h}_{t-1}+\boldsymbol{z}_t\tilde{\boldsymbol{h}}_t \tag{10}$$

式中：$\boldsymbol{x}_t$为当前t时刻下输入模型的负荷特征信息；$\boldsymbol{W}_{\mathrm{r}}$、$\boldsymbol{W}_{\mathrm{s}}$、$\boldsymbol{W}_{\mathrm{h}}$为可训练的参数矩阵；$\boldsymbol{r}_t$和$\boldsymbol{z}_t$分别为$t$时刻的更新门和重置门状态信息；$\boldsymbol{h}_t$为$t$时刻GRU网络隐藏状态输出；$\tilde{\boldsymbol{h}}_t$为$\boldsymbol{x}_t$和$\boldsymbol{h}_{t-1}$二者联合计算得到的状态信息；$\sigma(\cdot)$为sigmoid激活函数。连接不同时刻下的向量$\boldsymbol{h}_t$，最终获得表征长期时间模式的特征矩阵$\boldsymbol{H}_t$。

2.3　TAM

TAM是一种模拟人脑注意力资源分配的机制，旨在从大量信息中筛选并聚焦到少量重要时间特征[18]。本文引入TAM，旨在加强模型对漫长时序信息的记忆，强化局部短期信息关键特征，突出影响负荷变化的关键时间因素，提高模型的预测效果。所设计的TAM框架如附录C图C2所示，计算

方式为[24]：

$$\boldsymbol{\alpha}_t=\mathrm{softmax}(\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{W}_a\boldsymbol{H}_t) \tag{11}$$

$$\boldsymbol{c}_t=\boldsymbol{\alpha}_t\boldsymbol{H}_t \tag{12}$$

$$\boldsymbol{h}_t^{\mathrm{A}}=\tanh(\boldsymbol{W}[\boldsymbol{c}_t;\boldsymbol{h}_t]+\boldsymbol{b}) \tag{13}$$

式中：$\boldsymbol{H}_t=[\boldsymbol{h}_{t-w+1},\boldsymbol{h}_{t-w+2},\cdots,\boldsymbol{h}_{t-1}]$为含多个时刻信息的长期时间模式特征矩阵，来自与TAM相连的GRU网络，其时间窗口长度为w；$\boldsymbol{W}$和$\boldsymbol{W}_a$为可训练的权重参数矩阵；$\boldsymbol{\alpha}_t$为t时刻的注意力权重向量，表征状态矩阵中每个时刻下的隐藏状态信息重要程度，长度也为w；$\boldsymbol{c}_t$为t时刻赋权后表征时序关系的特征向量；$[\boldsymbol{c}_t;\boldsymbol{h}_t]$表示向量$\boldsymbol{c}_t$和$\boldsymbol{h}_t$的拼接；$\boldsymbol{b}$为偏置向量；$\boldsymbol{h}_t^{\mathrm{A}}$为特征向量与当前$t$时刻状态信息的拼接，也是TAM最终的输出结果。

与CNN相连的TAM也有着相似的计算流程，只是CNN的输出为一个表征多个时刻信息的局部特征矩阵，与之相连的TAM则是计算待预测时刻下局部耦合特征与过往时刻下局部耦合特征的关联性，为不同时刻下的局部特征赋予相匹配的权重。

TAM通常会采用矩阵张量点乘dot函数来实现状态特征信息的赋权，dot函数会指定两个待乘矩阵中需要进行点乘的轴，点乘运算完成后指定的轴在输出中会被消除，由此TAM可实现对GRU以及CNN输出的降维。

2.4　DA-LSTPNet预测模型结构

本文所提的短期负荷预测模型DA-LSTPNet由两个分支网络组成：一个是单阶段TAM下的CNN(single attention CNN，SA-CNN)；另一个则是单阶段TAM下的双层GRU(single attention GRU，SA-GRU)神经网络，如附录A图A2所示。

采用滑动窗口的方式来构建模型的输入与输出，每输入一个时间窗口的特征矩阵对应输出一个时间点下的负荷标幺值。输入信息在经过输入层后，信息同时流向SA-CNN和SA-GRU两个分支。SA-CNN负责局部变量之间的短时依赖信息的提取，并基于TAM对不同的局部信息赋予关键的预测权重；SA-GRU则实现对长期宏观信息的提取。之后采用add层对两个分支提取的特征信息进行融合，经过双层全连接层对相融合的结果进行降维。最终输出待预测时刻下的负荷标幺值，进行反归一化得到待预测时刻下的负荷有名值。如此循环预测96次即获得待预测日的日负荷序列曲线。预测模型相关参数如附录C表C1所示。

3　基于虚拟相似日和DA-LSTPNet的短期负荷预测算法

3.1　数据预处理

为了消除不同数据类型带来的量纲差异以及避免模型训练过程中导致的梯度问题，本文采用min-max方式将包含负荷、气象在内的各类特征数据归一化至(0,1)之间。

$$x^*=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{14}$$

式中：x为样本中待归一化的数据；x^*为归一化后的数据；$x_{\max}$和$x_{\min}$分别为样本数据中的最大值和最小值。

3.2　相似日选取评价指标

本文引入形态相似距离D_{MSD}作为相似日选取优劣的评价指标[25]。形态相似距离可以同时衡量曲线数值间距与形状相似程度。形状越相似，数值越接近，相似距离D_{MSD}的数值就越小。

$$D_{\mathrm{MSD}}(\boldsymbol{L}_{\mathrm{c}},\boldsymbol{L}_{\mathrm{f}})=\sqrt{\sum_{k=1}^{n}(l_{\mathrm{c},k}-l_{\mathrm{f},k})^2}\left(2-\frac{\left|\sum_{k=1}^{n}l_{\mathrm{c},k}-l_{\mathrm{f},k}\right|}{\sum_{k=1}^{n}\left|l_{\mathrm{c},k}-l_{\mathrm{f},k}\right|}\right) \tag{15}$$

式中：$\boldsymbol{L}_{\mathrm{c}}$和$\boldsymbol{L}_{\mathrm{f}}$分别为待比较相似日负荷序列与待预测日负荷序列，其中$\boldsymbol{L}_{\mathrm{c}}=[l_{\mathrm{c},1},l_{\mathrm{c},2},\cdots,l_{\mathrm{c},k},\cdots,l_{\mathrm{c},n}]^{\mathrm{T}}$，$\boldsymbol{L}_{\mathrm{f}}=[l_{\mathrm{f},1},l_{\mathrm{f},2},\cdots,l_{\mathrm{f},k},\cdots,l_{\mathrm{f},n}]^{\mathrm{T}}$，$n=96$为日负荷序列长度。

3.3　模型预测误差指标

为了评估模型的预测精度，本文选取均方根误差(RMSE)、平均百分比误差(MAPE)和准确度这3个指标进行评估。

$$R_{\mathrm{mse}}=\sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_{\mathrm{true},k}-y_{\mathrm{pred},k})^2} \tag{16}$$

$$M_{\mathrm{ape}}=\frac{1}{n}\sum_{k=1}^{n}\frac{\left|y_{\mathrm{true},k}-y_{\mathrm{pred},k}\right|}{y_{\mathrm{true},k}} \tag{17}$$

$$A_k=\left[1-\sqrt{\frac{1}{n_{\mathrm{a}}}\sum_{k=1}^{n}\left(\frac{y_{\mathrm{true},k}-y_{\mathrm{pred},k}}{y_{\mathrm{true},k}}\right)^2}\right]\times 100\% \tag{18}$$

式中：A_k为第k个采样点的准确度；n_{a}为预测的元素总数，$n_{\mathrm{a}}=96$；$y_{\mathrm{true},k}$和$y_{\mathrm{pred},k}$分别为第k个采样点的实际负荷值和预测负荷值。R_{mse}和M_{ape}的值越小代表负荷预测效果越好，准确度则反之。

3.4　基于DA-LSTPNet的日前短期负荷预测

基于虚拟相似日和DA-LSTPNet的日前短期

负荷预测流程如附录C图C3所示，具体步骤如下。

1）选取相似日。对非测试集的待预测日直接进行负荷虚拟相似日选取，对测试集待预测日先进行气象虚拟相似日选取。然后，根据气象虚拟相似日序列进行负荷虚拟相似日选取。

2）在选取获得虚拟相似日负荷后，与历史负荷、气象等特征数据一起组建成特征数据集，其中包含的具体特征元素如附录C表C2所示。

3）采用滑动窗口的方式在数据集中构建待预测日每个待预测时间点负荷对应的时间窗口矩阵。为了保证每个滑动时间窗口下输入预测模型的特征数据尽可能充足，使模型能够自适应地从大数据中学习到长期宏观以及短期局部的变化特征，本文将时间窗口矩阵的时间步长设置为96。附录C表C2所涉及的特征输入元素数目21即为矩阵的特征维度。

时间窗口矩阵的每一行为每个时刻下输入的特征向量；时间窗口矩阵的每一列为每个特征元素从过去时刻$t-95$至待预测时刻t的时间序列。

4）将训练集的输入和输出投入DA-LSTPNet模型进行训练，每次迭代训练后会在验证集样本中检验模型的训练效果，设置迭代次数50次，并以验证集损失函数最小为目标设置Model Check Point最优模型保存函数，保存训练过程中效果最好的模型参数。训练完成后最终输入测试集数据完成负荷预测。

4　负荷预测仿真及分析

4.1　数据介绍和实验平台

本文采用中国南方电网有限责任公司某地级市2015-01-04至2017-12-29实际电力负荷数据以及相关的气象等特征数据作为仿真数据集。负荷、气象数据采样频率均为15 min/点。选取2015年至2016年的数据作为模型的训练集，2017上半年数据作为验证集，下半年数据作为测试集。为了能够充分验证本文所提短期负荷预测方法的有效性和普适性，采用3种方式对所提方法进行仿真：变换不同相似日负荷下的对比仿真、针对模型结构的消融实验、变换不同负荷预测模型的对比仿真。

本文采用的硬件平台为Intel Core i7-8700 CPU和NVDIA GRID V100D-32Q GPU。软件代码均基于Python语言进行开发，预测模型均依靠tensorflow2.1支持下的keras深度学习框架进行搭建。Lightgbm则调用Lightgbm机器学习包实现。

4.2　虚拟相似日选取

分别采用历史相似日法与虚拟相似日法对待预测日2017-08-28进行相似日选取。采用加权灰色关联度选取得到关联度从大到小的3个相似日，分别为2017-08-10、2017-08-18、2017-08-25，各类相似日以及待预测日负荷序列曲线对比如图1所示，数值结果对比如表1所示。此外，图1还展示了气象虚拟相似日与待预测日2017-08-28温度序列之间的差异性，表1最后附上该日气象虚拟相似日的温度选取结果。

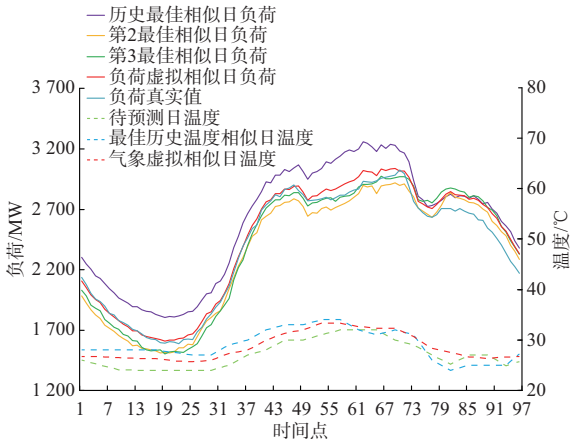


图1　2017-08-28选取相似日的负荷和温度曲线对比
Fig. 1　Comparison of load and temperature curves for selected similar days on August 28, 2017

表1　2017-08-28选取相似日结果比较
Table 1　Comparison of results for selected similar days on August 28, 2017

相似日类别	日期	灰色关联度	形态相似距离
历史最佳相似日	2017-08-10	0.881 6	0.628 9
第2最佳相似日	2017-08-18	0.880 5	0.416 9
第3最佳相似日	2017-08-25	0.862 2	0.568 7
负荷虚拟相似日			0.246 5
最佳历史温度相似日	2017-08-09	0.880 8	1.678 7
温度虚拟相似日			0.566 9

结合表1与图1可知：

1）由于负荷中内含的随机波动分量以及特征因素考虑不全面，特征相似度最高的一日并不意味着日负荷序列相似度一定为最高，需要修正和完善；

2）本文所提的虚拟相似日根据灰色关联度数值求出相关权重，自适应权重组合历史相似日负荷序列，有效地汇集了不同历史相似日负荷序列的相似特性，进而达到提高相似日负荷序列相似程度的目的。

采用与负荷虚拟相似日相类似的做法，构建出

气象虚拟相似日实时气象，实现了相似日实时气象序列相似度的提高。附录D表D1展示了气象虚拟相似日实时气象序列与实际气象序列在温度、湿度和风速上形态相似距离均值、最值的比较结果。

为验证气象虚拟相似日的实时气象特征给负荷预测带来的提升效果，本文做了含气象相似日与不含气象相似日下的预测仿真对比。其中，不含气象相似日的气象特征分量以日气象特征均值替代。两个不同的特征数据集分别投入DA-LSTPNet模型中进行训练预测，预测效果如附录D表D2所示。

针对负荷相似日，选取2015年至2017年的日负荷曲线，采用形态相似距离对虚拟相似日负荷以及历史相似日负荷进行比较。其中3类历史相似日以及虚拟相似日与待预测日间形态相似距离平均值、最值如附录D表D3所示。

同理，为验证负荷虚拟相似日选取给负荷预测带来的提升效果，本文分别利用历史最佳相似日负荷与虚拟相似日负荷来组建预测特征集，放入DA-LSTPNet模型进行预测，以该地区2017年7月至12月作为测试集来对比二者相似日的效果，结果如附录D表D4所示。

4.3　基于DA-LSTPNet模型的消融实验

为了验证本文所提预测模型DA-LSTPNet各部分的有效性，对该模型进行消融实验[26-27]仿真。消融实验有3个，分别摘除的结构为CNN、GRU网络以及双阶段TAM，因此，本文参与消融预测对比模型分别为SA-GRU、SA-CNN以及双层GRU网络，这些模型依旧保留原模型双层全连接层与激活函数，模型结构如附录A图A2所示，其余参数见附录C表C1。

以夏季典型日2017-08-10为例进行仿真分析，此日整体气象状况为：最低气温为24.5 ℃，最高气温为33 ℃，全天温度差异较大；最低湿度为63%，最高湿度达到100%，伴随有阵雨，局部气象有明显变化。且2017-08-09、2017-08-08均为高温湿雨天气，对2017-08-10有一定的气象累积影响效应。因此，对该日的预测不仅要求模型考虑局部短时气象与人们生存生活机制对负荷带来的局部变化影响，还要考虑多日长时气象累积对负荷的影响情况。

图2展示了消融实验的预测情况，可以看出：①单纯的GRU网络预测效果并不良好（准确度为95.48%），但加入TAM后，SA-GRU预测效果有了显著提升（准确度为96.56%），可见TAM的存在通过为GRU网络提取出来的时序特征信息进行加权，帮助预测模型把握长时序列变化特征，在一定程度上提高了预测的准确性；②SA-CNN预测出现了较大的数据水平偏离（准确度为95.19%）；③DA-LSTPNet模型通过整合CNN、GRU以及TAM，使其无论是在长期负荷水平的趋势预测上，还是在局部峰谷变化特性的把握上，效果均为最优（准确度97.81%）。

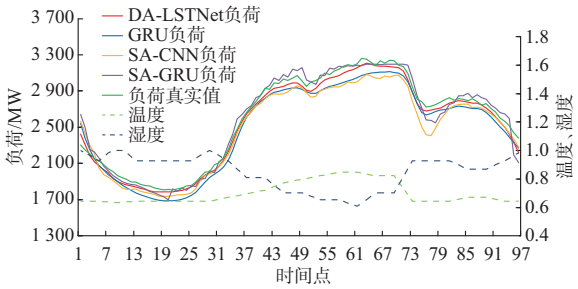


图2　基于DA-LSTPNet的消融实验
Fig. 2　Ablation test based on DA-LSTPNet

为充分验证DA-LSTPNet预测的普适性，对该地区进行了连续半年的日前负荷预测，并选出其中夏冬两季典型月7月和12月进行预测效果展示，如表2和表3所示。其中，各类指标结果为逐日计算后所取的平均值。

表2　模型消融实验下的夏冬两季典型月仿真结果对比
Table 2　Comparison of simulation results of typical months in summer and winter with model ablation test

模型	夏季7月			冬季12月		
	RMSE/MW	MAPE/%	准确度/%	RMSE/MW	MAPE/%	准确度/%
DA-LSTPNet	96.30	3.27	96.01	64.08	2.57	96.97
SA-GRU	104.50	3.64	95.60	79.39	3.20	96.29
SA-CNN	117.94	4.17	95.03	91.19	3.38	95.52
双层GRU	138.89	4.59	94.65	83.39	3.29	95.93

表3　模型消融实验下7至12月的仿真结果对比
Table 3　Comparison of simulation results from July to December with model ablation test

模型	RMSE/MW	MAPE/%	准确度/%
DA-LSTPNet	92.70	3.44	95.97
SA-GRU	106.88	3.95	95.44
SA-CNN	111.19	4.18	95.01
双层GRU	112.86	4.10	95.09

综合上述表2和表3的消融预测结果，显然可以看出：加入TAM之后，GRU网络的预测性能有了较为明显的提升，证明了TAM可以有效地对时

序特征进行加权，帮助提取时序特征中关键的信息，从而达到提高预测精度的目的。本文所提模型DA-LSTPNet通过结合CNN、GRU以及TAM，加强了对负荷与其余特征之间的长期宏观、短期局部时序变化联系特性的挖掘，预测精度得到进一步显著提升。无论在夏冬两季典型月还是连续半年的日前负荷预测中，其表现出来的预测性能均为最好，模型鲁棒性也为最优。

4.4　多模型预测结果仿真对比

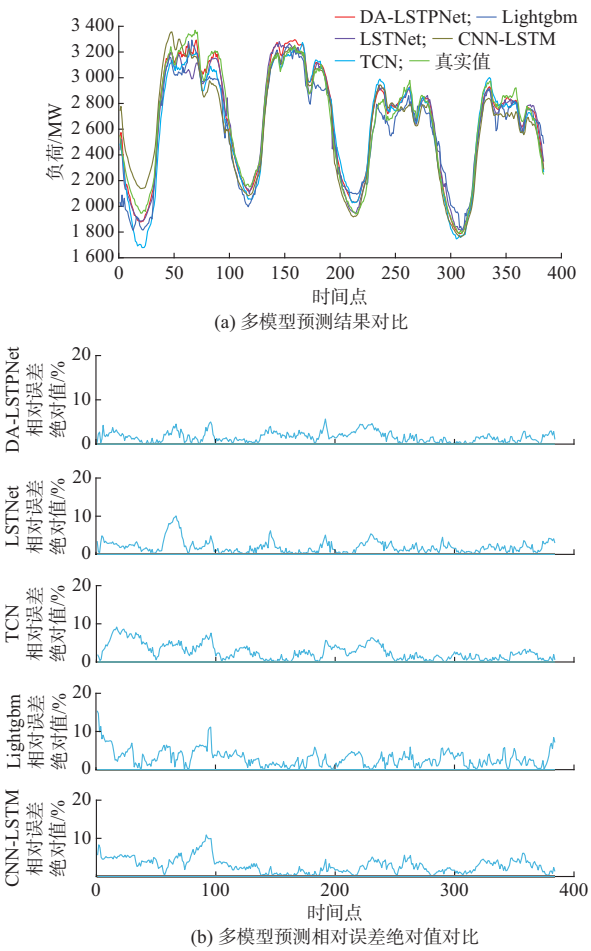
为了验证模型预测的有效性，本文采用多类目前时序预测流行的模型与本文所提模型作对比，包括时间卷积网络（TCN）、长短期时间序列网络（LSTNet）[26]、Lightgbm以及CNN-LSTM组合网络。其中，TCN模型由双层TCN和一层全连接网络组成，每层TCN均采用默认参数设置。LSTNet模型参数参照文献[26]所述理论进行设置。CNN-LSTM模型则由单层卷积层、单层最大池化层以及双层神经元均为64的LSTM网络层组合而成，其余参数设置与DA-LSTPNet一致。上述深度学习模型均设置迭代次数为50，激活函数均为带参数的线性整流函数（PReLU）。Lightgbm模型目标选择回归，训练方式采用梯度提升决策树(GBDT)，迭代次数为5 000，评估标准设置为RMSE，其余可调参数采用贝叶斯算法寻优确定。

选择2017-08-01至2017-08-04连续4天作为夏季典型高温降雨日进行分析，其气象状况如附录D表D5所示。连续4天夏季典型日的预测效果如图3所示，附录D表D6则以数值方式展示了夏季典型日预测效果。

从图3可以看出，在连续多天的夏季典型日预测中，无论是在峰谷还是处于负荷趋势形态的时间点上，本文所提模型DA-LSTPNet的相对误差绝对值数值均较低，拟合效果均优于其他3个模型。

同样地，为保证模型预测的普适性，本文也对这4个模型分别做了连续半年的日前负荷预测，并节选7月与12月作为夏冬两季典型月进行展示，展示的各类预测精度指标均为逐日计算后所取的平均值。预测结果如表4和表5所示，在表5中还加入了模型的训练和预测时间，用以对比不同模型的预测效率。

从表4和表5可以看出，在季节变换的情况下，DA-LSTPNet、TCN、LSTNet、CNN-LSTM模型预测精度均较高，均表现出了较好的预测稳定性。



(a) 多模型预测结果对比

(b) 多模型预测相对误差绝对值对比

图3　连续多天夏季典型日的预测对比

Fig. 3　Forecast comparison of several consecutive typical days in summer

表4　不同模型在夏冬两季典型月仿真结果对比

Table 4　Comparison of simulation results of typical months in summer and winter with different models

模型	夏季7月			冬季12月		
	RMSE/MW	MAPE/%	准确度/%	RMSE/MW	MAPE/%	准确度/%
DA-LSTPNet	96.30	3.27	96.01	64.08	2.57	96.97
TCN	106.34	3.60	95.70	67.63	2.72	96.80
LSTNet	105.19	3.59	95.54	64.71	2.69	96.86
Lightgbm	137.89	4.80	94.26	73.21	2.91	96.59
CNN-LSTM	108.96	3.87	95.36	85.86	3.36	96.07

Lightgbm则在气象多变、高温累积效应严重的夏季表现出了较差的预测效果，在气象变化相对稳定的冬季效果较好，预测普适性较弱。CNN-LSTM的预测效果在季节变换的情况下虽较为稳定，但精度并不突出。TCN和LSTNet模型的预测效果较为接近，但是LSTNet模型的训练时间均较长，所耗费预测资源相对较多。在预测时间相当的LSTNet和

表5　不同模型在7至12月的仿真结果以及训练预测时间对比

Table 5　Comparison of simulation results and training prediction time from July to December with different models

模型	RMSE/MW	MAPE/%	准确度/%	训练时间/s	预测时间/s
DA-LSTPNet	92.70	3.44	95.97	724	11
TCN	103.06	3.94	95.33	787	7
LSTNet	100.74	3.76	95.58	871	11
Lightgbm	118.00	3.92	94.77	8	1
CNN-LSTM	108.03	3.95	95.27	625	8

DA-LSTPNet模型当中，预测效果最好的是本文所提的DA-LSTPNet模型，其预测结果比LSTNet预测结果在RMSE和MAPE上分别降低了7.98%和8.51%，预测准确度则提升了0.41%，而且训练时间也比LSTNet降低了16.88%。

5　结语

为提高地区电网短期负荷预测的准确度，本文提出了基于虚拟相似日和DA-LSTPNet的短期负荷预测方法，该方法特点如下。

1)从粗粒度的日特征气象数据出发，构建了气象虚拟相似日选取算法，获得与日负荷曲线相匹配的细粒度96点实时气象数据，为负荷精细化预测提供了匹配有效的支撑。

2)采用MIC量化负荷与气象特征之间的非线性关联程度，以细粒度气象数据作为负荷相似日选取的输入，采用MIC赋权的加权灰色关联度算法与关联度加权法选取获得负荷虚拟相似日曲线。在保留日负荷序列变化特性的基础上，将各历史相似日负荷序列各不相同的局部相似特性进行整合，从而进一步提高了相似日选取的相似度。

3)提出了双分支网络结构下的DA-LSTPNet模型，实现了对负荷与特征间宏观长期和局部短期的特征信息提取；引入TAM对GRU网络和CNN提取出来的时间信息进行加权，加强了网络对时序关系的捕捉，获得了更好的预测效果。

本文构建的特征当中还缺乏实时电价因素的考虑，后续可以进一步考虑更加丰富特征数据选择，发挥大数据优势，提高负荷预测的精度。

附录见本刊网络版(http://www.aeps-info.com/aeps/ch/index.aspx)，扫英文摘要后二维码可以阅读网络全文。

参考文献

[1] 于道林.智能电网下需求响应机理及其短期负荷预测模型研究[D].青岛:青岛大学,2017.
YU Daolin. Research on demand response mechanism and short-term load forecasting model in smart grid[D]. Qingdao: Qingdao University, 2017.

[2] 黎灿兵,李晓辉,赵瑞,等.电力短期负荷预测相似日选取算法[J].电力系统自动化,2008,32(9):69-73.
LI Canbing, LI Xiaohui, ZHAO Rui, et al. A novel algorithm of selecting similar days for short-term power load forecasting[J]. Automation of Electric Power Systems, 2008, 32(9): 69-73.

[3] 吴潇雨,和敬涵,张沛,等.基于灰色投影改进随机森林算法的电力系统短期负荷预测[J].电力系统自动化,2015,39(12):50-55.
WU Xiaoyu, HE Jinghan, ZHANG Pei, et al. Short-term load prediction of power system based on grey projection improved random forest algorithm [J]. Automation of Electric Power Systems, 2015, 39(12): 50-55.

[4] 李滨,黄佳,吴茵,等.基于分形特性修正气象相似日的节假日短期负荷预测方法[J].电网技术,2017,41(6):1949-1955.
LI Bin, HUANG Jia, WU Yin, et al. Holiday short-term load forecasting based on fractal characteristic modified meteorological similar day[J]. Power System Technology, 2017, 41(6): 1949-1955.

[5] 吴云,雷建文,鲍丽山,等.基于改进灰色关联分析与蝙蝠优化神经网络的短期负荷预测[J].电力系统自动化,2018,42(20):67-72.
WU Yun, LEI Jianwen, BAO Lishan, et al. Short-term load forecasting based on improved grey relational analysis and optimized bat neural network[J]. Automation of Electric Power Systems, 2018, 42(20): 67-72.

[6] 李滨,陆明珍.考虑实时气象耦合作用的地区电网短期负荷预测建模[J].电力系统自动化,2020,44(17):60-68.
LI Bin, LU Mingzhen. Modeling of regional power grid short-term load forecasting considering real-time meteorological coupling[J]. Automation of Electric Power Systems, 2020, 44(17): 60-68.

[7] 王剑锋,向铁元,徐富祥,等.基于分段多目标相似日选取法的短期负荷预测[J].武汉大学学报,2016,49(3):435-440.
WANG Jianfeng, XIANG Tieyuan, XU Fuxiang, et al. Short-term load forecasting based on piecewise similar selecting method with multiobjective[J]. Journal of Wuhan University, 2016, 49(3): 435-440.

[8] TAYLOR J W. An evaluation of methods for very short-term load forecasting using minute-by-minute British data [J]. International Journal of Forecasting, 2008, 24(4): 645-658.

[9] LI W, ZHANG Z. Based on time sequence of ARIMA model in the application of short-term electricity load forecasting [C]// International Conference on Research Challenges in Computer Science, December 28-29, 2009, Shanghai, China: 11-14.

[10] 吴倩红,高军,侯广松,等.实现影响因素多源异构融合的短期

负荷预测支持向量机算法[J].电力系统自动化，2016，40(15)：67-72.
WU Qianhong, GAO Jun, HOU Guangsong, et al. Short-term load forecasting support vector machine algorithm based on multi-source heterogeneous fusion of load factors [J]. Automation of Electric Power Systems, 2016, 40(15): 67-72.

[11] 熊军华，牛珂，张春歌，等.基于小波变异果蝇优化支持向量机短期负荷预测方法研究[J].电力系统保护与控制，2017，45(13)：71-77.
XIONG Junhua, NIU Ke, ZHANG Chunge, et al. Research on short-term load forecasting method based on wavelet mutation drosophila optimization support vector machine [J]. Power System Protection and Control, 2017, 45(13): 71-77.

[12] VEERAMSETTY V, DESHMUKH R. Electric power load forecasting on a 33/11 kV substation using artificial neural networks[J]. SN Applied Sciences, 2020, 2(5): 1-10.

[13] HERNÁNDEZ L, BALADRÓN C, AGUIAR J M, et al. Artificial neural network for short-term load forecasting in distribution systems[J]. Energies, 2014, 7: 1576-1598.

[14] 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. Multivariate load prediction of integrated energy system based on LSTM and multi-task learning [J]. Automation of Electric Power Systems, 2021, 45(5): 63-70.

[15] 陈锦鹏，胡志坚，陈纬楠，等.二次模态分解组合DBiLSTM-MLR的综合能源系统负荷预测[J].电力系统自动化，2021，45(13)：85-94.
CHEN Jinpeng, HU Zhijian, CHEN Weinan, et al. Load prediction of integrated energy system based on combination of quadratic modal decomposition and deep bidirectional long short-term memory and multiple linear regression[J]. Automation of Electric Power Systems, 2021, 45(13): 85-94.

[16] 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)：65-74.
CHEN Haiwen, WANG Shouxiang, WANG Shaomin, et al. Load aggregation prediction method based on gated cyclic unit network and model fusion [J]. Automation of Electric Power Systems, 2019, 43(1): 65-74.

[17] 王增平，赵兵，纪维佳，等.基于GRU-NN模型的短期负荷预测方法[J].电力系统自动化，2019，43(5)：53-58.
WANG Zengping, ZHAO Bing, JI Weijia, et al. Short-term load forecasting method based on GRU-NN model [J]. Automation of Electric Power Systems, 2019, 43(5): 53-58.

[18] 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. Short-term power load forecasting method based on CNN-GRU attention mechanism [J]. Power System Technology, 2019, 43(12): 4370-4376.

[19] 孙谦，姚建刚，赵俊，等.基于最优交集相似日选取的短期母线负荷综合预测[J].中国电机工程学报，2013，33(4)：126-134.
SUN Qian, YAO Jiangang, ZHAO Jun, et al. Comprehensive forecasting of short-term busbar load based on optimal selection of similary day of intersection [J]. Proceedings of the CSEE, 2013, 33(4): 126-134.

[20] 李啸骢，李春涛，从兰美，等.基于动态权值相似日选取算法的短期负荷预测[J].电力系统保护与控制，2017，45(6)：1-8.
LI Xiaocong, LI Chuntao, CONG Lanmei, et al. Short-term load prediction based on dynamic weight similary day selection algorithm [J]. Power System Protection and Control, 2017, 45(6): 1-8.

[21] WANG C, DAI D, LI X, et al. Accelerating computation of large biological datasets using MapReduce framework [J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2016, 14(4): 783-795.

[22] 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.

[23] LAI G, CHANG W, YANG Y, et al. Modeling long- and short-term temporal patterns with deep neural networks [C]// 41st International ACM SIGIR Conference, July 8-12, 2018, Ann Arbor, USA.

[24] LUONG M T, PHAM H, MANNING C D. Effective approaches to attention-based neural machine translation [C]// Conference on Empirical Methods in Natural Language Processing, September 17-21, 2015, Lisbon, Portugal: 1412-1421.

[25] 李中，刘洋洋，张铁峰.基于形态相似距离的时间序列相似度计算[J].计算机工程与设计，2016，37(3)：679-683.
LI Zhong, LIU Yangyang, ZHANG Tiefeng. Time series similarity measurement based on morphology similarity distance [J]. Computer Engineering and Design, 2016, 37(3): 679-683.

[26] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, USA: 580-587.

[27] CHENG J, HUANG K, ZHENG Z. Towards better forecasting by fusing near and distant future visions [C]// AAAI Conference on Artificial Intelligence, February 7-12, 2020, New York, USA.

李　滨(1975—)，女，博士，副教授，硕士研究生导师，主要研究方向：电力系统最优化。E-mail：lizhen@gxu.edu.cn

高　枫(1996—)，男，通信作者，硕士研究生，主要研究方向：电网短期负荷预测。E-mail：335783642@st.gxu.edu.cn

（编辑　王梦岩）

Short-term Load Forecasting for Regional Power Grid Based on Virtual Similar Days and Dual-stage Attention-based Long and Short Time Pattern Network

LI Bin, GAO Feng

(Key Laboratory of Guangxi Electric Power System Optimization and Energy-saving Technology (Guangxi University), Nanning 530004, China)

Abstract: A regional short-term load forecasting method based on virtual similar days and dual-stage attention-based long and short time pattern network (DA-LSTPNet) is proposed for the demand of the refinement of short-term load forecasting. To obtain the fine-grained real-time meteorological data matching the load, a meteorological virtual similar day containing fine-grained meteorological data is firstly obtained using gray correlation and correlation weighting method based on the coarse-grained meteorological data. Then, the maximum information coefficient (MIC) is used to analyze the nonlinear correlation between meteorological feature information and load. And the MIC-weighted selection algorithm for the load virtual similar day is constructed to solve the problem of excessive local similarity or even non-similarity caused by selecting historical days as traditional load similar days. Finally, in order to address the problem that the relationship between characteristic factors and local load fine-grained variation, the DA-LSTPNet is constructed to effectively explore the characteristics of long-term macroscopic and short-term local variation of load feature data for the day-ahead short-term load forecasting. Based on the actual load data of the power grid in a certain area of southern China, the various forms of simulation are used to demonstrate the higher prediction accuracy and universality of the proposed forecasting method.

This work is supported by National Natural Science Foundation of China (No. 51767004) and Guangxi Special Fund for Innovation-driven Development (No. AA19254034).

Key words: short-term load forecasting; virtual similar day; dual-branch neural network; time attention mechanism