

# 基于 Attention-BiLSTM 神经网络和气象数据修正的短期负荷预测模型

王继东, 杜冲

(天津大学 智能电网教育部重点实验室, 天津 300072)

**摘要:**基于循环神经网络的负荷预测模型大多将历史负荷数据和影响负荷的其他因素如气象数据等共同作为预测模型的输入特征,但气象数据内部规律性不强,不适合作为循环神经网络的输入。针对该问题,提出一种基于 Attention-BiLSTM 神经网络和气象数据修正的短期负荷预测模型。采用最大信息系数分析影响负荷的主要因素;考虑到负荷序列较长且存在双向信息流,利用 BiLSTM 神经网络进行预测;引入注意力机制,通过注意力权重突出关键因素的影响,挖掘负荷数据的内部规律;利用核极限学习机结合气象数据进行误差预测和修正,完成负荷预测。以我国东部某地区真实数据作为实际算例,实验结果表明,所提模型与其他模型相比有更好的预测效果。

**关键词:**短期负荷预测;最大信息系数;注意力机制;双向长短时记忆神经网络;核极限学习机

**中图分类号:**TM 715

**文献标志码:**A

**DOI:**10.16081/j.epae.202112017

## 0 引言

准确高效的短期电力负荷预测是保证电力系统安全可靠运行的前提,也是电网合理安排发电计划的依据。电力市场化改革的进行,对短期电力负荷预测的精度提出了更高的要求。

用于短期电力负荷预测的方法主要分为统计学方法和机器学习方法两大类。统计学方法主要有回归分析法<sup>[1]</sup>、自回归积分滑动平均模型<sup>[2]</sup>、波动模型<sup>[3]</sup>等,这些方法原理简单,计算速度较快,但是对于变量间非线性关系的处理能力有限。机器学习方法主要有支持向量机法<sup>[4-5]</sup>、反向传播(BP)神经网络<sup>[6]</sup>、随机森林法<sup>[7]</sup>等,这些方法的共同缺点是难以挖掘负荷序列的时序性,且不适合处理大量数据。

近年来,深度学习发展迅速,为短期负荷预测方法的研究提供了新的思路。典型的深度学习方法有卷积神经网络、深度信念网络和循环神经网络 RNN (Recurrent Neural Network)等。其中,以 RNN 为主体的预测模型由于在处理时序数据方面的优势而被广泛应用。长短时记忆 LSTM (Long Short-Term Memory) 神经网络是一种特殊的 RNN,其通过增加门控单元,解决了传统 RNN 存在的梯度消失问题<sup>[8]</sup>,能够更好地学习时间序列中的历史信息。文献<sup>[9]</sup>使用 LSTM 神经网络进行负荷预测,相较于 BP 神经网络, LSTM 神经网络提高了预测精度,但是由于输入特征较多,难以突出对当前负荷影响更大的因素。文献<sup>[10]</sup>在负荷预测前首先使用卷积神经网络进行特

征提取,但是其仅将历史负荷数据作为模型的输入,而没有考虑其他因素对负荷数据的影响。文献<sup>[11]</sup>在使用 LSTM 神经网络进行预测的同时,引入注意力(Attention)机制,通过给输入特征赋予不同的权重,突出关键特征的影响,增强了模型学习序列中远距离依赖的能力。

上述方法存在的共同问题是, LSTM 神经网络只能考虑单向的数据信息,忽视了序列的反向信息对预测的影响。另外,这些方法通常是历史负荷数据和影响负荷的其他因素如气象数据等作为模型的输入特征共同输入预测模型。文献<sup>[12]</sup>通过对真实气象数据序列特征的分析发现,气象数据在几个星期到几年的范围内可近似为白噪声,序列内部规律性差。而 RNN 及其变体处理内部规律性不强的数据的能力较差<sup>[13]</sup>。因此,气象数据等不适合作为 RNN 模型的输入特征。

针对以上问题,本文提出一种基于 Attention-BiLSTM (Attention Bidirectional Long Short-Term Memory) 和气象数据修正的短期负荷预测模型。在采用最大信息系数 MIC (Maximum Information Coefficient) 分析影响负荷的主要因素的基础上,针对规律性较强的历史负荷数据以及规律性较差的气象数据分别建立预测模型和误差修正模型。首先,使用双向长短时记忆 (BiLSTM) 神经网络学习历史负荷数据的正、反向规律,并引入 Attention 机制突出关键因素的影响;然后,使用核极限学习机 KELM (Kernel Extreme Learning Machine) 结合气象数据进行误差预测,并对预测模型的结果进行修正,旨在提高模型的整体预测精度。实验结果表明,本文模型相较于其他模型具有更高的预测精度。

收稿日期:2021-04-02;修回日期:2021-10-27

基金项目:国家重点研发计划项目(2016YFB0901102)

Project supported by the National Key Research and Development Program of China(2016YFB0901102)

### 1 影响负荷的因素分析

由于负荷序列具有一定的周期性和趋势性,因此历史负荷数据是预测当前时刻负荷的最主要依据。此外,气象数据特别是温度和相对湿度对负荷数据也有一定影响。因此,本文重点考虑历史负荷及温度、相对湿度对当前时刻负荷的影响。

由于 MIC 在处理相关性分析问题时具有较好的鲁棒性和较小的计算复杂度<sup>[14]</sup>,因此本文采用 MIC 分析各变量与当前负荷数据之间的相关性。

#### 1.1 MIC

MIC 是一种基于互信息的关联性分析算法。它采用网格划分的方法,不仅可以衡量数据集中变量之间的线性和非线性函数关系,还可以广泛地挖掘变量之间的非函数依赖关系,具有较好的普适性和公平性。MIC 计算方法描述如下。

在一个二元数据集  $D \in \mathbf{R}^2$  中,给定 2 个变量  $X = \{x_i, i=1, 2, \dots, n\}$  和  $Y = \{y_i, i=1, 2, \dots, n\}$ ,  $n$  为变量  $X$  和  $Y$  的维度,则其互信息  $I(X; Y)$  定义为:

$$I(X; Y) = \sum_{y_i \in Y} \sum_{x_i \in X} p(x_i, y_i) \log_2 \frac{p(x_i, y_i)}{p(x_i)p(y_i)} \quad (1)$$

式中:  $p(x_i, y_i)$  为  $X$  和  $Y$  的联合概率密度;  $p(x_i)$ 、 $p(y_i)$  分别为  $X$ 、 $Y$  的边缘概率密度。

将数据集  $D$  中由  $x_i$  和  $y_i$  构成的散点图划分为  $a$  行  $b$  列的网格,记为  $G = (a, b)$ 。分别计算各个网格中的互信息  $I_{D|G}(X; Y)$  ( $D|G$  表示数据集  $D$  在网格  $G$  下的划分),并选取最大值作为划分  $G$  的互信息值,则最大互信息值  $I_{D|G}^*(X; Y)$  可定义为:

$$I_{D|G}^*(X; Y) = \max_G I_{D|G}(X; Y) \quad (2)$$

将得到的互信息标准化,求出 MIC:

$$M_{D|G}(X; Y) = \frac{I_{D|G}^*(X; Y)}{\lg \min(a, b)} \quad (3)$$

$$F(D)_{\text{MIC}} = \max_{ab < B(n)} \{M_{D|G}(X; Y)\} \quad (4)$$

式中:  $M_{D|G}(X; Y)$  为标准化的最大互信息值;  $\lg \min(a, b)$  为标准化系数;  $F(D)_{\text{MIC}}$  为 MIC;  $B(n)$  为网格划分数的上界,一般取  $B(n) = n^{0.6}$ 。MIC 是一种归一化的最大互信息,其取值范围为  $[0, 1]$ , 2 个变量间的 MIC 值越接近 1, 表示变量间的关联性越强, MIC 值越接近 0, 表示变量间的关联性越弱。

#### 1.2 历史负荷与当前负荷的相关性分析

附录 A 图 A1 为我国东部某地区一个月的负荷数据变化情况。由图可知,电力负荷数据具有明显的周期性:随着人们生活作息的变化,白天负荷水平逐渐提高,夜晚负荷水平逐渐降低,这体现出负荷的日周期性;一周中每天的负荷遵从上述变化,其中周一至周五为工作日,负荷以工业用电为主,负荷水平

较高,而周六、周日由于工厂等用电大户休息,负荷以居民用电为主,处于较低水平,这体现出负荷的周周期性。此外,负荷数据还具有连续性,中间一般不会出现阶跃和离散现象。

设  $p_i(t)$  为第  $i$  天  $t$  时刻的负荷数据,  $P_i(t)$  为第  $i$  天以负荷项  $p_i(t)$  为尾项,向第  $i-1$  天方向扩展的长度为 24 个时刻的负荷序列。考虑到负荷的日周期性、周周期性以及连续性,分别计算一个月内  $P_i(t-1)$ 、 $P_i(t-2)$ 、 $P_{i-1}(t)$ 、 $P_{i-1}(t-1)$ 、 $P_{i-7}(t)$ 、 $P_{i-7}(t-1)$  和当前负荷序列  $P_i(t)$  之间的相关性,计算结果如表 1 所示。

表 1 历史负荷与当前负荷的相关性

Table 1 Correlation between historical load and current load

负荷序列	MIC	负荷序列	MIC
$P_i(t-1)$	0.9075	$P_{i-1}(t-1)$	0.7600
$P_i(t-2)$	0.7587	$P_{i-7}(t)$	0.8501
$P_{i-1}(t)$	0.8152	$P_{i-7}(t-1)$	0.7815

由表 1 可知,所选负荷序列与当前负荷序列的 MIC 值均大于 0.7, 相关性较强<sup>[15]</sup>。其中,  $P_i(t-1)$  与  $P_i(t)$  时间相距最近,相关性最强。另外,所选负荷数据的周周期性略强于日周期性,计算结果符合负荷变化的实际情况。

#### 1.3 气象数据与当前负荷的相关性分析

除了历史负荷,气象也被认为是影响当前负荷的重要因素,本文主要考虑温度和相对湿度对负荷的影响。

设  $P_i$  为第  $i$  天的负荷数据,  $T_i$  为第  $i$  天的温度数据,  $H_i$  为第  $i$  天的相对湿度数据。由于气象数据对负荷数据的影响可能存在季节性差异,因此计算各季节典型月中每日气象数据和负荷数据的相关性。本文认为 MIC 值大于 0.2 则表示两序列之间存在一定的相关性。分别计算各季节典型月中气象数据与负荷数据之间的 MIC 值大于 0.2 的天数占比,计算结果如表 2 所示。

表 2 气象数据与负荷的相关性

Table 2 Correlation between meteorological data and load

典型月	MIC 值大于 0.2 的天数占比 / %	
	温度	相对湿度
4 月(春)	36.7	16.7
7 月(夏)	51.6	22.6
10 月(秋)	32.3	16.1
12 月(冬)	45.2	12.9

表 2 结果表明,温度和相对湿度是负荷重要的外部影响因素,且存在季节性差异。其中温度与负荷的相关性更强,对负荷的影响更大。综合来看,气象数据在夏季时对负荷的影响最为明显。

## 2 基于 Attention-BiLSTM 神经网络的预测模型

### 2.1 BiLSTM 神经网络

LSTM 神经网络是一种时间 RNN,是目前最成功的 RNN 架构之一。LSTM 神经网络基本结构见附录 A 图 A2。

LSTM 神经网络的核心是门控单元,其中遗忘门负责选择性地遗忘历史信息,输入门负责保留当前信息并将其与历史信息融合,输出门决定当前状态对隐含层的影响。LSTM 神经网络对应的参数计算如下:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (5)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (6)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (7)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (8)$$

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

式中:  $f_t$ 、 $i_t$ 、 $o_t$ 、 $c_t$  分别为当前  $t$  时刻遗忘门、输入门、输出门和状态单元的状态;  $h_{t-1}$  为前一时刻隐含层的状态;  $x_t$  为当前  $t$  时刻的输入;  $W_f$ 、 $W_i$ 、 $W_o$ 、 $W_c$  和  $b_f$ 、 $b_i$ 、 $b_o$ 、 $b_c$  分别为对应的权重系数矩阵和偏置项;  $\sigma(\cdot)$  表示 Sigmoid 激活函数。

LSTM 神经网络存在的一个不足是,其仅能够利用从正向序列中传来的历史信息,这使其在对波动过程进行预测时可能会导致相位滞后。

为验证负荷数据存在双向信息流,选取数据集中连续一个月的负荷数据,以其中中心点为原点向正、反 2 个方向延伸,得到正向负荷序列和反向负荷序列。分别计算正、反向负荷序列的自相关系数,结果如图 1 所示。

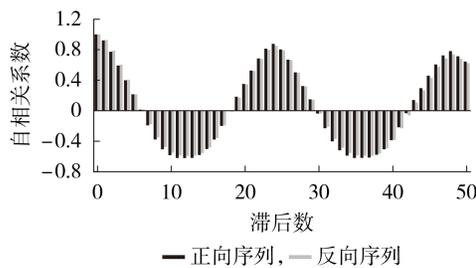


图1 正、反向负荷序列的自相关系数

Fig.1 Autocorrelation coefficients of positive and negative load sequences

由图 1 可知,负荷序列具有明显的正、反向规律性,因此,在负荷预测时要同时考虑历史负荷和未来负荷对预测精度的影响。BiLSTM 神经网络由正向和反向 2 个 LSTM 神经网络组成,相较于传统的 LSTM 神经网络,BiLSTM 神经网络同时考虑了前、后向数据的内在规律<sup>[16]</sup>,是从历史和未来 2 个方向展开预测,使用 2 个独立的隐含层分别处理来自正向和反向的数据,然后将正向和反向的隐含层输出均

作为输出层的输入,从而中和相位误差,提高预测精度。BiLSTM 神经网络结构示意图见附录 A 图 A3。

### 2.2 Attention 机制

深度学习中的 Attention 机制是一种模拟人脑 Attention 的模型。在处理特定信息时,人脑会将 Attention 集中在关键区域,以获得更多需要关注的信息<sup>[17]</sup>。Attention 机制给输入特征赋予不同的权重,以突出强相关因素的影响,减少弱相关因素的影响,相较于分心模型,其可以得到更好的预测效果。Attention 结构见附录 A 图 A4,其各参数计算公式为:

$$e_t = \nu \tanh(W h_t + b) \quad (10)$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^m e_j} \quad (11)$$

$$C_t = \sum_{j=1}^m \alpha_j h_j \quad (12)$$

式中:  $e_t$ 、 $\alpha_t$  分别为当前  $t$  时刻不同特征对应的权重分数和 Attention 权重;  $\nu$ 、 $W$  为计算 Attention 权重时的多层感知机的权重;  $b$  为计算 Attention 权重时的多层感知机的偏置参数;  $m$  为预测模型输入向量的维度;  $C_t$  为 Attention 机制在  $t$  时刻的输出。Attention 机制自适应地计算并调整原始输入特征对应的隐含层状态值,达到突出重要特征以及弱化次要特征的效果,最后将加权过的隐含层状态值输入 BiLSTM 神经网络进一步挖掘数据的内部规律。

## 3 基于 KELM 的气象数据修正模型

为了进一步提高模型的预测精度,本文利用 KELM 结合气象数据对 Attention-BiLSTM 神经网络模型的预测结果进行修正。

KELM 是在极限学习机的基础上,结合核函数提出的一种改进算法<sup>[18]</sup>。极限学习机的输出与连接权重计算公式为:

$$\begin{cases} f(x) = h(x)\beta = H\beta \\ \beta = H^T(I/C + HH^T)^{-1}T \end{cases} \quad (13)$$

式中:  $x$  为模型输入;  $f(x)$  为模型输出;  $h(x)$ 、 $H$  为隐含层特征映射矩阵;  $\beta$  为隐含层与输出层连接权值;  $I$  为对角矩阵;  $C$  为惩罚系数;  $T$  为目标向量。

为了提升极限学习机的泛化能力和稳定性, KELM 算法使用核矩阵  $\Omega_{ELM}$  代替极限学习机中的随机矩阵  $H$ 。核矩阵定义为:

$$\begin{cases} \Omega_{ELM} = HH^T \\ \Omega_{ELM,i,j} = h(x^i)h(x^j) = k(x^i, x^j) \end{cases} \quad (14)$$

式中:  $x^i$  和  $x^j$  分别为样本中第  $i$  个和第  $j$  个输入向量;  $k(x^i, x^j)$  为核函数,通常定义为径向基函数(RBF)核,如式(15)所示。

$$k(x^i, x^j) = \exp(-\gamma \|x^i, x^j\|^2) \quad (15)$$

式中: $\|\mathbf{x}^i, \mathbf{x}^j\|$ 为样本的欧氏范数; $\gamma$ 为核参数。

根据以上计算,KELM的输出与连接权值为:

$$\begin{cases} f(\mathbf{x}) = \begin{bmatrix} k(\mathbf{x}, \mathbf{x}^1) \\ k(\mathbf{x}, \mathbf{x}^2) \\ \vdots \\ k(\mathbf{x}, \mathbf{x}^N) \end{bmatrix}^T (I/C + \mathbf{\Omega}_{ELM})^{-1} \mathbf{T} \\ \boldsymbol{\beta} = (I/C + \mathbf{\Omega}_{ELM})^{-1} \mathbf{T} \end{cases} \quad (16)$$

式中: $N$ 为输入样本数。

KELM属于单隐含层前馈神经网络,计算速度快,其拟合能力相较于普通的极限学习机和支持向量机有较大提升,且由于核函数直接采用内积形式,隐含层节点无需单独设置。综上,本文选择 KELM 算法建立修正模型。

本文构建基于 Attention-BiLSTM 神经网络的预测模型,挖掘历史负荷数据的内在规律性。同时,为了综合考虑影响负荷数据的气象因素,进一步提高预测精度,建立基于 KELM 的气象数据修正模型。建模流程如附录 A 图 A5 所示。建模的具体步骤如下。

1)数据预处理。首先对输入数据序列进行异常值处理,然后再将数据序列按式(17)归一化至区间 $[0,1]$ 中。

$$y^* = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \quad (17)$$

式中: $y$ 为原始数据; $y^*$ 为归一化后的数据; $y_{\max}$ 、 $y_{\min}$ 分别为数据集中的最大值、最小值。

2)负荷预测和误差序列生成。按照第 2 节建立 Attention-BiLSTM 神经网络预测模型,将处理后的历史负荷数据作为输入特征,采用滚动预测的方式进行模型的训练,分别得到训练集上的预测值序列 $y'_i$ 和测试集上的预测值序列 $y''_i$ ,在此基础上求取真实值 $y_i$ 与 $y'_i$ 之差,得到训练集上的误差序列 $y_{res}$ :

$$y_{res} = y_i - y'_i \quad (18)$$

3)误差预测。将误差序列 $y_{res}$ 与处理后的气象数据作为输入特征,训练 KELM,得到测试集的误差预测值序列 $y'_{res}$ 。

4)误差修正。用步骤 3)中得到的误差预测值序列 $y'_{res}$ 修正步骤 2)中的预测结果 $y''_i$ ,得到最终的预测值 $\tilde{y}_i$ :

$$\tilde{y}_i = y''_i + y'_{res} \quad (19)$$

#### 4 算例仿真

为验证本文所提方法的科学性和可靠性,选取我国东部某地区连续两年的真实负荷数据和气象数据作为实验数据集进行仿真验证。

本文实验所用 PC 机配置为 Windows 10 64 位

操作系统,CPU 为 Inter Core i5-8400,内存为 8 GB。软件平台为 Spyder,程序运行环境为 keras2.0.6 和 Tensorflow1.13.1。

本文误差指标选用平均绝对百分比误差(MAPE) $e_{MAPE}$ 和均方根误差(RMSE) $e_{RMSE}$ ,计算公式分别见附录 A 式(A1)、(A2)。

#### 4.1 Attention 机制分析

为研究 Attention 机制对预测模型的影响,选取数据集中一个月的负荷数据,对不同输入特征分配的 Attention 权重如图 2 所示。

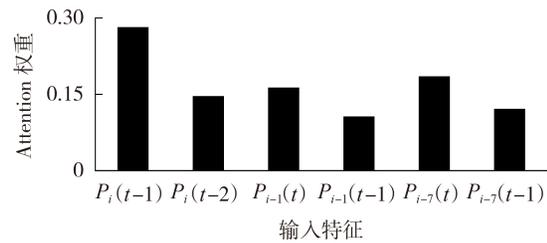


图 2 不同输入特征的 Attention 权重分配情况

Fig.2 Distribution condition of Attention weight for different input features

由图 2 可知,Attention 机制关联了不同输入特征序列对预测结果的影响,其中模型对前一时刻负荷序列 $P_i(t-1)$ 分配了最高的 Attention 权重,其对当前负荷的预测结果影响最大,而 $P_{i-1}(t-1)$ 的权重系数最小,因此对预测结果的影响最小。

图 3 为所选数据集时间跨度由 5 d 增加至 50 d 时,各输入序列对应 Attention 权重的变化情况。由图可知,各输入序列的 Attention 权重随着时间跨度的增加而动态变化并逐渐趋于稳定。Attention 机制不是简单地将各影响因素贡献率平均化,而是在各时刻自主挖掘不同输入特征对预测结果的影响程度,进而提高模型对关键特征信息的提取能力。

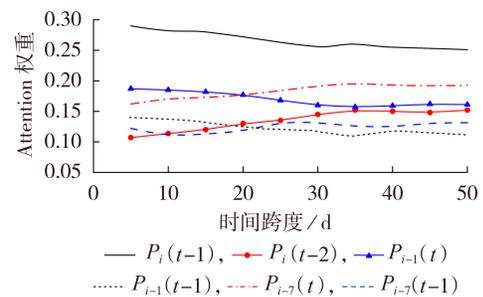


图 3 Attention 权重的变化

Fig.3 Change of Attention weight

#### 4.2 气象修正模型对预测精度的影响分析

由 1.3 节可知,气象数据对负荷的影响存在季节性差异。为研究气象修正模型在不同季节时对预测精度的影响,将该地区一年的负荷数据和气象数据划分成 4 个数据集,分别为春季(3 至 5 月)、夏季(6

至8月)、秋季(9至11月)和冬季(1、2、12月)。将4个数据集按照7:3的比例划分为训练集和测试集,使用多步迭代预测的方法对比本文基于 Attention-BiLSTM 神经网络和气象数据修正的模型(模型1)与 Attention-BiLSTM 神经网络(模型2)的预测精度。其中,模型1用历史负荷数据训练 Attention-BiLSTM 神经网络模型,得到初步的预测误差,再结合气象数据作为 KELM 模型的输入进行误差的预测和修正。模型2的输入特征为历史负荷数据和对应时刻的气象数据。为保证实验的可靠性,两模型中 BiLSTM 神经网络的结构相同,均为2层 BiLSTM 神经网络层,各层神经元个数分别为50、40个,时间步为10,迭代次数为500次。

各季节负荷在测试集上的预测误差如表3所示。气象数据修正模型使得整体预测精度在春、夏、秋、冬季分别提高了0.54%、0.73%、0.45%、0.57%。这表明,相较于模型2使用 Attention-BiLSTM 神经网络模型处理所有的特征,模型1考虑不同类型特征各自的特点,根据特征序列内部规律性强弱分别对其进行处理,充分发挥了 Attention-BiLSTM 神经网络模型处理时序数据的优势,同时建立基于 KELM 的气象修正模型,避免了将规律性差的气象数据输入 BiLSTM 神经网络,有效提高了预测精度。此外,夏季时气象因素与负荷序列相关性最强,气象修正效果最明显。

表3 气象修正模型对预测精度的影响

Table 3 Influence of meteorological correction model on prediction accuracy

季节	模型	$e_{MAPE} / \%$	$e_{RMSE} / MW$
春季	1	1.56	36.26
	2	2.10	52.21
夏季	1	1.92	49.33
	2	2.65	66.93
秋季	1	1.68	37.76
	2	2.13	52.98
冬季	1	1.78	39.16
	2	2.35	53.50

#### 4.3 不同模型的预测结果及分析

为进一步研究本文模型的预测性能,分别用本文模型、Attention-LSTM 神经网络模型、BiLSTM 神经网络模型、LSTM 神经网络模型和 BP 神经网络模型对该地区进行一天和连续一周的预测。首先将数据集进行重新划分,前18个月的数据为训练集,后6个月的数据为测试集。实验结果如表4所示。

由表4可知:相较于其他对比模型,本文模型在一天和一周的预测结果中都达到了最好的效果,这说明本文模型具有较高的精度和较强的鲁棒性;相较于 LSTM 神经网络模型,增加 Attention 机制使得 Attention-LSTM 神经网络模型整体预测精度在一天

表4 不同模型的预测误差对比

Table 4 Comparison of prediction errors among different models

模型	一天		一周	
	$e_{MAPE} / \%$	$e_{RMSE} / MW$	$e_{MAPE} / \%$	$e_{RMSE} / MW$
本文模型	1.82	37.31	1.92	39.48
Attention-LSTM 神经网络	2.55	59.16	2.86	75.92
BiLSTM 神经网络	2.33	48.66	2.62	70.01
LSTM 神经网络	2.76	78.39	3.15	86.36
BP 神经网络	4.87	123.54	5.31	133.80

和一周范围内分别提高了0.21%和0.29%,这表明 Attention 机制可以更好地突出关键特征对预测结果的影响;相较于 LSTM 神经网络模型,BiLSTM 神经网络模型的整体预测精度分别提高了0.43%和0.53%,这表明 BiLSTM 神经网络可以更好地挖掘负荷数据的正、反向规律,进而提高预测精度。

为了更直观地展示不同模型的预测效果,选取不同模型一天的负荷预测值和真实值进行对比,如图4所示。

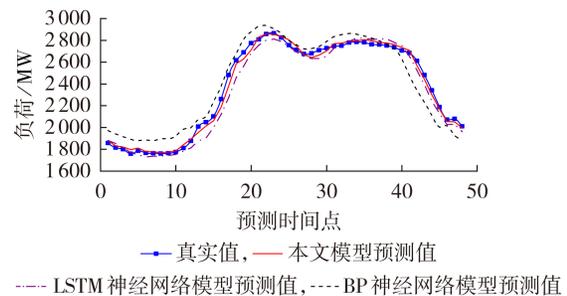


图4 不同模型的预测结果

Fig.4 Predicted results of different models

由图4可知,所选预测日的负荷曲线近似为双峰曲线,在10:00—20:00时间段(图中预测时间点20—40)负荷曲线波动较大,相较于其他模型,本文模型在负荷曲线波峰和波谷区域的波动性最小。此外,在早晚负荷曲线的上升和下降阶段,本文模型也能很好地捕捉负荷变化的规律。总体上看,本文模型的预测结果与真实值拟合程度最高,变化趋势基本一致。

## 5 结论

针对电力系统短期负荷预测精度要求日益提高的要求,本文提出一种基于 Attention-BiLSTM 神经网络和气象数据修正的短期负荷预测模型,并通过实验进行验证。本文主要结论如下。

1)采用 MIC 方法验证了历史负荷数据和气象数据中的温度和相对湿度数据对当前负荷数据存在较大的影响。

2)建立基于 Attention-BiLSTM 神经网络的预测

模型, BiLSTM 神经网络可以充分挖掘历史负荷数据的时序性和正、反向信息, Attention 机制通过给输入特征赋予不同的权重, 突出了对负荷预测精度影响更大的特征。

3) 针对 BiLSTM 神经网络在处理温度数据、湿度数据等内部规律性不强的数据时精度会降低的问题, 建立基于 KELM 的气象数据修正模型, 充分发挥了 BiLSTM 神经网络处理时序数据的优势和 KELM 学习速度快、非线性映射能力强的特点。另外, 引入气象修正模型可以减少时序预测模型输入特征的维度, 进而节约时序预测模型的运算成本。

未来可以研究在考虑降雨量、日类型等影响因素更复杂环境下的短期负荷预测问题, 并且可以与其他负荷预测方法相结合进一步提高模型的预测精度和普适性。

附录见本刊网络版 (<http://www.epae.cn>)。

#### 参考文献:

- [1] 刘继东, 韩学山, 楚成博, 等. 考虑非气象因素的电网夏季降温负荷研究[J]. 电力自动化设备, 2013, 33(2): 40-46.  
LIU Jidong, HAN Xueshan, CHU Chengbo, et al. Cooling load of summer grid considering non-meteorological factors[J]. Electric Power Automation Equipment, 2013, 33(2): 40-46.
- [2] 李妮, 江岳春, 黄珊, 等. 基于累积式自回归平均传递函数模型的短期负荷预测[J]. 电网技术, 2009, 33(8): 93-97, 103.  
LI Ni, JIANG Yuechun, HUANG Shan, et al. Short-term load forecasting based on ARIMA transfer function model[J]. Power System Technology, 2009, 33(8): 93-97, 103.
- [3] 陈昊, 王玉荣. 基于随机波动模型的短期负荷预测[J]. 电力自动化设备, 2010, 30(11): 86-89.  
CHEN Hao, WANG Yurong. Short-term load forecasting based on SV model[J]. Electric Power Automation Equipment, 2010, 30(11): 86-89.
- [4] 刘渝根, 陈超. 基于人工蜂群算法优化支持向量机的接地网腐蚀速率预测模型[J]. 电力自动化设备, 2019, 39(5): 182-186, 200.  
LIU Yugen, CHEN Chao. Corrosion rate prediction model of grounding grid based on support vector machine optimized by artificial bee colony algorithm[J]. Electric Power Automation Equipment, 2019, 39(5): 182-186, 200.
- [5] 程启明, 陈路, 程尹曼, 等. 基于 EEMD 和 LS-SVM 模型的风电功率短期预测方法[J]. 电力自动化设备, 2018, 38(5): 27-35.  
CHENG Qiming, CHEN Lu, CHENG Yinman, et al. Short-term wind power forecasting method based on EEMD and LS-SVM model[J]. Electric Power Automation Equipment, 2018, 38(5): 27-35.
- [6] 张平, 潘学萍, 薛文超. 基于小波分解模糊灰色聚类和 BP 神经网络的短期负荷预测[J]. 电力自动化设备, 2012, 32(11): 121-125, 141.  
ZHANG Ping, PAN Xueping, XUE Wenchao. Short-term load forecasting based on wavelet decomposition, fuzzy gray correlation clustering and BP neural network[J]. Electric Power Automation Equipment, 2012, 32(11): 121-125, 141.
- [7] 王守相, 周凯, 苏运. 基于随机森林算法的台区合理线损率估计方法[J]. 电力自动化设备, 2017, 37(11): 39-45.  
WANG Shouxiang, ZHOU Kai, SU Yun. Line loss rate estimation method of transformer district based on random forest algorithm[J]. Electric Power Automation Equipment, 2017, 37(11): 39-45.
- [8] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157-166.
- [9] 李鹏, 何帅, 韩鹏飞, 等. 基于长短期记忆的实时电价条件下智能电网短期负荷预测[J]. 电网技术, 2018, 42(12): 4045-4052.  
LI Peng, HE Shuai, HAN Pengfei, et al. Short-term load forecasting of smart grid based on long-short-term memory recurrent neural networks in condition of real-time electricity price[J]. Power System Technology, 2018, 42(12): 4045-4052.
- [10] 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.
- [11] 彭文, 王金睿, 尹山青. 电力市场中基于 Attention-LSTM 的短期负荷预测模型[J]. 电网技术, 2019, 43(5): 1745-1751.  
PENG Wen, WANG Jinrui, YIN Shanqing. Short-term load forecasting model based on Attention-LSTM in electricity market[J]. Power System Technology, 2019, 43(5): 1745-1751.
- [12] 时少英, 刘式达, 付遵涛, 等. 天气和气候的时间序列特征分析[J]. 地球物理学报, 2005, 48(2): 259-264.  
SHI Shaoying, LIU Shida, FU Zuntao, et al. The characteristic analysis of weather and climate time series[J]. Chinese Journal of Geophysics, 2005, 48(2): 259-264.
- [13] 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.
- [14] KINNEY J B, ATWAL G S. Equitability, mutual information, and the maximal information coefficient[J]. PNAS, 2014, 111(9): 3354-3359.
- [15] 吴田, 曹志博, 方玉群, 等. 基于最大信息系数的同塔双回输电线路感应影响因素敏感性分析[J/OL]. 电测与仪表. [2020-12-15]. <http://kns.cnki.net/kcms/detail/23.1202.TH.20201016.1612.010.html>.
- [16] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks[J]. IEEE Transactions on Signal Processing, 1997, 45(11): 2673-2681.
- [17] 崔宇, 侯慧娟, 胥明凯, 等. 基于双重注意力机制的变压器油中溶解气体预测模型[J]. 中国电机工程学报, 2020, 40(1): 338-347, 400.  
CUI Yu, HOU Huijuan, XU Mingkai, et al. A prediction method for dissolved gas in power transformer oil based on dual-stage attention mechanism[J]. Proceedings of the CSEE, 2020, 40(1): 338-347, 400.
- [18] HUANG G B. An insight into extreme learning machines: random neurons, random features and kernels[J]. Cognitive Computation, 2014, 6(3): 376-390.

#### 作者简介:



王继东

王继东(1977—),男,教授,博士,研究方向为电能质量、分布式发电和微电网、智能用电等(E-mail: jidongwang@tju.edu.cn);

杜冲(1995—),男,硕士研究生,主要研究方向为电力系统负荷预测(E-mail: wsdc@tju.edu.cn)。

(编辑 王锦秀)

(下转第 224 页 continued on page 224)

## Convolutional neural network based on attention mechanism for reading recognition of pointer-type meter images

ZHANG Sen<sup>1</sup>, WAN Jilin<sup>1</sup>, WANG Huifang<sup>1</sup>, GUAN Minyuan<sup>2</sup>, YANG Bin<sup>2,3</sup>, LI Fan<sup>2</sup>

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China;

2. Huzhou Power Supply Company of State Grid Zhejiang Electric Power Co., Ltd., Huzhou 313000, China;

3. Huzhou Electric Power Design Institute Co., Ltd., Huzhou 313000, China)

**Abstract:** At present, most researches on the reading recognition of pointer-type meters are based on the detection of pointer line segments. However, this method has many processes and low reading recognition efficiency. Besides, the errors accumulated in the intermediate processes such as meter image calibration and pointer segment fitting can easily cause the pointer inclination angle to deviate from its true value. The image feature mapping method for meter readings is studied from another perspective because this method has the advantage of fewer processes and higher efficiency. Firstly, a two-stream heterogeneous convolutional neural network fused with CBAM (Convolutional Block Attention Module) is constructed, which strengthens the extraction of meter image features, improves the type and distribution of features and the accuracy of meter reading recognition. Then a soft stagewise regression strategy is adopted, which greatly simplifies the size of the model and makes the model easy to deploy. Finally, the proposed method is compared with deep learning models based on the pointer line segment detection and other machine learning models on the recognition accuracy and efficiency of pointer-type meter reading recognition by example. Results show that the proposed method maintains a good balance between the accuracy and efficiency of the meter image reading recognition.

**Key words:** pointer-type meter; reading recognition; convolutional neural network; attention mechanism; stage-wise regression; soft range

(上接第177页 continued from page 177)

## Short-term load prediction model based on Attention-BiLSTM neural network and meteorological data correction

WANG Jidong, DU Chong

(Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin 300072, China)

**Abstract:** Most of the load prediction models based on recurrent neural network take historical load data and other factors such as meteorological data together as input features of prediction model, but the internal regularity of meteorological data is poor, so it is not suitable for the input of recurrent neural network. Aiming at the problem, a short-term load prediction model based on Attention-BiLSTM (Attention Bi-directional Long Short-Term Memory) neural network and meteorological data correction is proposed. The maximum information coefficient is adopted to analyze the main factors that affecting the loads. Considering that the load sequence is long and there exists bi-directional information flow, BiLSTM neural network is used for prediction. The attention mechanism is introduced, the influence of key factors is highlighted by the weight of attention, and the internal regularity of load data is excavated. The kernel extreme learning machine is used for error prediction and correction combined with meteorological data, and the load prediction is completed. The real data of a certain area in eastern China is taken as practical example, and the experimental results show that the proposed model has better prediction effect than other models.

**Key words:** short-term load prediction; maximum information coefficient; Attention mechanism; bi-directional long short-term memory neural network; kernel extreme learning machine

附录 A:

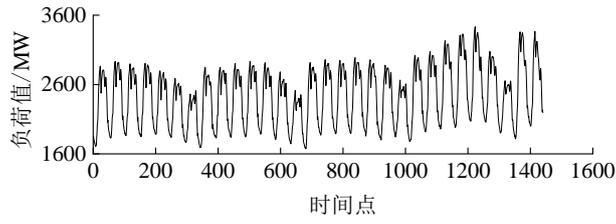
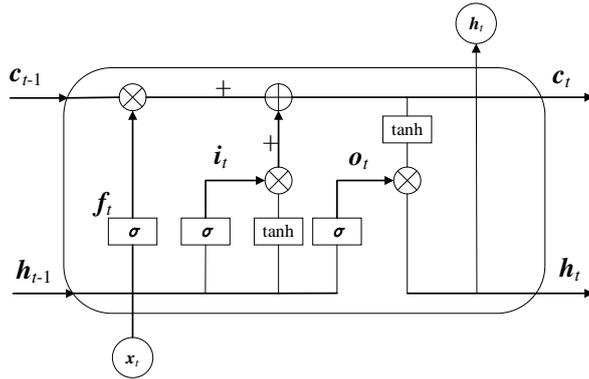


图 A1 历史负荷数据

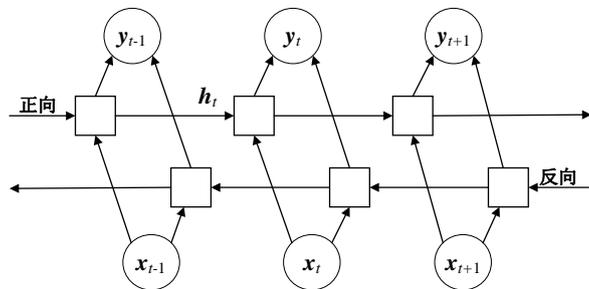
Fig.A1 Historical load data



注：符号“ $\otimes$ ”表示求向量积运算，符号“ $\oplus$ ”表示求和运算。

图 A2 LSTM 神经网络结构

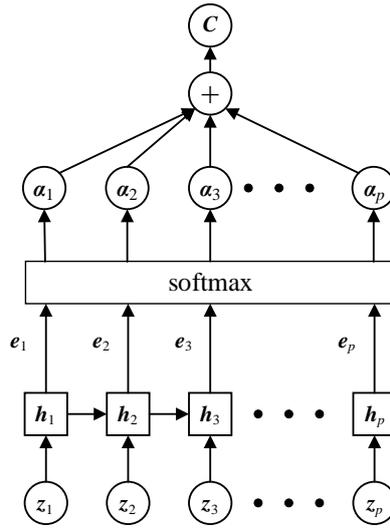
Fig.A2 Structure of LSTM neural network



注：符号“ $\square$ ”表示 BiLSTM 的两层独立的隐含层中的神经元。

图 A3 BiLSTM 神经网络结构图

Fig.A3 Structure of BiLSTMneural network



注：  $z_1, z_2, \dots, z_p$  表示输入序列，  $h_1, h_2, \dots, h_p$  表示对应隐藏层的状态值。

图 A4 Attention 机制

Fig.A4 Attention mechanism

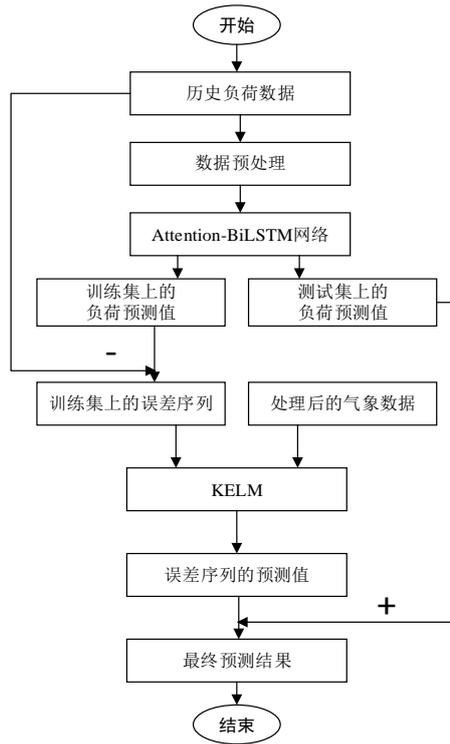


图 A5 建模流程图

Fig.A5 Flowchart of modeling

$$e_{\text{MAPE}} = \frac{1}{k} \sum_{i=1}^k \frac{|l_i - l'_i|}{l_i} \times 100\% \quad (\text{A1})$$

$$e_{\text{RMSE}} = \sqrt{\frac{1}{k} \sum_{i=1}^k (l_i - l'_i)^2} \quad (\text{A2})$$

式中：  $k$  为预测点数；  $l_i$  为第  $i$  个预测点的真实值；  $l'_i$  为第  $i$  个预测点的预测值。