基于改进变分模态分解与深度学习的多因素电力负荷预测

赖小玲，贺嫚嫚*，胡伟，张艺，杜璞良，刘蕊，宋晓彤，郑婷婷
（上海电力大学经济与管理学院，上海 201306）

摘　要：针对传统电力负荷预测方法存在精度较低、负荷数据噪声大等问题，提出一种基于改进变分模态分解（VMD）、卷积神经网络（CNN）和形变长短期记忆（Mogrifier LSTM）网络的多因素电力负荷预测方法。首先，运用麻雀搜索算法（SSA）对变分模态分解进行优化，得到最佳效果的分解子序列，有效减轻负荷数据噪声对预测精度的影响；其次，分析各因素对负荷预测的影响机理，利用皮尔逊相关系数推导各影响因素与负荷之间的相关性，剔除冗余特征，大幅降低模型失准的发生概率；最后，采用 CNN 提取特征向量，将分解后的负荷数据及温度、湿度等特征数据输入到 CNN-Mogrifier LSTM 深度网络模型中，通过 CNN-Mogrifier LSTM 深度网络模型对特征数据进行多维分析，提高短期负荷预测精度。算例分析结果表明，所提出的多因素电力负荷预测模型具有较好的适配性和预测效果，与次优 VMD-CNN-Mogrifier LSTM 模型相比，在两份所用真实数据集上的预测精度分别提升 0.5 和 2.4 百分点，为短期电力负荷预测提供一种可行的解决思路。

关键词：负荷预测；麻雀搜索算法；变分模态分解；长短期记忆网络；相关分析
中图分类号：TP399　　**文献标志码：**A　　**DOI：**10.19678/j.issn.1000-3428.0069241

Multi-factor Power Load Forecasting Based on Improved Variational Mode Decomposition and Deep Learning

LAI Xiaoling, HE Manman*, HU Wei, ZHANG Yi, DU Puliang, LIU Rui, SONG Xiaotong, ZHENG Tingting
(School of Economics and Management, Shanghai University of Electric Power, Shanghai 201306, China)

【Abstract】 To address the problems of low accuracy and large noise of load data in traditional power load forecasting methods, this study proposes a multi-factor power load forecasting method based on improved Variational Modal Decomposition (VMD), Convolutional Neural Network (CNN), and deformed length short-term memory network (Mogrifier LSTM). First, the Sparrow Search Algorithm (SSA) is used to optimize the VMD and obtain the decomposition subsequence with the best effect, which effectively reduces the influence of load data noise on the prediction accuracy. Second, the influence mechanism of each factor on load prediction is analyzed, the correlation between each influencing factor and load is derived using the Pearson′s correlation coefficient, and redundant features are removed, which greatly reduces the probability of model inaccuracy. Finally, a CNN is employed to extract feature vectors. The decomposed load data and feature data such as temperature and humidity are fed into the CNN-Mogrifier LSTM deep network model. The feature data are analyzed in multiple dimensions in this model to improve the short-term load prediction accuracy. The results show that the multi-factor power load prediction model proposed in this study has good adaptability and prediction effects. Compared with the suboptimal VMD-CNN-Mogrifier LSTM model, the prediction accuracy of the proposed model on two real datasets is improved by 0.5 and 2.4 percentage points, respectively, which provides a feasible solution for short-term power load forecasting.

【Key words】 load forecasting; Sparrow Search Algorithm (SSA); variational mode decomposition; Long Short-Term Memory (LSTM) network; correlation analysis

0　引言

电力系统负荷预测是电力系统发电计划的重要组成部分，也是电力系统经济运行的基础，预测的准确与否对整个电力系统的运行、检修、规划都有着至关重要的作用[1]。大规模分布式电源及电动汽车充电桩等多样性负荷的接入带来了更大的负荷波动性、时变性和随机性，进一步加大了预测难度[2]。传统的负荷预测方法比如回归分析法[3]、时间序列法[4]、指数平滑法[5]、卡尔曼滤波法[6]等预测精度较低，不能适应电力负荷序列具有的非线性特点。

收稿日期：2024-01-16　　**修回日期：**2024-03-17
基金项目：国家社科基金项目（19BGL003）。
通信作者 E-mail：*1693600817@qq.com

计算机工程 www.ecice06.com

考虑到负荷序列具有时序性和非线性的特点，现阶段国内外学者展开了大量深入研究。反向传播(BP)神经网络是应用最广泛的神经网络模型之一，但是其具有训练速度慢且容易陷入局部最小值的缺点[7]。长短期记忆(LSTM)能够捕捉到长期依赖信息、缓解梯度消失或爆炸现象，卷积神经网络(CNN)在提取数据特征方面具有出色的能力，所以不少学者将 CNN 与 LSTM 相结合来提高预测精度[8]。文献[9-11]运用 LSTM 深度神经网络并分别与不同的方法相结合对电力负荷进行预测，文献[12-14]将 CNN-LSTM 引入预测模型，有效提高了预测精度，但是 LSTM 存在上下文信息流失的缺陷，而 Mogrifier LSTM 可以有效避免此种缺陷[15]。上述方法能够有效提高电力负荷预测的精度，但是未能考虑到负荷数据噪声对负荷预测的影响。

针对数据降噪的问题，传统的方法主要包括小波域降噪、经验模态分解 (EMD)和变分模态分解(VMD)。文献[16]利用 EMD 将原始电力负荷数据分解为有限数量的模态分量和残差分量，对得到的子序列运用粒子群算法优化的门控循环单元神经网络模型进行预测，但是运用 EMD 分解得到的本征模态函数(IMF)分量往往存在模态混合以及端点效应。文献[17]采用 VMD 将历史电力负荷数据分解为不同频率的分量，并将其输入到改进鲸鱼算法优化的 LSTM 神经网络中进行负荷预测。但是 VMD 的分解模态数 K 和惩罚因子 α 会直接影响其分解效果，所以不少学者将研究聚焦于 VMD 参数的选取优化上。文献[18-19]分别运用粒子群优化(PSO)算法与花朵授粉算法对 VMD 的分解参数 α 与 K 进行寻优，达到了较好的预测效果。麻雀搜索算法(SSA)具有更高性能的全局搜索能力，在精度、稳定性和鲁棒性方面要优于差分进化(DE)算法、灰狼优化(GWO)算法和粒子群优化算法等群智能优化算法[20-21]。

本文提出一种基于改进变分模态分解、卷积神经网络和形变长短期记忆网络的多因素电力负荷预测模型。通过引入 Pearson 系数对特征因素间的相关性进行衡量，剔除不相关因素，并运用 SSA 优化 VMD，得到最佳分解参数，将负荷数据分解为最佳子序列以减轻数据噪声对电力负荷预测的影响。在上述基础上，将分解后的负荷子序列同温度、湿度等特征因素输入到 CNN-Mogrifier LSTM 深度神经网络中进行训练与测试。

1 麻雀搜索算法优化变分模态分解

1.1 变分模态分解

VMD 是受 EMD 启发的最新数据分解技术，其作为一种非递归自适应的信号处理算法，克服了模式混合效应、终点效应等问题，可以有效地将一个多分量信号分解为几个带宽有限的调幅调频信号(IMF)。VMD 不仅具有坚实的理论基础，而且还具有良好的抗噪声鲁棒性，在解决大多数传统信号分解方法中的模态混合问题方面发挥着巨大的优势。

首先构造变分问题：

$$\begin{cases}\min\limits_{\{u_k\},\{\omega_k\}}\left\{\sum\limits_{k=1}^{k}\left\|\partial_t\left[\delta(t)+\dfrac{j}{\pi t}*u_k(t)\right]\mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2\right\}\\ \text{s.t.}\ \sum\limits_{k}^{K}u_k=S\end{cases} \tag{1}$$

式中：$\{u_k\}$ 为各模态函数的集合；$\{\omega_k\}$ 为各模态中心频率的集合；K 为预设分量个数；S 为原输入信号。

然后解决上述的约束最优化问题，将约束变分问题转变为非约束变分问题，利用二次惩罚项和拉格朗日乘子法的优势，引入增广拉格朗日(Lagrangian)函数：

$$\begin{aligned}&L(\{u_k\},\{\omega_k\},\lambda)=\\&\alpha\sum_{k=1}^{K}\left\|\partial_t\left\{\left[\delta(t)+\frac{j}{\pi t}\right]u_k(t)\right\}\mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2+\\&\left\|f(t)-\sum_{k=1}^{K}u_k(t)\right\|_2^2+\\&\langle\lambda(t),f(t)-\sum_{k=1}^{K}u_k(t)\rangle\end{aligned} \tag{2}$$

式中：α 为惩罚参数；λ 为拉格朗日乘子。

利用乘法算子交替方向法求解拉格朗日乘法算子，迭代更新算法如下：

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum\limits_{i\neq k}\hat{u}_i(\omega)+\hat{\lambda}(\omega)/2}{1+2a(\omega-\omega_k)^2} \tag{3}$$

$$\omega_k^{n+1}=\frac{\int_0^{\infty}\omega|\hat{u}_k(\omega)|^2\mathrm{d}\omega}{\int_0^{\infty}|\hat{u}_k(\omega)|^2\mathrm{d}\omega} \tag{4}$$

$$\hat{\lambda}^{n+1}(\omega)=\hat{\lambda}(\omega)+\tau\left(\hat{f}(\omega)-\sum_{k=1}^{K}\hat{u}_k^{n+1}(\omega)\right) \tag{5}$$

式中：τ 表示噪声容限。

满足：

$$\frac{\sum\limits_{k}\|u_k^{n+1}-u_k^{n}\|_2^2}{\|u_k^{n}\|_2^2}<\varepsilon \tag{6}$$

计算机工程 www.ccice06.com

从 VMD 模型可以看出：K 和 α 取值不同，VMD 分解性能也不同。对于 K 而言，若 K 设置较小，则负荷数据便不能被充分分解，可能会出现多个分量都包含在同一个模态中，即欠分解现象；若 K 设置较大，则会出现一个模态分量包含在多个模态中，导致模态混叠，即过分解现象。对于 α 而言，如果 α 值设置较大，那么带宽限制将变得更加狭窄，这将导致有效的频率成分被抑制；相反，冗余的频率成分将会得到保留。

1.2 麻雀搜索算法

SSA 算法是一种革命性的优化技术，它以寻找-跟随-预警的抽象模型为基础，模拟了麻雀的觅食过程，以期获得更优的解决方案，它是由 XUE 等[22]于 2020 年提出的一种新的优化算法。其寻找者位置更新公式如下：

$$X_{ij}^{t+1}=\begin{cases}X_{ij}^{t}\cdot \mathrm{e}^{\left(\frac{-i}{\alpha\cdot M}\right)}, & R_2<S_{ST}\\ X_{ij}^{t}+Q\cdot L, & R_2\geqslant S_{ST}\end{cases} \tag{7}$$

式中：t 表示当前迭代次数；α 为(0,1]的随机数；Q 为服从正态的随机数；R_2 为预警值；S_{ST} 为安全值。

其跟随者位置更新公式如下：

$$X_{ij}^{t+1}=\begin{cases}Q\cdot \mathrm{e}^{\left(\frac{X_{wj}^{t}-X_{ij}^{t}}{i^2}\right)}, & i>\dfrac{n}{2}\\ X_{pj}^{t+1}+\left|X_{ij}^{t}-X_{pj}^{t+1}\right|\cdot \boldsymbol{A}^{+}\cdot L, & \text{other}\end{cases} \tag{8}$$

式中：X_{pj} 表示 $t+1$ 次迭代时最好的位置；X_{wj} 表示 t 次迭代处在全局最差的位置；$\boldsymbol{A}$ 为值为−1 或 1 的 $1\times j$ 矩阵。

为了验证 SSA 算法的性能，采用 cec2005 测试函数集对 SSA、PSO、DE 和 GWO 进行对比分析。cec2005 是使用最广泛且比较经典的测试集，包含 23 个基准函数，其中，$f_1\sim f_5$ 为单峰函数，能够衡量算法的收敛速度与寻优能力，$f_6\sim f_{12}$ 为基本的多峰函数，$f_{13}\sim f_{14}$ 为扩展的多峰函数，$f_{15}\sim f_{23}$ 为多峰组合函数，能够衡量算法的局部最优逃逸能力和勘探能力。选择的对比算法 PSO、DE 和 GWO 分别是经典群智能优化算法、进化算法和新型群智能优化算法的典型代表。

设定种群数为 50，最大迭代次数为 500 次。选择对优化算法挑战性极大的 f_5、f_8 及其他共 4 个函数对 SSA 的性能进行测试，如图 1 所示。

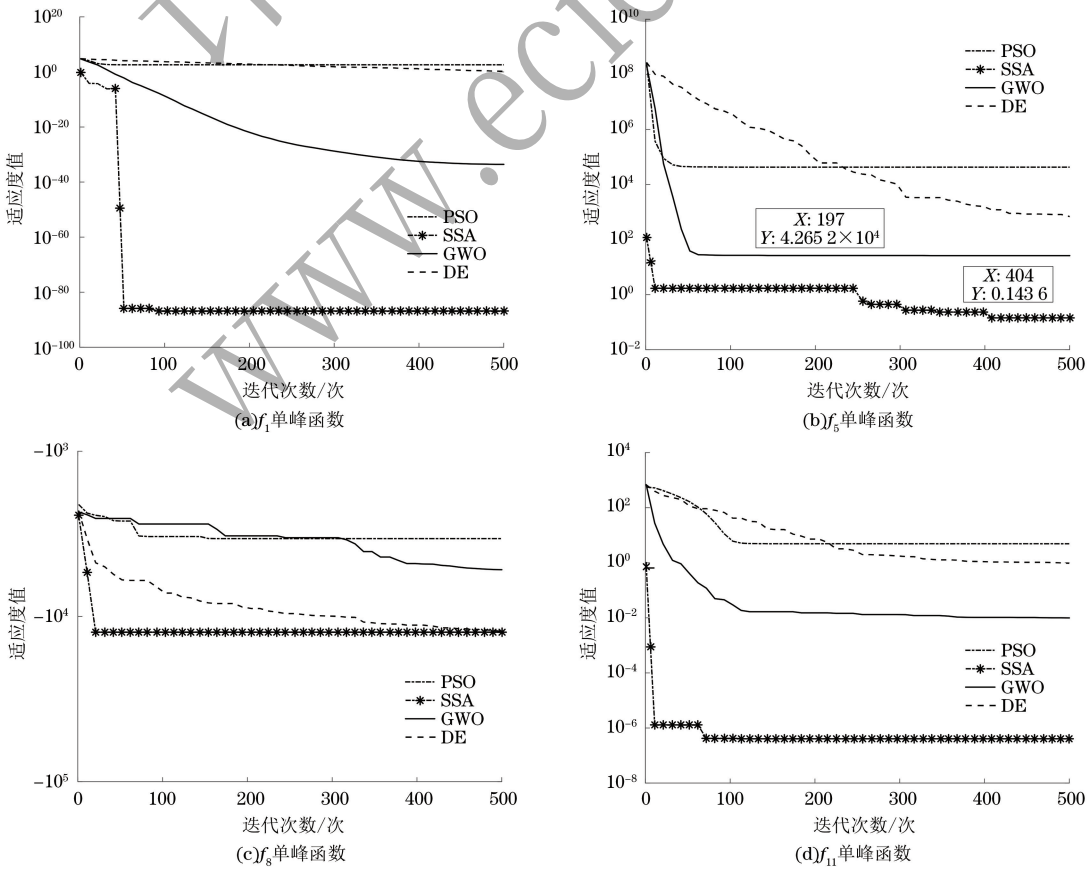


图 1 4 种优化算法适应度曲线

Fig. 1 Fitness curves for four optimization algorithms

从图 1(a)可以看出，相比 GWO，SSA、PSO 和 DE 在迭代 500 次后都能够收敛，但 SSA 的收敛速度较快，其寻优精度也优于其余 3 种优化算法；从图 1(b)可以看出，除 DE 迭代 500 次后不能收敛

计算机工程 www.ecice.com

外，其余 3 种优化算法均收敛，SSA 相比 GWO 收敛速度较慢，但 SSA 寻优得到的值最小，为 0.143 6，非常接近 f_5 基准函数的理论极小值 0；由图 1(c)可知，SSA 能迅速收敛得到最优值，且其最优值非常接近 f_8 函数的理论最小值，为 −12 569.5；如图 1(d)所示，相比其他 3 种优化算法，SSA 的适应度曲线存在多个局部最优值，但 SSA 能够及时跳出局部最优的陷阱，向最优值收敛。综上所述，SSA 具有良好的性能，无论是在收敛速度、寻优精度，还是在局部最优逃逸能力方面都有良好的表现。

1.3 麻雀搜索算法优化变分模态分解

目前，关于 VMD 参数的选取尚无统一的方法，大多数研究人员仍然要通过反复实验确定其参数[23]，另外一部分专家则利用智能优化算法对 VMD 的两个重要参数进行寻优求解。前文对 SSA 的性能进行了测试，证明了其具有较快的收敛速度以及较好的收敛精度，并且不易陷入局部最优值，故本文运用 SSA 优化 VMD，通过寻优确定分解模态数 K 和惩罚因子 α 的最佳值。选用最小包络熵作为 SSA 优化 VMD 的适应度函数。包络熵代表原始信号的稀疏特性：在 IMF 中的噪声较多、特征信息相对较少时，包络熵较大；而当 IMF 中的噪声较少、特征信息相对较多时，包络熵则会较小。根据 ZHANG 等[24]研究中的包络熵概念，包络熵的计算公式如下：

$$E_p = -\sum_{j=1}^{N} p_j \lg p_j \tag{9}$$

$$p_j = \frac{a(j)}{\sum_{j=1}^{N} a(j)} \tag{10}$$

式中：$a(j)$ 是由 VMD 分解的 K 个模态分量经 Hilbert 解调后的包络信号；p_j 是 $a(j)$ 通过计算的归一化得到的概率分布序列；N 为采样点数。

SSA 优化 VMD 的总体流程如图 2 所示，具体步骤如下：

步骤 1 将麻雀种群数以及最大迭代次数初始化，设定 VMD 的分解层数 K 和惩罚因子 α 的范围。

步骤 2 计算初始各麻雀的适应度大小，确定当前最佳的适应度值及其对应位置。

步骤 3 经过一次迭代，重新评估每只麻雀的适应度值，并且实时更新它们的位置信息。根据麻雀当前状态，确定最佳的适应度值和位置。

步骤 4 如果满足最佳适应度，则输出最佳麻雀个体[K，α]，否则，继续执行步骤 2，以确定最佳适应度值。

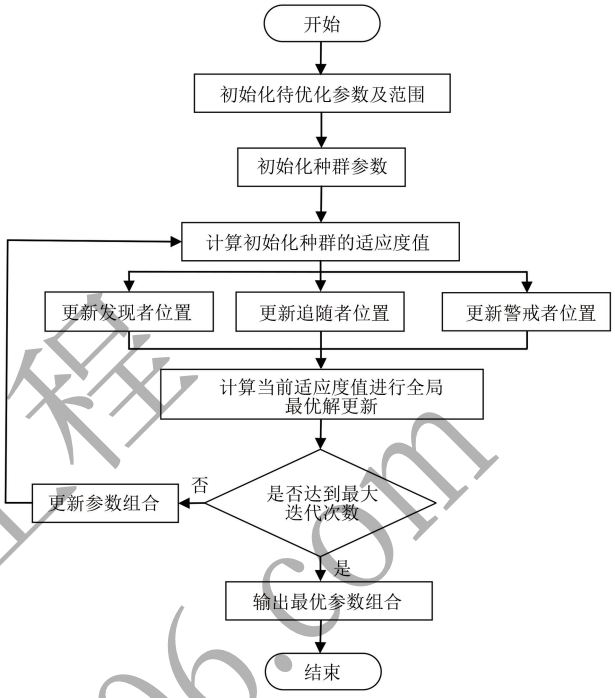


图 2 VMD 的总体流程

Fig.2 Overall procedure VMD

2 CNN-Mogrifier LSTM 深度神经网络

2.1 卷积神经网络

卷积神经网络(CNN)由许多卷积层和池化层组成，以一系列全连接层结束，在每个卷积层与池化层之间都有 ReLU 激活函数来加速模型的收敛。卷积层是 CNN 的核心层，其主要由一些卷积核组成，卷积核的计算公式为：

$$y^{l(i,k)} = \sum_{k'=0}^{n-1} Q_i^{l(k')} x^{l(k+k')} \tag{11}$$

式中：$Q_i^{l(k')}$ 为第 l 层第 i 个卷积核的第 k' 个权值；$x^{l(k+k')}$ 为第 l 层中第 k 个被卷积的局部区域；n 为卷积核的宽度；$y^{l(i,k)}$ 为对应的输出。

2.2 形变长短时记忆网络

长短期记忆(LSTM)是一种递归神经网络(RNN)，它可以学习过去和未来的信息。与传统的 RNN 不同，它可以捕获长序列的文本。LSTM 不仅是用于长序列数据的最成功的深度神经网络，而且它擅于从输入数据中学习隐性知识。其基本单元结构如图 3 所示，计算公式如下：

$$F_t = \sigma(W_F \cdot [h_{t-1}, X_t]) + b_F \tag{12}$$

$$A_t = \sigma(W_F \cdot [h_{t-1}, X_t]) + b_A \tag{13}$$

$$B_t = \sigma(W_B \cdot [h_{t-1}, X_t]) + b_B \tag{14}$$

$$O_t = \sigma(W_O \cdot [h_{t-1}, X_t]) + b_O \tag{15}$$

$$h_t = O_t \times \tanh(C_t) \tag{16}$$

式中：W 表示各类权重；b 为各类偏置；X 代表输

计算机工程 www.ecice06.com

入；h 表示隐层处理状态。

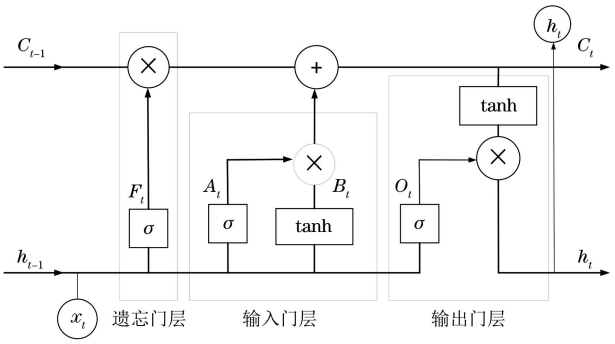


图 3 LSTM 基本单元结构

Fig.3 LSTM basic unit structure

由图 3 的 LSTM 基本单元结构所示，其当前节点的输入 x_t 与上一节点的输出 h_{t-1} 并无直接联系，随着模型变得更加复杂，当前输入和上下文之间的相关性往往会消失。而形变长短期记忆(Mogrifier LSTM)网络的提出解决了 LSTM 上下文无关输入的瓶颈，它以当前输入和先前输出相互选择的形式扩展了 LSTM，将序列输入与上层上下文交互以保持相关性，并通过迭代计算提取隐式特征。其交互计算过程如下：

$$x^i = 2\sigma(\boldsymbol{Q}^i \boldsymbol{h}^{i-1}) \odot \boldsymbol{x}^{i-2}, \text{for odd } i \in [1,2,\cdots,r] \tag{17}$$

$$h^i = 2\sigma(\boldsymbol{R}^i \boldsymbol{x}^{i-1}) \odot \boldsymbol{h}^{i-2}, \text{for even } i \in [1,2,\cdots,r] \tag{18}$$

式中：$\boldsymbol{Q}^i$ 和 $\boldsymbol{R}^i$ 为初始矩阵，i 为超参数，当 $i=0$ 时，整个模型就退化成原始的 LSTM。经过 Sigmoid 运算后，其值分布在(0,1)范围内，因此在式中乘以 2 以保证其数值的稳定性。

Mogrifier LSTM 的基本单元结构如图 4 所示。由图 4 可知，Mogrifier LSTM 将当前输入 x_t 与先前输出 h_{t-1} 进行充分交互，以增强时间序列中的主要特征，然后将更新后的 x_t 和 h_{t-1} 输入到当前节点，进而提升了模型提取数据时间相关性的能力。

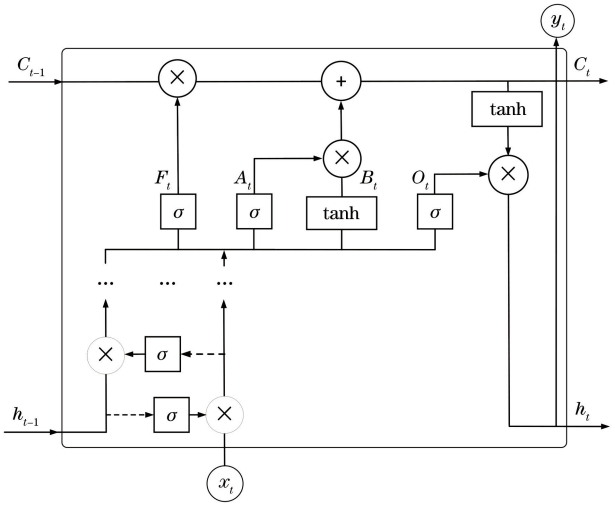


图 4 Mogrifier LSTM 基本单元结构

Fig.4 Mogrifier LSTM basic unit structure

本文所构建的负荷预测模型是基于历史负荷信息以及温度、湿度等特征数据来预测未来某个时点的负荷值，核心在于分析历史数据中的时间序列特征。根据上述分析可知，相比 LSTM，Mogrifier LSTM 在处理复杂的时间序列数据时具有更高的性能，能够充分考虑上下文联系以减少某些主要特征的丢失，因此本文运用 Mogrifier LSTM 进行多因素电力负荷预测能够得到良好的预测效果。

3 改进 VMD-CNN-Mogrifier LSTM 的组合预测模型

3.1 数据预处理

1)数据清洗。对于数据中的缺失值采用线性插值法进行填充。如果在某处缺失单个值，该值可用前一时刻和后一时刻的平均值填充；如果在某处连续缺失多个值，对于某一时刻的值，则可以采用与此时刻距离最近、非缺失的两个历史相同时刻的平均值来填充。对于数据中的异常值，将其视为缺失值，采用缺失值的填充方法对异常值进行处理。

2)数据归一化。由于各项数据的量纲不一致，为了消除量纲不同带来的影响，对数据进行归一化处理。归一化表达式如式(19)所示：

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{19}$$

式中：x 为原始数据；$x_{\min}$ 和 $x_{\max}$ 分别为原始数据集的最大值和最小值。

3.2 特征选择的输入

为了减轻冗余特征对电力负荷预测的影响以及减少不必要信息参与训练导致运行时间过长情况的发生，引入 Pearson 相关系数计算各因素与电力负荷的相关性。Pearson 系数可以衡量变量之间的线性相关性，其值介于[−1,1]之间。两个因素相关系数的绝对值越大，代表这两个因素的相关性越高。根据计算结果，选择与电力负荷相关的特征因素参与后续的负荷预测。

3.3 模型误差评价指标

为了评价所提模型的预测性能，采用决定系数(R^2)、平均绝对百分比误差(MAPE)、平均绝对误差(MAE)以及均方根误差(RMSE)作为模型误差评价指标。

R^2 是解释预测值和真实值之间差异的度量，其值在 0～1 之间，越接近 1 表示预测效果越好。计算

计算机工程 www.ccjen06.com

公式如式(20)所示：

$$R^2=1-\frac{\sum_{i=1}^{N}(\hat{y}_i-y_i)^2}{\sum_{i=1}^{N}(\bar{y}_i-y_i)} \tag{20}$$

MAPE 反映了预测值和真实值的差异，以百分比的形式表示。其值越小，说明模型的预测精度越高。计算公式如式(21)所示：

$$M_{\mathrm{MAPE}}=\frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{y}_i-y_i|}{y_i}\times 100\% \tag{21}$$

MAE 是用来衡量预测值和真实值之间的平均绝对误差，是一个非负值，其值越小表示模型越好。计算公式如式(22)所示：

$$M_{\mathrm{MAE}}=\frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i-y_i| \tag{22}$$

RMSE 是预测值和真实值之间的标准差，其值越小越好。计算公式如式(23)所示：

$$R_{\mathrm{RMSE}}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i-y_i)^2} \tag{23}$$

式中：$\hat{y}_i$ 为模型负荷预测值；y_i 为模型负荷真实值；$\bar{y}_i$ 为模型负荷平均值；N 为预测点数目。

3.4 整体预测步骤

所提模型通过 SSA 优化的 VMD 分解历史负荷数据，得到较为平滑的子序列，去除部分数据噪声；接着利用 CNN-Mogrifier LSTM 层提取数据特征并运用提取到的数据特征进行电力负荷预测。

改进 VMD-CNN-Mogrifier LSTM 组合预测模型的总体流程如图 5 所示，具体步骤如下：

步骤 1 数据预处理。对原始负荷数据进行预处理，包括数据缺失值填充和异常值剔除等步骤。

步骤 2 输入特征选择。运用 Pearson 相关性系数法计算各影响因素与负荷数据之间的相关性，对不相关的因素予以剔除。

步骤 3 参数寻优。初始化麻雀种群参数，进行 SSA 参数寻优，将得到的最佳参数代入 VMD 中对原始负荷数据进行分解，得到若干模态分量。

步骤 4 负荷预测。将各模态分量以及温度、湿度等相关影响因素特征集并排输入到 CNN-Mogrifier LSTM 模型中进行电力负荷预测。

步骤 5 误差评估。将各分量预测值进行重构叠加得到总的负荷预测值，进行模型误差评估。

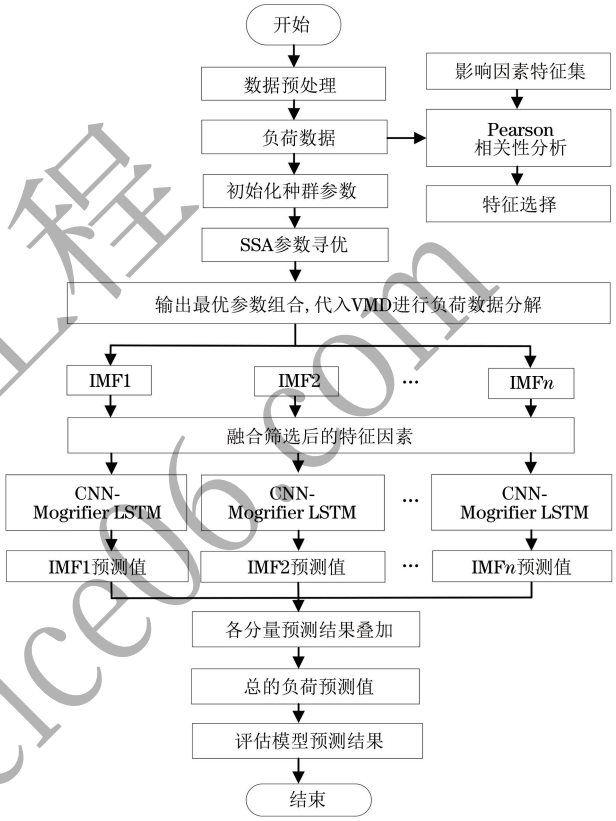


图 5 改进 VMD-CNN-Mogrifier LSTM 的组合预测模型流程

Fig.5 Combinatorial prediction model procedure of improve VMD-CNN-Mogrifier LSTM

4 算例分析

4.1 数据集

本文算例分析的数据来源于南方某地区的实际负荷数据，采样时间为 2018 年 1 月 11 日—2018 年 7 月 10 日，采样间隔为 15 min，共计 17 375 条数据。考虑到电力负荷预测会受到气象等因素的影响，而该数据集包含了最高温度、最低温度、平均温度、相对湿度以及降雨量 5 项特征因素，因此可以认为该数据集具备较强的实用性。部分数据展示如表 1 所示。

表 1 部分数据展示

Table 1 Partial data presentation

采样时间	时间间隔/min	最高温度/℃	最低温度/℃	平均温度/℃	相对湿度(平均)/%RH	降雨量/mm	电力负荷/MW
	0:00	21.6	12.5	15.8	67	0	5 340.320
2018/1/11	0:15	21.6	12.5	15.8	67	0	5 229.377
	0:30	21.6	12.5	15.8	67	0	5 142.122
…				…			
	23:15	34.5	27.7	30	80	5.6	8 845.428
2018/7/10	23:30	34.5	27.7	30	80	5.6	8 640.312
	23:45	34.5	27.7	30	80	5.6	8 491.707

计算机工程 www.ecice06.com

4.2 数据集划分及参数设置

将数据集按照 6∶2∶2的比例对训练集、验证集和测试集进行划分。设定学习率为 0.001，batchsize 为 64。CNN 模型卷积核大小为 2×2，激活函数选择 ReLU，LSTM 模型隐藏层神经元个数为 64。

4.3 输入特征选取结果

采用 Pearson 相关系数法计算最高温度、最低温度、平均温度、相对湿度以及降雨量同电力负荷的相关系数，结果如图 6 所示。

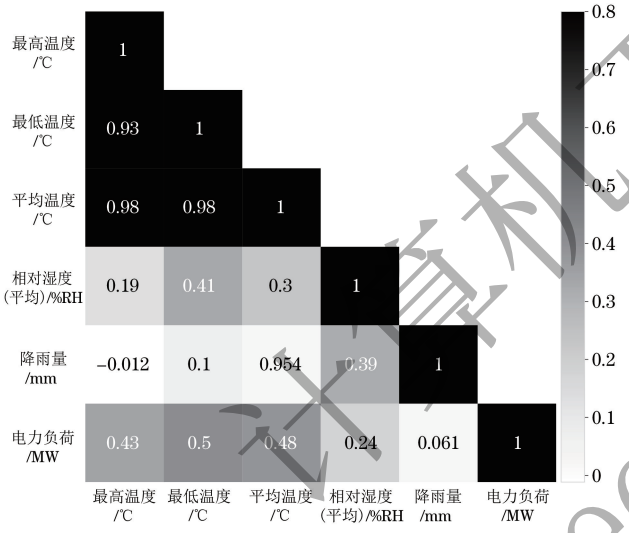


图 6 Pearson 相关系数热力图

Fig.6 Pearson correlation coefficient heat map

由图 6 可知，最高温度、最低温度、平均温度以及相对湿度与电力负荷的线性相关关系较为明显，而降雨量与电力负荷之间的相关系数为 0.061，可近似地认为它们之间无相关关系。故将降雨量舍去，选择最高温度、最低温度、平均温度、相对湿度作为特征因素参与后续的电力负荷预测。

4.4 VMD 分解结果

采用 SSA 算法对 VMD 的最佳参数[K，α]进行寻优，初始参数设置如表 2 所示。适应度函数变化如图 7 所示。

表 2 初始化参数

Table 2 Initialization parameters

最大迭代数/次	采样点/个	α	K
20	30	[100，2 500]	[3，10]

由表 2 和图 7 可知，经过 5 次迭代后包络熵值保持稳定不变，迭代停止，对应的最佳分解参数[K，α]为[5，100]。对负荷数据进行分解，选取前 5 664 个采样点做可视化，其原始数据分布和 VMD 分解结果如图 8 所示。从图 8 中可以看出，原负荷序列随时间变化呈现出一定的波动性，在各种不确定性因素影响

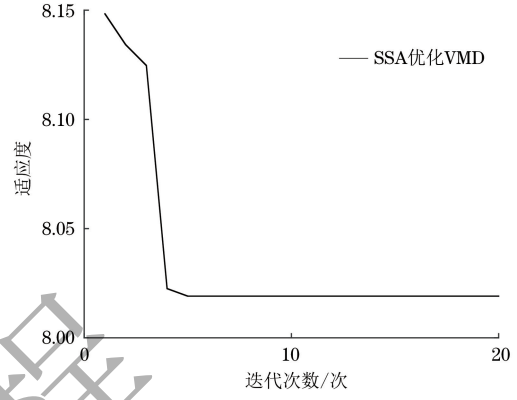


图 7 SSA 优化 VMD 的包络熵变化曲线

Fig.7 SSA optimizes the envelope entropy change curve of VMD

下，负荷数据含有噪声。经过 VMD 分解，负荷波动性显著降低，IMF1 也显示出原始负荷序列整体的变化趋势。

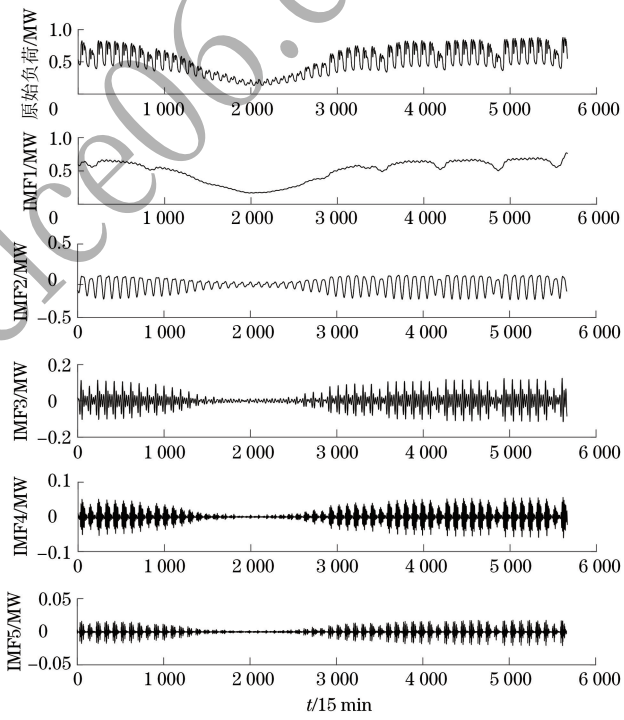


图 8 优化后 VMD 分解结果

Fig.8 The results of VMD decomposition after optimization

4.5 模型对比实验

4.5.1 消融实验

为了证明所提模型 SSA-VMD-CNN-Mogrifier LSTM 不同模块的实际作用，实验设计 5 种模型参与消融实验，即本文所提模型、VMD-CNN-Mogrifier LSTM、CNN-Mogrifier LSTM、Mogrifier LSTM 以及 LSTM 模型。评价指标结果如表 3 所示。

由消融实验的结果可以看出，引入 SSA 优化算法和 VMD 能够显著提高模型的预测性能。SSA 优化算法在收敛速度、寻优精度以及局部最优逃逸能

www.ecice06.com 计算机工程

表 3 消融实验评价指标结果

Table 3 The results of the evaluation index of the ablation experiment

模型	R^2	MAPE/%	MAE/MW	RMSE/MW
SSA-VMD-CNN-Mogrifier LSTM(本文模型)	0.983	0.87	54.27	70.59
VMD-CNN-Mogrifier LSTM	0.978	4.61	269.47	344.34
CNN-Mogrifier LSTM	0.820	11.40	739.21	994.57
Mogrifier LSTM	0.526	23.77	1 360.31	1613.53
LSTM	0.374	31.19	1 600.28	1855.75

力方面都有良好的表现，运用 SSA 对 VMD 进行优化与直接采用 VMD 相比，MAPE、MAE、RMSE 分别降低了 3.74 百分点、215.2 MW、273.75 MW。VMD 能够将时间序列数据分解为多个 IMF 分量，VMD-CNN-Mogrifier LSTM 模型与 CNN-Mogrifier LSTM 模型相比，MAPE、MAE、RMSE 分别降低了 6.79 百分点、469.74 MW、650.23 MW。结果表明，运用 SSA 优化 VMD 进行数据分解，能够得到最佳子序列，减轻数据噪声对负荷预测的影响，提高预测精度。

Mogrifier LSTM 能将先前输入与当前输出信息进行充分交互，其与 LSTM 模型相比，MAPE、MAE、RMSE 分别降低了 7.42 百分点、239.27 MW、242.22 MW。另外，引入 CNN 也在一定程度上提升了模型的预测性能，其具备提取数据集潜在特征的能力，与未引入 CNN 的 Mogrifier LSTM 模型相比，MAPE、MAE、RMSE 分别降低了 12.37 百分点、621.1 MW、618.96 MW。

综上所述，消融实验结果证明了本文所提模型 SSA-VMD-CNN-Mogrifier LSTM 各部分的有效性，SSA 优化、VMD 分解、引入 CNN 以及 Mogrifier LSTM 都有助于提高模型的预测性能，为短期电力负荷预测提供了一种可行的解决思路。

4.5.2 组合模型对比实验

为了验证所提模型 SSA-VMD-CNN-Mogrifier LSTM 的性能，设计以下几种模型进行对比实验：EMD-CNN-Mogrifier LSTM，SSA-VMD-CNN-LSTM，VMD-CNN-LSTM 以及 CNN-LSTM。评价指标结果如表 4 所示。模型预测结果选取某日 96 个采样点做可视化展示，如图 9 所示。

表 4 组合模型对比实验的评价指标结果

Table 4 Evaluation index results of comparative experiments of composite models

模型	R^2	MAPE/%	MAE/MW	RMSE/MW
SSA-VMD-CNN-Mogrifier LSTM(本文模型)	0.983	0.87	54.27	70.59
EMD-CNN-Mogrifier LSTM	0.969	4.66	299.69	412.02
SSA-VMD-CNN-LSTM	0.968	4.82	279.10	352.32
VMD-CNN-LSTM	0.948	8.98	411.36	534.29
CNN-LSTM	0.494	25.91	1 428.20	1 667.86

由表 4 结果可知，相比其他 4 种模型，本文所提模型的预测精度最高，达到了 98.3%。下文对实验结果展开具体分析。

1)不同分解方法对比分析。

综合表 3 及表 4 结果，VMD-CNN-Mogrifier LSTM 模型的预测性能高于 EMD-CNN-Mogrifier LSTM 模型；相比 EMD-CNN-Mogrifier LSTM 模型和 VMD-CNN-Mogrifier LSTM 模型，本文模型的 MAPE 分别降低了 3.79 和 3.74 百分点，MAE 分别降低了 245.42 MW 和 215.2 MW，RMSE 分别降低了 341.43 MW 和 273.75 MW。上述结果表明，EMD 分解效果劣于 VMD 和 SSA-VMD。EMD 分解得到的负荷子序列往往存在着模态混合以及端点效应，VMD 能够克服 EMD 分解过程中产生的模态混叠现象，但是 VMD 参数如果选取不当，则会出现过分解和欠分解现象，影响预测精度；而 SSA 优化 VMD 并对负荷进行分解能够得到最佳分解子序列，改善模型预测性能。

2)不同神经网络模型对比分析。

由表 4 评价指标及图 9 可视化结果可以看出，相比于 SSA-VMD-CNN-LSTM，本文所提模型的预测曲线更贴合真实值，其 MAPE 降低了 3.95 百分点，MAE 降低了 244.83 MW，RMSE 降低了 281.73 MW。表明采用 Mogrifier LSTM 模型进行电力负荷预测能比 LSTM 模型得到更好的效果。

另外将本文模型与 SSA-VMD-CNN-LSTM、VMD-CNN-LSTM 及 CNN-LSTM 3 种模型进行再

计算机工程 www.ciceo.com

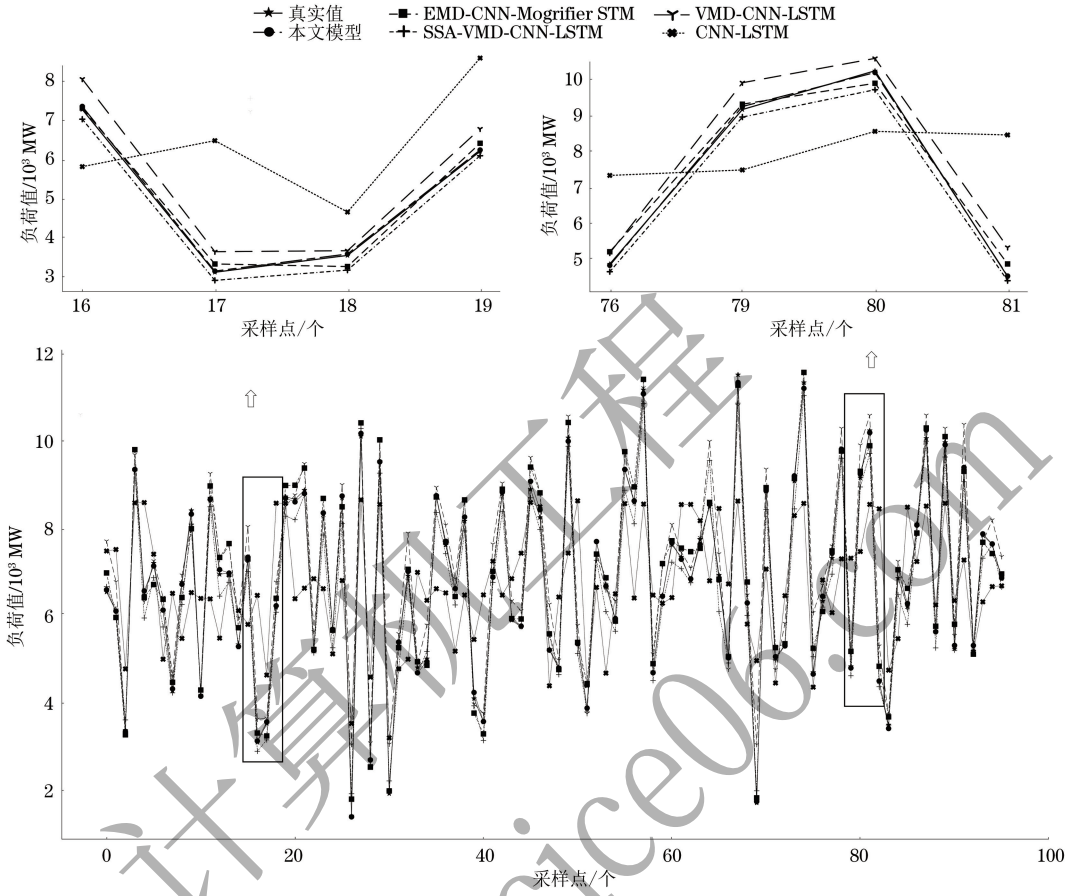


图 9 组合模型对比实验的预测曲线对比

Fig.9 Comparison of prediction curves of combination model comparison experiments

次分析验证，可以得出如下结论：采用 SSA 优化 VMD 比单独使用 VMD 分解负荷数据并进行预测的效果好；在数据输入到后续神经网络中进行预测之前，对负荷数据进行分解去除噪声能够得到较好的预测效果；采用 Mogrifier LSTM 模型进行负荷预测的效果优于 LSTM 模型。这一结论更加有力地支撑了前文的分析结果。

为了进一步体现所提模型的优越性，选择 CNN-BAS-GRU、VMD-CNN-GRU-Attention 以及 VMD-LSTM-LightGBM 进行对比分析。评价指标结果如表 5 所示。模型对比预测结果选取某日 96 个采样点做可视化展示，如图 10 所示。

表 5 4 种模型对比实验的评价指标结果

Table 5 The evaluation index results of the comparison experiment of four models

模型	R^2	MAPE/%	MAE/MW	RMSE/MW
本文模型	0.983	0.87	54.27	70.59
VMD-LSTM-LightGBM	0.966	5.74	329.56	432.86
VMD-CNN-GRU-Attention	0.958	6.53	389.82	526.35
CNN-BAS-GRU	0.934	10.21	548.03	852.73

根据对比实验结果可知，本文所提模型的 4 项误差评价指标值皆优于后 3 种比较模型，具有良好的预测性能，体现出了本文模型的优越性，同时也证明了本文所提模型的有效性。

4.6 所提模型鲁棒性分析

鲁棒性是指模型在陌生环境或者噪声干扰下仍然保持优异性能的能力。通过对数据集施加随机噪声干扰，观察模型预测性能的变化能够验证模型是否存在鲁棒性。本文对数据集施加随机噪声干扰，将干扰后的数据运用本文所提模型进行电力负荷预测，并与不存在噪声干扰数据集的预测结果进行对比分析，选取某日进行可视化，结果如图 11 所示。

从图 11 可以看出，存在噪声干扰的预测曲线与不存在噪声干扰的预测曲线以及真实值曲线的趋势基本一致，不存在异常点。同时，3 条曲线较为接近，说明模型不易受外部噪声干扰的影响，验证了本

计算机工程 www.ecice06.com

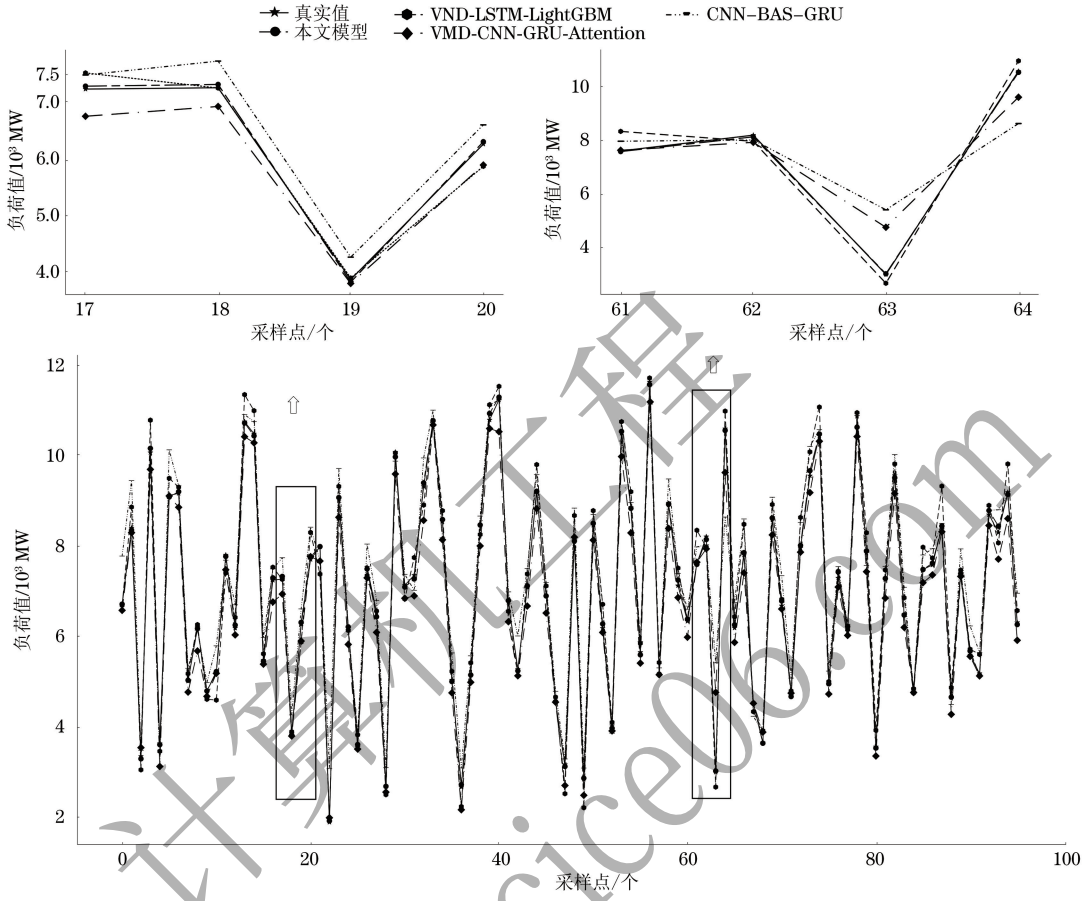


图 10 4 种模型对比实验的预测曲线对比

Fig. 10 Comparison of prediction curves of comparison experiments of four models

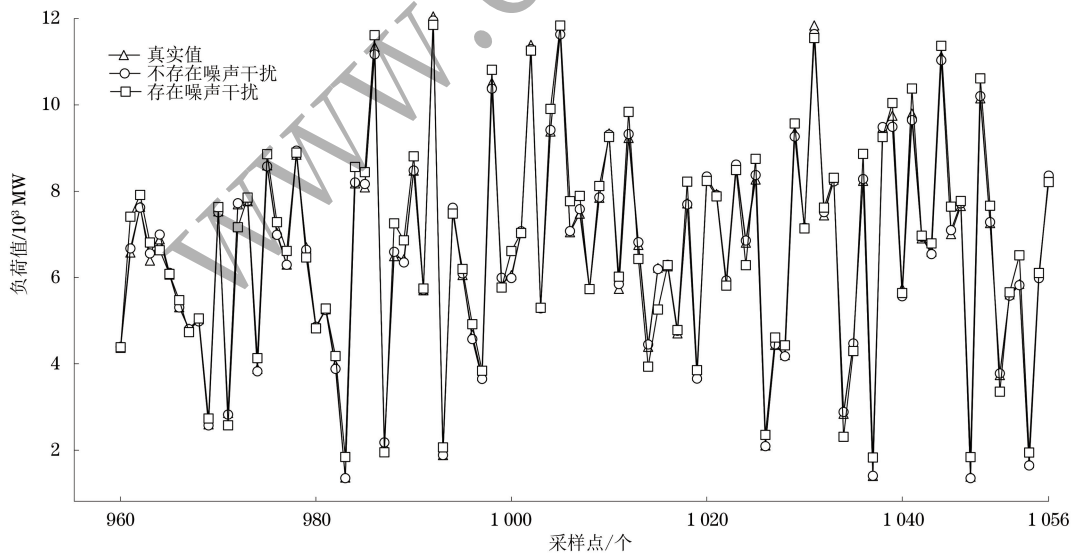


图 11 鲁棒性验证实验预测曲线对比

Fig. 11 Comparison of prediction curves for robustness verification experiments

文所提模型具有鲁棒性。

4.7 所提模型泛化能力验证

为了验证所提模型的预测效果与泛化能力，增加一份数据集进行性能测试。数据样本来源于澳大利亚某州的真实负荷数据集，包含干球温度、露点温度、湿球温度、湿度以及电价 5 项特征因素，采样时间间隔为 30 min，选取采样时间为 2010 年 1 月 1 日—2010 年 5 月 31 日的数据，共计 7 248 条数据，基于本文所提模型进行负荷预测。数据预处理步骤同第 3.1 节，输入特征选择同第 3.2 节，实验设置同第 4.2 节，选取第 4.4 节对比模型中预测效果最好的一个模型进行对比实验，其评价指标结果如

计算机工程 www.cice06.com

表 6 所示。由表 6 可知，本文模型在澳大利亚某州数据集上表现出了良好的预测效果，说明本文模型在面对不同数据集时仍然具有较强的预测能力，证明所提模型具备较高的泛化能力。

表 6 泛化能力验证的评价指标结果

Table 6 Evaluation index results of generalization ability verification

模型	R^2	MAPE/%	MAE/MW	RMSE/MW
本文模型	0.984	0.90	80.83	102.17
VMD-CNN-Mogrifier LSTM	0.960	1.69	156.47	200.73

4.8 模型参数量及预测时间对比

从模型的参数量以及预测时间两方面对所提模型的计算复杂性进行分析。模型参数量是评估模型是否轻量化的关键因素，而预测时间则可以对模型的运行性能进行评估。部分模型参数量及预测时间对比结果如表 7 所示。由表 7 可知，本文模型参数量为 104.28×10^3，相比于 SSA-VMD-CNN-LSTM 模型，参数量减少了 519.17×10^3；本模型与 CNN-Mogrifier LSTM 和 EMD-CNN-Mogrifier LSTM 的参数量较为接近。另外，虽然本文模型的参数量远远大于 Mogrifier LSTM 与 BPNN，但是 Mogrifier LSTM 与 BPNN 结构简单，无法满足实际负荷预测的需要。对各模型运行 1 天、2 天、4 天以及 8 天(96 个采样点/天)的运行时间进行统计，可知本文所提模型运行时间较长一点，但是与运行时间短的模型差距不大，在既有条件下运行仍然可以较快地得到预测结果。虽然本文模型牺牲了一定的计算效率，但是其在负荷预测精度方面却得到了较大提升。

综上所述，本文所提模型的计算复杂性不高，能够满足实际在线应用运算要求。

表 7 模型参数量及时间开销对比结果

Table 7 Comparison results of model parameter quantity and time overhead

模型	参数量/10^3	不同采样点下的预测时间/s			
		1 天	2 天	4 天	8 天
本文模型	104.28	2.71	5.46	10.75	21.22
CNN-Mogrifier LSTM	103.92	2.88	5.67	11.04	21.87
Mogrifier LSTM	78.95	5.12	10.40	20.65	42.73
EMD-CNN-Mogrifier LSTM	104.93	2.61	5.03	10.36	20.64
SSA-VMD-CNN-LSTM	623.45	0.76	1.42	2.84	5.63
VMD-CNN-LSTM	623.66	0.62	1.36	2.73	5.31
CNN-LSTM	623.09	0.71	1.31	2.62	5.30
BPNN	25.70	0.11	0.20	0.36	0.66

4.9 模型应用的可能限制讨论

1)数据获取难度。

在实际情况中，电力负荷会受到多种因素的影响，比如气象、节假日等，但由于这些数据可能来自于不同的机构或者传感器，从而造成数据信息采集异常困难。比如作为重要影响因素之一的气象因素就常常面临着数据缺失与不足的情况，这进一步加重了在建模阶段的小样本问题。

2)数据获取实时性。

在数据的采集、传输和处理过程中可能会存在延迟，从而无法获取到实时性数据。数据的质量被视为电力系统的负荷预测建模的重要基础和预测性能的关键保证[25]，若无法获取到大量真实有效且具有实时性的数据，模型的预测性能也会大打折扣。

3)难以集成到现有的电力系统管理工具中。

首先将负荷预测模型接入到现有的电力系统管理工具中时可能会存在模型与工具接口不兼容的问题。其次负荷预测模型与电力系统管理工具集成需要考虑其与电力系统管理工具中其他模块之间的协同，对此需要重新对系统进行设计与开发。最后负荷预测模型成功集成到现有的电力系统管理工具中后，还需要对相关工作人员进行培训。上述问题所产生的时间和成本使得负荷预测模型难以集成到现有的电力系统管理工具中。

5 结束语

本文提出一种基于改进变分模态分解和深度学习的多因素电力负荷预测方法。运用 SSA 改进的 VMD 对负荷数据进行分解去噪，以减轻其对电力负荷预测精度造成的不利影响；同时充分考虑各种特征因素对负荷产生的外部影响，通过 Pearson 系数分析各因素与负荷之间的相关性，并进行输入特征选取，以减少不相关因素对负荷预测精度的影响；最后利用 CNN 提取数据集潜在特征的能力以及

计算机工程 www.ecice06.com

Mogrifier LSTM 捕捉数据长期成分的能力，将经 SSA-VMD 分解的各模态向量与相关特征因素集输入到 CNN-Mogrifier LSTM 深度神经网络中进行负荷预测。在真实数据集上的预测结果表明，所提模型预测性能良好，具有较优的鲁棒性与泛化能力。下一步将研究所提模型在不同噪声水平数据集下的鲁棒性。

参考文献

[1] 吴铁洲，邹智，姜奔，等. 基于 TLBGA-GRU 神经网络的短期负荷预测[J]. 计算机工程，2022，48(11)：69-76.
WU T Z，ZOU Z，JIANG B，et al. Short-term load forecasting based on TLBGA-GRU neural network[J]. Computer Engineering，2022，48(11)：69-76.（in Chinese）

[2] HAN J Y，YAN L，LI Z Y. A task-based day-ahead load forecasting model for stochastic economic dispatch[J]. IEEE Transactions on Power Systems，2021，36(6)：5294-5304.

[3] AKHTAR S，SHAHZAD S，ZAHEER A，et al. Short-term load forecasting models：a review of challenges，progress，and the road ahead[J]. Energies，2023，16(10)：4060.

[4] ACQUAH M A，JIN Y W，OH B C，et al. Spatiotemporal sequence-to-sequence clustering for electric load forecasting[J]. IEEE Access，2023，11：5850-5863.

[5] YAN H F，YU X Y，LI D W，et al. Research on commercial sector electricity load model based on exponential smoothing method[M]. Berlin，Germany：Springer，2022.

[6] LIANG Z，ZHENG C Y，ZHAO Z G，et al. Short-term load forecasting based on Kalman filter and nonlinear autoregressive neural network[C]//Proceedings of the 33rd Chinese Control and Decision Conference. Kunming，China：[s. n.]，2021：3747-3751.

[7] BIAN H H，WANG Q，XU G Z，et al. Research on short-term load forecasting based on accumulated temperature effect and improved temporal convolutional network[J]. Energy Reports，2022，8：1482-1491.

[8] 曹正志，叶春明. 基于并联 CNN-SE-Bi-LSTM 的轴承剩余使用寿命预测[J]. 计算机应用研究，2021，38(7)：2103-2107.
CAO Z Z，YE C M. Prediction of bearing remaining useful life based on parallel CNN-SE-Bi-LSTM[J]. Application Research of Computers，2021，38(7)：2103-2107.（in Chinese）

[9] CHEN Z X，ZHANG D L，JIANG H R，et al. Load forecasting based on LSTM neural network and applicable to loads of “Replacement of coal with electricity”[J]. Journal of Electrical Engineering & Technology，2021，16(5)：2333-2342.

[10] HAN M C，ZHONG J W，SANG P，et al. A combined model incorporating improved SSA and LSTM algorithms for short-term load forecasting[J]. Electronics，2022，11(12)：1835.

[11] LIN J，MA J，ZHU J G，et al. Short-term load forecasting based on LSTM networks considering attention mechanism[J]. International Journal of Electrical Power and Energy Systems，2022，137：107818.

[12] REN C，JIA L，WANG Z L. A CNN-LSTM hybrid model based short-term power load forecasting[C]//Proceedings of Power System and Green Energy Conference. [S. l.]：IEEE Press，2021：182-186.

[13] YANG Z D，LI X F，KONG X Y，et al. A method of short-term load prediction of renewable energy power system based on CNNLSTM[C]//Proceedings of the 25th International Conference on Electrical Machines and Systems. Washington D. C.，USA：IEEE Press，2022：1-5.

[14] SAEED F，PAUL A，SEO H. A hybrid channel-communication-enabled CNN-LSTM model for electricity load forecasting[J]. Energies，2022，15(6)：2263.

[15] SHEN X D，ZHAO H X，XIANG Y，et al. Short-term electric vehicles charging load forecasting based on deep learning in low-quality data environments[J]. Electric Power Systems Research，2022，212：108247.

[16] ZHANG W H，WANG T. Short-term power load forecasting model design based on EMD-PSO-GRU[J]. Scientific Programming，2022，22：4755519.

[17] ZHUANG Z Y，ZHENG X D，CHEN Z X，et al. A reliable short-term power load forecasting method based on VMD-IWOA-LSTM algorithm[J]. IEEJ Transactions on Electrical and Electronic Engineering，2022，17(8)：1121-1132.

[18] 袁东辉，朱愉洁，齐咏生，等. 一种增强型的滚动轴承故障诊断[J]. 计算机仿真，2022，39(10)：526-532.
YUAN D H，ZHU Y J，QI Y S，et al. An enhanced fault diagnosis of rolling bearing[J]. Computer Simulation，2022，39(10)：526-532.（in Chinese）

[19] YANG J J，YAN K，WANG Z，et al. A novel denoising method for partial discharge signal based on improved variational mode decomposition[J]. Energies，2022，15(21)：8167.

[20] YAN S Q，LIU W D，LI X Q，et al. Comparative study and improvement analysis of sparrow search algorithm[J]. Wireless Communications and Mobile Computing，2022，22：4882521.

[21] YUE Y G，CAO L，LU D W，et al. Review and empirical analysis of sparrow search algorithm[J]. Artificial Intelligence Review，2023，56(10)：10867-10919.

[22] XUE J K，SHEN B. A novel swarm intelligence optimization approach：sparrow search algorithm[J]. Systems Science & Control Engineering，2020，8(1)：22-34.

[23] LI J X，LIU Z W，QIU M，et al. Fault diagnosis model of rolling bearing based on parameter adaptive VMD algorithm and sparrow search algorithm-Based PNN[J]. Systems Science & Control Engineering，2020，8(1)：22-34.

[24] ZHANG Z M，LIU C L，WANG R，et al. Mechanical fault diagnosis of a disconnector operating mechanism based on vibration and the motor current[J]. Energies，2022，15(14)：5194.

[25] 韩富佳，王晓辉，乔骥，等. 基于人工智能技术的新型电力系统负荷预测研究综述[J]. 中国电机工程学报，2023，43(22)：8569-8592.
HAN F J，WANG X H，QIAO J，et al. Review on artificial intelligence based load forecasting research for the new-type power system[J]. Proceedings of the CSEE，2023，43(22)：8569-8592.（in Chinese）

编辑 索书志

计算机工程 www.ecice06.com