

文章编号: 2095-2163(2022)12-0142-06

中图分类号: TP301.6

文献标志码: A

# 一种变分模态分解与 Adam 优化的 LSTM 电价预测方法

马丽莹, 魏云冰

(上海工程技术大学 电子电气工程学院, 上海 201620)

**摘要:** 电价预测有助于电力市场的优化调度。长短期记忆(long short-term memory, LSTM)网络作为一种特殊的循环神经网络(recurrent neural network, RNN),在解决此类问题时具有良好的性能。为了提高电价预测的准确性,本文将变分模态分解(variational mode decomposition, VMD)、LSTM和Adam优化算法相结合,构建了一种VMD-Adam-LSTM混合模型。经VMD分解,将原始复杂的电价序列分解为波动简单且数量有限的固有模态函数(intrinsic mode function, IMF)。VMD分解能够有效克服经验模态分解(empirical mode decomposition, EMD)中存在的模态混叠问题。先将高效的随机梯度优化器Adam与LSTM结合,再对分解后的每个IMF进行预测,可以精确捕捉到电价的波动行为。本文将VMD-Adam-LSTM混合模型应用于实际电价数据。并与其他模型相比较,验证了该模型在电价预测上具有良好的性能。

**关键词:** 电价预测;长短期记忆网络;变分模态分解;Adam优化算法

## A LSTM electricity price forecasting method based on variational modal decomposition and Adam optimization

MA Liying, WEI Yunbing

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**[Abstract]** Price forecasting is helpful to optimize the power market. Long short-term memory (LSTM) network, as a special recurrent neural network (RNN), has a good performance in solving electricity price forecasting problems. In order to improve the accuracy of electricity price prediction, a VMD-Adam-LSTM hybrid model is constructed by combining VMD, LSTM and Adam optimization algorithm. The intrinsic mode function (IMF) is used to decompose the original complex electricity price sequence into the simple and limited intrinsic mode function (IMF) after VMD decomposition. VMD decomposition could effectively overcome the modal aliasing problem in empirical mode decomposition (EMD). Firstly, the efficient stochastic gradient optimizer Adam is combined with LSTM, and then the decomposed IMF is predicted, which can accurately capture the fluctuation behavior of electricity price. In this paper, the VMD-Adam-LSTM hybrid model is applied to the actual electricity price data. Compared with other models, it is verified that this model has good performance in electricity price prediction.

**[Key words]** electricity price forecast; long and short term memory network; variational modal decomposition; Adam optimization algorithm

## 0 引言

电价预测问题影响着现代电力系统运行的各个过程。这些过程与其他当代科学和工程问题有关,如最优电网调度、能源消耗、能源开发、温室气体排放、电力系统模拟和电力负荷需求建模<sup>[1-2]</sup>。因此,电价预测是电力系统领域的一个牵涉到多学科的综合研究课题。但是电价预测受到多种因素的影响,例如节假日、温度、节日、季节等,具有一定的波动性。此外,由于电力资源的波动和不平衡有可能造成电力负荷增加,导致电价升高<sup>[3-4]</sup>。近年来,许多研究人员已经证明,电力相关预测是一项艰巨的任

务,因为电力需求在很大程度上要取决于多种因素<sup>[5-7]</sup>。

为了实现对电价的高精度预测,一些学者开始提出一些电价预测的方法,大致可分为2类:传统方法和人工智能方法。传统方法主要包括自回归综合移动平均<sup>[8]</sup>、模糊推理系统<sup>[9]</sup>和贝叶斯随机波动率<sup>[10]</sup>等方法。由于电价变化较快,具有一定的波动性和非线性,传统的预测方法无法取得较好的预测效果。

为了解决非线性的问题,一些研究人员采用人工智能方法实现电价预测。例如极限学习机、支持向量机(support vector machines, SVM)和人工神经

**作者简介:** 马丽莹(1995-),女,硕士研究生,主要研究方向:电力市场中长期交易、电价预测;魏云冰(1970-),男,博士,教授,主要研究方向:电力市场、电力系统自动化、电价预测。

**通讯作者:** 魏云冰 Email: wei.yunbing@sues.edu.cn

收稿日期: 2022-03-07

网络等。Shrivastava 等人<sup>[11]</sup>将极限学习机与小波技术相结合提高了电价预测的精度和可靠性。Yan 等人<sup>[12]</sup>提出了一种基于多 SVM 的中期电力市场出清价格预测模型。Panapakidis 等人<sup>[13]</sup>将人工神经网络与聚类算法相结合提出了一种混合模型, 来实现日前电价预测。然而, 由于电价数据具有一定的记忆性、即存在长期依赖关系, 传统的神经网络在电价时间序列分析的研究中面临着一定挑战。

LSTM 作为一种特殊的 RNN, 能够应用于长期依赖关系的序列预测。Liu 等人<sup>[14]</sup>利用经验小波变换和 LSTM 的混合模型预测风速。实际电价数据具有非线性、非平稳和长期依赖性的特点, 而 LSTM 很好地解决了这些问题。Peng 等人<sup>[15]</sup>在预测电价时采用了差分进化算法来识别 LSTM 的参数, 并通过实验验证了长短时记忆网络在电价预测精度方面明显优于传统的智能计算方法。其主要原因在于 LSTM 能更好地捕捉并处理电价不规则性的特征。基于随机梯度的方法在深度学习优化领域具有十分重要的实际意义, 对预测的准确性起着至关重要的作用。Adam 是一种常见且高效的优化算法<sup>[16-17]</sup>, 其优势在于对内存需求较少, 计算效率高, 可以很好地处理稀疏梯度, 在非线性和非平稳设置下表现出优异的性能。该算法可以通过寻找一系列参数使目标函数最小化来达到优化深度学习模型的目的。因此, Adam 优化的 LSTM 神经网络被认为是预测电价的有力工具。

将 VMD 分解算法和 Adam 优化的 LSTM 神经网络相结合用于电价时间序列的分析和预测是一种比较新颖的想法。VMD 分解方法有效解决了 EMD 中存在的模态混叠问题。因此, 本文提出了一种基于 VMD 和 Adam 优化的 LSTM 网络的混合模型, 即 VMD-Adam-LSTM 模型。通过美国 PJM 的电价数据来验证该模型的预测效果。为了说明该模型的性能, 将其与传统模型和混合模型进行了对比来验证该模型的预测精度。

本文的框架结构如下所示: 第 1 节介绍了 VMD 分解和 Adam 优化 LSTM 的基本理论; 第 2 节详细描述了所提出的混合模型; 在第 3 节中, 给出了电价预测案例; 结论见第 4 节。

## 1 基本算法

在本节中, 将介绍提出的混合模型中采用的 2 种算法。在 1.1 节中, 简要介绍 VMD 分解算法原理, 在 1.2 节中介绍 Adam 优化的 LSTM 神经网络。

络。

### 1.1 变分模态分解

2014 年, Dragomiretskiy 等人<sup>[18]</sup>提出了 VMD 方法<sup>[18-19]</sup>。VMD 方法解决了局部均值分解、经验模态分解等方法中存在的模态混叠问题<sup>[19]</sup>。作为一种非递归算法, VMD 将输入信号同时分解为多个 IMF。这里需用到的数学公式为:

$$u_k(t) = A_k(t) \cos(\phi_k(t)) \quad (1)$$

其中,  $A_k(t)$  是  $u_k(t)$  的瞬时幅值,  $\omega_k = \phi_k'(t)$  是瞬时频率。约束变分问题描述为:

$$\begin{cases} \min_{\{u_k(t)\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left( \frac{e^{j\omega_k t}}{\omega_k} * u_k(t) \right) \right\|_2^2 \right\} \\ \text{subject to } \sum_{k=1}^K u_k = x(t) \end{cases} \quad (2)$$

其中,  $x(t)$  是要分解的原始信号;  $u_k, \omega_k$  分别是 IMF 和中心频率; “\*” 是卷积符号。引入惩罚因子  $a$  和拉格朗日乘数  $\lambda$  来解决方程 (2) 的约束问题:

$$\begin{aligned} L(u_k, \omega_k, \lambda) = & a \sum_k \left\| \partial_t \left( \frac{e^{j\omega_k t}}{\omega_k} * u_k(t) \right) \right\|_2^2 + \\ & \left\| x(t) - \sum_k u_k(t) \right\|_2^2 + \langle \lambda(t), x(t) - \sum_k u_k(t) \rangle \end{aligned} \quad (3)$$

$u_k, \omega_k$  和  $\lambda$  的更新公式见如下:

$$\hat{u}_k^{n+1}(\omega) = \frac{x(\omega) - \sum_{i=1, i \neq k}^K \hat{u}_i^{n+1}(\omega) + \frac{\lambda^n(\omega)}{2}}{1 + 2\varepsilon(\omega - \omega_k^n)^2} \quad (4)$$

$$\hat{\omega}_k^{n+1} = \frac{\int_{\forall} \omega |u_k^{n+1}(\omega)|^2 d\omega}{\int_0^{\forall} |u_k^{n+1}(\omega)|^2 d\omega} \quad (5)$$

$$\hat{\lambda}^{n+1}(\omega) = \lambda^n(\omega) + \tau \left( x(t) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \quad (6)$$

直到满足条件  $\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2 / \|\hat{u}_k^n\|_2^2 < \xi$ , 最终将原始信号分解成  $K$  个 IMF 分量。

### 1.2 Adam 优化的 LSTM 神经网络

LSTM 神经网络拥有很多层连接的记忆块, 每个记忆块包含管理块的输出门和状态门。LSTM 神经网络具有一定的记忆性, 即下一个状态的信息不仅与当前状态有关, 还与历史状态有关。LSTM 每个单元内主要包括遗忘门、输入门和输出门, 如图 1 所示。对此拟展开阐释分述如下。

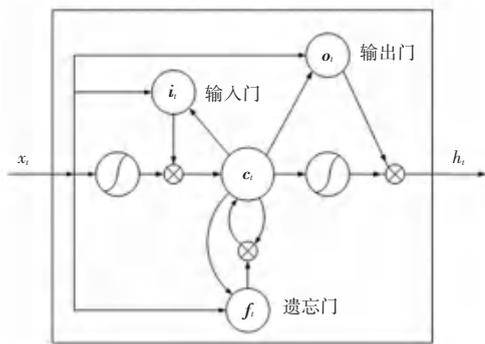


图1 LSTM神经网络结构

Fig. 1 LSTM neural network structure

(1) 遗忘门  $f_t$ : 有条件地决定块中保留的信息。推得的数学公式具体如下:

$$f_t = \sigma(\omega_f[h_{t-1}, x_t] + b_f) \quad (7)$$

其中,  $\omega_f$  和  $b_f$  分别表示权重矩阵和偏置向量。 $h_{t-1}$  是先前状态的输出;  $x_t$  表示输入值;  $\sigma(\cdot)$  表示 sigmoid 函数。

(2) 输入门  $i_t$ : 有条件地确定输入哪些值来更新记忆状态。推得的数学公式具体如下:

$$i_t = \sigma(\omega_i[h_{t-1}, x_t] + b_i) \quad (8)$$

(3) 输出门  $o_t$ : 根据输入和块记忆有条件地决定输出。推得的数学公式具体如下:

$$o_t = \sigma(\omega_o[h_{t-1}, x_t] + b_o) \quad (9)$$

在每个时间  $t$  内, 输入特征由输入  $x_t$  和之前隐藏状态  $h_{t-1}$  计算, 并且 tanh 函数取值范围在 -1 到 1 之间。由此推得的数学公式为:

$$\tilde{c}_t = \tanh(\omega_c[h_{t-1}, x_t] + b_c) \quad (10)$$

记忆单元由调节的输入特征  $\tilde{c}_t$  和前一个记忆单元  $c_{t-1}$  的遗忘部分更新。研究推得的公式为:

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \quad (11)$$

最终由输出门  $o_t$  和记忆单元  $c_t$  计算得到隐藏状态输出  $h_t$ 。研究用到的公式为:

$$h_t = o_t * \tanh(c_t) \quad (12)$$

在式(7)~(12)中, 矩阵  $\omega_f, \omega_i, \omega_o$  和  $\omega_c$  是权重矩阵; 向量  $b_f, b_i, b_o$  和  $b_c$  是偏置向量;  $h_{t-1}$  是前一个隐藏状态计算的结果;  $h_t$  是  $t$  时刻隐藏状态的输出;  $x_t$  表示输入信息。

Adam 是一种随机梯度下降的一阶优化算法, 在深度学习中有广泛的用途<sup>[20]</sup>。该方法的优势在于具有高效的计算能力, 对内存的需求不高, 实现过程中对需要调节的参数较少、且操作原理简单。Adam 优化算法的具体步骤如下所示。

**步骤 1** 初始化步长  $\alpha = 0.001$ , 初始化矩估计

指数衰减率  $\beta_1, \beta_2 \in [0, 1)$ , 这里取  $\beta_1 = 0.9$ ,  $\beta_2 = 0.99$ , 设定误差值  $\varepsilon = 10^{-8}$ ,  $f(\theta)$  是参数为  $\theta$  的随机目标函数。这里使用均方误差, Adam 的目标是找到一组使均方误差最小化的参数  $\theta$ 。

**步骤 2** 初始化第一个矩向量  $m_0 = 0$ , 第二个矩向量  $v_0 = 0$ , 初始时间  $t = 0$ 。

**步骤 3** 更新参数。相关系列公式见如下:

$$(1) t = t + 1$$

$$(2) \text{运算得到 } t \text{ 时刻的梯度随机目标: } g_t = \tilde{N}\theta_t f(\theta_{t-1})$$

$$(3) \text{更新有偏差的第一矩估计: } m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$$

$$(4) \text{更新有偏差的二阶原始矩估计: } v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$$

(5) 计算偏差校正的第一矩估计, 即:

$$mb_t = m_t / (1 - \beta_1^t)$$

(6) 计算偏差校正的二阶原始矩估计, 即:

$$vb_t = v_t / (1 - \beta_2^t)$$

$$(7) \theta_t = \theta_{t-1} - \alpha \cdot mb_t / (\sqrt{vb_t} + \varepsilon)$$

**步骤 4** 判断  $\theta_t$  是否收敛, 若不收敛, 继续执行步骤 3, 直到满足条件时循环结束。得到返回值  $\theta_t$ 。

如这里的 Adam 算法所述, 首先需要确定参数  $\alpha, \beta_1, \beta_2$  以及随机目标函数  $f(\theta)$ , 接下来初始化参数向量、第一个矩向量、第二个矩向量和时间。再进行多次循环迭代, 更新计算参数向量 ( $t, g_t, m_t, v_t, mb_t, vb_t$  和  $\theta_t$ ), 当参数  $\theta$  收敛时则认为循环停止。Adam 通过优化目标函数  $f(\theta)$  适用于 LSTM 神经网络, 找到一组使均方误差最小化的参数。

## 2 预测模型

### 2.1 电价预测过程

利用 VMD 分解和 Adam 优化 LSTM 的混合模型预测电价的过程大致分为以下 6 个步骤:

**步骤 1** 本文研究采用美国 PJM 电力市场 2012 年 1 月的节点边际电价数据进行实验分析, 来验证 VMD-Adam-LSTM 混合模型的有效性。

**步骤 2** 数据预处理。为了进一步提高模型对价格序列的预测精度, 根据  $3\sigma$  准则对原始数据训练集的异常值进行处理。通过计算一组电价序列的均值和标准差, 并根据一定的概率确定一个范围, 认为超出这个范围的数据属于异常值, 通过  $3\sigma$  准则剔除异常数据。具体计算流程如下:

$$(1) \text{计算样本电价数据的均值 } \mu \text{ 和方差 } \sigma^2。$$

(2) 根据  $3\sigma$  准则判断数据是否在  $(\mu - 3\sigma, \mu + 3\sigma)$  范围内。超出范围则剔除异常数据。

**步骤 3** 电价数据的 VMD 分解。由式(2)可知, VMD 方法的输入不仅是原始信号, 还需要确定参数  $K$  的具体值。 $K$  决定了分解过程中提取的模态数量。如果提取模态太多, 会导致精度降低和计算复杂性变大。但是, 如果模态数量太少, 模态中包含的信息不足以训练出具有较高精度的预测模型。因此, 如何为  $K$  选择一个合适值是非常重要的。本文通过中心频率确定 VMD 分解个数  $K$ , 并对分解后  $K$  个 IMF 和残差进行预测。

**步骤 4** Adam 优化的 LSTM 模型预测。经过上述 VMD 分解后, 利用 LSTM 神经网络对 VMD 分解后的时间序列进行训练和预测, 从而得到更准确的预测结果。该网络采用单输入/单输出的结构, 隐含层个数为 4, 使用默认的 *sigmoid* 激活函数。网络训练 1 000 个周期, 前一时刻步长作为输入变量来预测下一个时刻的数据。作为一种高效的随机梯度优化器, Adam 被用于 LSTM 的电价预测中。

**步骤 5** 计算原始电价预测结果。由式(2)的约束条件可知, 将  $K$  个 IMF 分别预测的结果和残差预测的结果相加, 即为原始电价预测结果。

**步骤 6** 预测模型评估。将预测结果与实际电价数据比较, 计算 3 种评估指标, 研究内容详见第 3 节。

根据上述预测步骤, 基于 VMD 分解和 Adam 优化 LSTM 神经网络的混合模型预测电价的过程如图 2 所示。

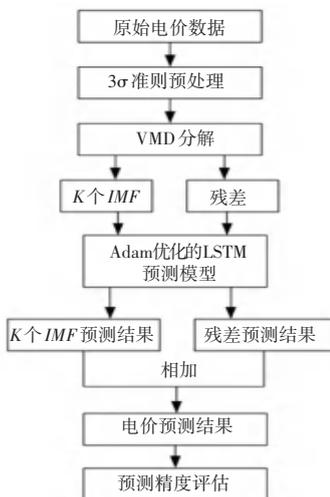


图 2 基于 VMD-Adam-LSTM 混合预测模型流程图

Fig. 2 Flow chart of VMD-Adam-LSTM hybrid prediction model

## 2.2 评估指标

为了评估预测模型的精度, 使用平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE) 和  $R^2$  的典型误差评估指标。对此可做探讨阐述如下。

(1) MAE。反映了绝对误差的平均值。MAE 越小, 预测越准确, 推得的数学公式可写为:

$$MAE = \frac{1}{m} \sum_{t=1}^m |x(t) - \tilde{x}(t)| \quad (13)$$

其中,  $x(t)$  为  $t$  时刻的原始电价数据;  $\tilde{x}(t)$  为  $t$  时刻的预测电价;  $m$  为预测点数。

(2) RMSE。描述了预测误差的离散程度。RMSE 越小, 预测越准确。推得的数学公式可写为:

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^m (x(t) - \tilde{x}(t))^2} \quad (14)$$

(3)  $R^2$ 。是拟合程度的统计, 反映了解释的程度。 $R^2$  结果越接近 1, 预测精度越高。推得的数学公式可写为:

$$R^2 = 1 - \frac{\sum_{t=1}^m (x(t) - \bar{x}(t))^2}{\sum_{t=1}^m (x(t) - \bar{x}(t))^2} \quad (15)$$

其中,  $\bar{x}(t)$  是  $x(t)$  的均值。

## 3 算例分析

### 3.1 数据介绍

为了验证本文提出的 VMD-Adam-LSTM 混合模型的有效性, 采用美国 PJM 电力市场 2012 年 1 月的节点边际电价数据进行实验仿真验证。在此数据采集过程中, 每隔 1 h 统计得到一个电价数据, 选取 1 月份前 30 天的数据, 共计 720 个数据作为实验数据。

### 3.2 VMD 分解和模态个数的确定

不同模态个数会直接影响预测结果, 因此在采用 VMD 算法对电价数据进行模态分解之前, 需要先确定分解的模态个数。如果分解的模态个数较少, 可能导致原始电价数据的信息丢失; 如果模态个数较多, 相邻 IMF 分量的中心频率较近还可能出现频率混叠, 也会导致精度降低和计算复杂性较高。根据中心频率不宜较近的方法选择合适的模态个数, 通过比较较为接近的 2 个中心频率之间的差值, 如果较小则不能选择此次分解的模态个数。将原始电价数据通过  $3\sigma$  准则对异常值处理后经 VMD 分解, 得到不同模态个数下的中心频率, 其计算结果见表 1。

表1 不同模态个数下VMD分解的中心频率

Tab. 1 Center frequency of VMD decomposition under different number of modes

模态个数		中心频率/Hz				
2	1 439	2 916	-	-	-	-
3	925	2 213	2 984	-	-	-
4	506	1 136	2 227	2 986	-	-
5	504	1 131	2 199	2 828	3 048	-

从表1中可以看出IMF个数为5时,中心频率2 828和3 048较为接近,可能造成频率混叠。因此选择分解的模态个数为4,具体分解结果如图3所示。

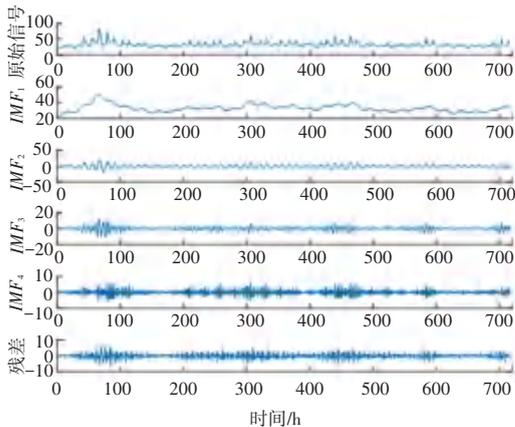


图3 电价数据的VMD分解结果

Fig. 3 VMD decomposition results of electricity price data

### 3.3 预测结果与评估

接下来,采用Adam优化的LSTM分别对各个IMF分量和残差进行预测。选取前29天数据,即0到696 h,作为训练集,最后一天697到720 h数据作为测试集。LSTM神经网络采用单输入/单输出的方式,隐含层个数为4,隐含层神经元个数为10,激励函数为sigmoid函数。将4个IMF分量预测结果和残差预测结果相加得到原始电价数据的预测结果。

为了验证本文提出的VMD-Adam-LSTM混合模型的有效性,分别与单一的LSTM预测模型和EMD-Adam-LSTM混合预测模型相比,其预测结果如图4所示。3种预测模型的预测误差结果见表2。

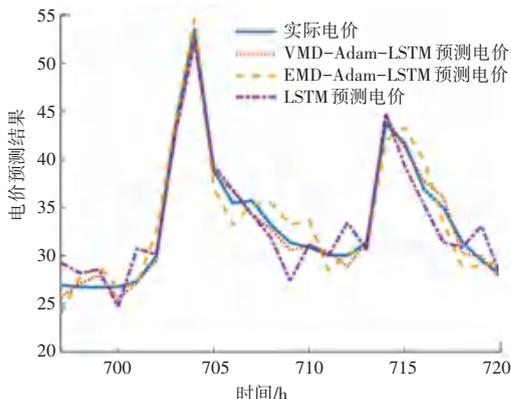


图4 不同模型的电价预测结果

Fig. 4 Electricity price prediction results of different models

表2 3种预测模型误差结果

Tab. 2 Error results of the three prediction models

预测模型	MAE	RMSE	R <sup>2</sup>
VMD-Adam-LSTM	0.750 8	0.858 4	0.983 3
EMD-Adam-LSTM	1.524 3	1.776 3	0.932 2
LSTM	1.661 2	2.041 1	0.890 7

从图4和表2中可以看出,LSTM预测模型对电价的拟合效果最低,VMD-Adam-LSTM混合模型相比于EMD-Adam-LSTM混合模型的拟合效果更高。究其原因则在于LSTM是单一的预测模型,没有对原始电价数据进行预处理,而VMD相对于EMD分解方法更能抑制模态混叠问题,使预测精度提高。

## 4 结束语

本文提出了一种VMD-Adam-LSTM混合模型,并采用美国PJM的电价数据进行实验验证。

该混合模型具有LSTM能够处理非线性和长依赖性时间序列的特点,同时也结合了VMD相对于EMD更能抑制模态混叠的优点。在VMD分解时,通过比较中心频率选择合适的IMF分量个数,防止过拟合现象的发生。在数据预处理时,采用 $3\sigma$ 准则避免了异常值的影响。实验结果表明,本文提出的VMD-Adam-LSTM混合模型相比于单一的LSTM模型和EMD-Adam-LSTM混合模型,具有更高的预测精度,验证了该模型的有效性和准确性。

## 参考文献

- [1] 宋坤, 武志锴, 南哲等. 基于多步电价预测的储能系统调度策略优化方法[J]. 沈阳工业大学学报, 2021, 43(05):493-499.
- [2] 王俊秀, 戴泽宇. 基于多维精益管理的电价预测与电费风险控制分析研究[J]. 电工技术, 2021, 16:135-136.
- [3] HE Kaijian, YU L, TANG Ling. Electricity price forecasting with a BED (Bivariate EMD Denoising) methodology [J]. Energy, 2015, 91:601-609.
- [4] ELAMIN N, FUKUSHIGE M. Modeling and forecasting hourly electricity demand by SARIMAX with interactions [J]. Energy, 2018, 165:257-268.
- [5] 赵雅雪, 王旭, 蒋传文等. 基于最大信息系数相关性分析和改进多层级门控LSTM的短期电价预测方法[J]. 中国电机工程学报, 2021, 41(01):135-146.
- [6] SINGH N, MOHANTY S R, SHUKLA R D. Short term electricity price forecast based on environmentally adapted generalized neuron [J]. Energy, 2017, 125:127-139.
- [7] LAGO J, De RIDDER F, VRANCX P, et al. Forecasting day-ahead electricity prices in Europe: The importance of considering market integration [J]. Applied Energy, 2018, 211:890-903.
- [8] ZHOU Ming, YAN Zheng, NI Yixin, et al. A novel arima approach on electricity price forecasting with the improvement of predicted error [J]. Proceedings of the CSEE, 2004, 24(12):63-68.