

# 基于 EEMD-LSTM 的区域能源短期负荷预测

马梦冬, 彭道刚, 王丹豪

(上海电力大学 自动化工程学院, 上海 200090)

**摘要:** 短期负荷预测既是电网规划的重要组成部分, 也是系统可靠、高效运行的前提和基础。采用 EEMD(集合经验模态分解)方法将区域能源互联网历史负荷数据分解成若干分量, 再对各个分量分别建立模型, 运用 LSTM(长短期记忆神经网络)设置对应的隐藏层数, 对各个分量分别进行预测, 最后叠加预测值得出最终预测结果。将 EEMD-LSTM 算法与 LSTM 算法、EMD-LSTM 算法以及 Elman 算法进行比较, 结果表明 EEMD-LSTM 算法在区域能源互联网负荷预测中能够实现较好的预测精度, 具有很好的应用前景和推广价值。

**关键词:** 短期负荷预测; 区域能源; 集合经验模态分解; 长短期记忆神经网络

文章编号: 1007-1881(2020)04-0029-07

DOI: 10.19585/j.zjdl.202004005

中图分类号: TM615

文献标志码: A

开放科学(资源服务)标识码(OSID):



## Short-term Load Forecasting Based on EEMD-LSTM for Regional Energy

MA Mengdong, PENG Daogang, WANG Danhao

(College of Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China)

**Abstract:** Short-term load forecasting is not only an important part of power grid planning but a prerequisite and basis for the reliable and efficient operation of the system. The ensemble empirical mode decomposition method (EEMD) is used to decompose historical load data of the regional energy Internet into several components, and then models for the components built separately. The long-short-term memory (LSTM) is used to set the corresponding hidden layer number to predict each component separately and conclude the final prediction result by superimposing the prediction values. Compared with LSTM algorithm, EMD-LSTM algorithm and Elman algorithm, it is verified that EEMD-LSTM can achieve better prediction accuracy in regional energy Internet load forecasting and has a good application prospect and is worthy of promotion.

**Keywords:** short-term load forecasting; regional energy; EEMD; LSTM

## 0 引言

随着全球能源需求的不断增长, 化石能源日益短缺<sup>[1]</sup>。为缓解能源危机, 减小环境污染带来的影响, 新能源发电越来越受关注。2011年, 美国学者杰里米·里夫金(Jeremy Rifkin)在其著作《第三次工业革命》中, 以新能源技术和信息技术的深入结合为特征, 首次提出能源互联网的概念<sup>[2]</sup>。能源互联网包括广域能源互联网和区域能源互联网, 广域能源互联网一般承担远距离的大型输电,

连接两地大型能源基地; 区域能源互联网作为能源互联网的重要表现形式, 主要进行近距离输配电, 多用于经济开发区、工业园区、电动汽车充放电等<sup>[3]</sup>。能源互联网主要由天然气、风电、光伏、储能提供电热, 因此推进了多种能源形式协同运行, 通过燃气内燃机、燃气轮机、风光发电以及生物质能发电, 实现了能源的梯级利用和优势互补, 符合“融合”“统一”“高效”“清洁”的理念<sup>[4]</sup>。能源互联网以其强大的耦合性为可再生能源产业发展提供保障, 同时也成为智能电网的重要组成部分。

针对电网的可靠运行问题, 很多学者进行了相关研究。文献[5]介绍了神经网络的特点, 说明了能够有效处理电力短期负荷预测数据的非线性

基金项目: 国家自然科学基金项目(71871160); 上海市“科技创新行动计划”高新技术领域项目(19511101600)

变化问题。文献[6]能够达到较高的精度,但其优化寻优算法考虑的因素较多,模型复杂,训练时间长,给实际应用带来困难。文献[7-8]采用人工神经网络对负荷参数模型进行预测,但是精度不够理想。文献[9]采用改进的遗传算法搜索全局最优值,以优化极限学习机算法等建立负荷预测的模型,预测结果基本可以满足精度要求,但预测模型考虑了较多因素,模型构建复杂。文献[10]对配电网进行负荷预测,并对不同层数的神经网络模型进行了对比,结果显示对于差异性较大的模型尚不能达到最优效果。文献[11-12]阐述深度学习在电力负荷预测中的应用,列举了常用的深度学习算法,对比分析之后验证出 LSTM(长短期记忆神经网络)算法具有较高的精度,能够较好地预测负荷的短期功率,因此本文算例分析中采用 LSTM 算法进行建模。文献[13]提出了一种改进的最小二乘向量机负荷预测模型,但模型预测结果依然存在较大误差,无法保证整体相对平稳的预测精度。

与大电网相比,区域能源互联网负荷随机性很强,历史负荷曲线相似度低,基本没有规律可循;并且用户容量有限,各用户间负荷特征不够明显。本文算例分析中采用 EEMD(集合经验模态分解)方法对原始数据进行分解,能够有效避免模态混叠现象<sup>[14]</sup>;同时基于 LSTM 较精准的预测效果,利用 LSTM 对 EEMD 方法得到的每个分量分别进行预测,以获得更好的预测结果。

## 1 EEMD 方法

EEMD 由 EMD(经验模态分解)加入白噪声后改进得出。对于非平稳数据,不需要考虑其他因素影响,可以使用 EMD 方法得到一组能够表现信号特征的数据。与小波变换等方法相比,这种方法是直观的和自适应的。因为分解是基于信号序列时间尺度的局部特征,而 EEMD 作为一种白噪声辅助分析方法,在原始信号中加入多组随机的白噪声信号,能够更有效地降低原始序列的波动性,其分解过程如图 1 所示。利用 EEMD 方法可以将能源互联网负荷序列分解成数个相对平稳的分量。相比于小波分析,EEMD 不需要对信号进行预先分析和研究,而是直接进行分解。

在本次模型建立中,将能源互联网原始负荷

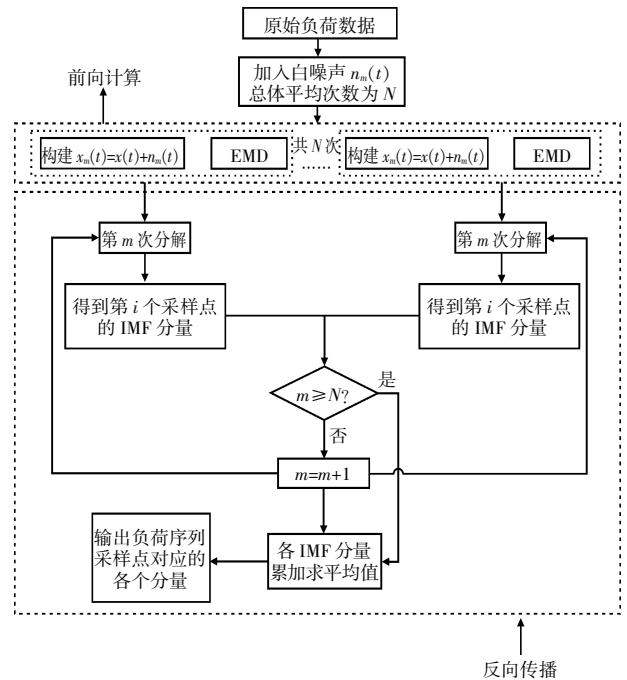


图 1 EEMD 过程

序列  $x(t)$  中加入随机白噪声序列  $n_m(t)$ , 得到加入噪声信号后的待处理序列:

$$x_m(t) = x(t) + n_m(t), m = 1, 2, \dots, N, \quad (1)$$

式中:  $N$  为 EMD 总体平均次数。

运用 EMD 方法加入 100 组有差别的随机白噪声序列, 标准差取 0.2。将区域能源互联网原始负荷序列分解一次之后再加入均方根相等的白噪声序列, 经过  $N$  次 EMD, 使信号重构达到较好的效果。通过此次分解, 原始数据 68 个典型日的采样点共得到了 7 个 IMF 分量和 1 个剩余分量, 也就是区域能源互联网原始负荷序列的 EEMD 结果。EEMD 能够克服基函数无自适应的问题, 把时间序列分解成若干个 IMF 分量和残余项  $R(t)$ , 即某个时刻采样点的数值等于对应的各个 IMF 分量与残余项相加。

## 2 LSTM 算法

为了解决 RNN(循环神经网络)在处理长期依赖性时产生的梯度消失问题, 提出了在 RNN 基础上的改进算法 LSTM, 该算法在每个隐藏层中引入栅极和存储单元的概念。一个 LSTM 内存块主要由输入门、遗忘门、输出门和自连接存储单元 4 个部分组成。输入门控制对存储单元的激活输入, 能够学习如何过滤和输出到连续网络的激

活单元。遗忘门帮助网络忘记过去的输入数据并重新设置记忆细胞, 由 sigmoid 单元将权重设置为 0 和 1 之间的值: 当权重为 0 时, 表示将上一状态的信息全部舍弃; 当权重为 1 时, 表示将上一状态的信息全部保留。此外, 应用乘法门使存储单元可以长时间访问和存储信息, 每个单元中都包含存储前一时刻信息的记忆细胞。因此, LSTM 作为一种深度神经网络, 具有更多的隐藏层, 增加了计算的精度优势。LSTM 的典型结构如图 2 所示, 单个 LSTM 神经元的预测结构如图 3 所示。LSTM 记忆细胞的运算过程为:

$$\text{遗忘门: } f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

$$\text{输入门: } i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

候选时刻记忆单元状态:

$$c_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (4)$$

当前时刻记忆单元状态:

$$\tilde{c}_t = f_t c_{t-1} + i_t c_t \quad (5)$$

$$\text{输出门: } o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

$$\text{输出值: } h_t = o_t \tanh(\tilde{c}_t) \quad (7)$$

式中:  $W_f, W_i, W_c, W_o$  分别为遗忘门、输入门、细胞状态、输出门的权重矩阵;  $b_f, b_i, b_c, b_o$  分别为对应的偏置矢量;  $x_t$  为输入值;  $\sigma$  为 sigmoid 函数; 下标  $t$  表示当前时刻,  $t-1$  表示前一时刻。

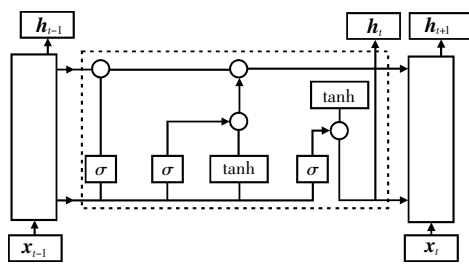


图 2 LSTM 典型结构

从图 3 中可以看出, LSTM 包括各个向量的交互, 使得结构相比于单一的神经网络具有更多的层级。图中“+”表示矩阵相加, “x”表示矩阵相乘。在进行负荷预测时, 结合当前时刻的负荷输入值与前一时刻的负荷输出值赋予不同的权重矩阵和偏置矢量, 使用 tanh 函数和 sigmoid 单元进行细胞状态的更新。本次试验设置隐藏层数为 2, 采用前 48 个时刻预测后 24 个时刻, 训练集和测试集设置比例为 7:3。首先利用 EEMD 方法对原始负荷数据进行分解, 得出一组具有不同特

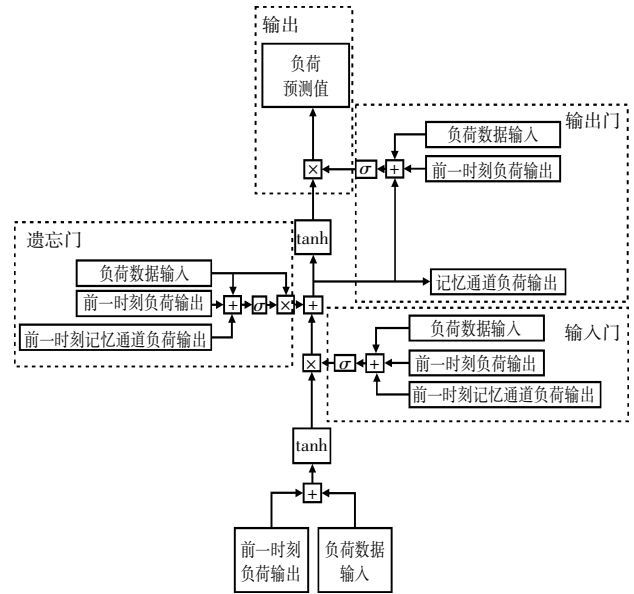


图 3 单个 LSTM 神经元的预测结构

征的分量; 然后将分解后的各个分量分别送入 LSTM 进行预测; 最后将各个分量值累加, 得到预测结果。

### 3 基于 EEMD-LSTM 的预测模型建立

本文提出基于 EEMD-LSTM 的区域能源互联网系统日前小时负荷预测模型, 对应流程如图 4 所示。对于历史区域能源互联网小时出力序列, 首先采用 EEMD 方法分解得到一组分量, 每个分量数据归一化到 [0, 1] 范围内, 然后建立对应的 LSTM 模型进行预测, 将结果进行反归一化, 并叠加各个分量即日前区域能源互联网系统小时负荷预测值, 最后将输出的预测结果与原始负荷数据进行对比, 并进行误差分析。

为了更直观地对比预测结果, 本次算例选用预测结果误差分析常用的 APE(相对百分比误差, 量符号为  $E_{AP}$ )进行作图分析, 对比各个算法的误差结果。使用 MAPE(平均绝对百分比误差, 量符号为  $E_{MAP}$ )、MSE(均方误差, 量符号为  $E_{MS}$ )和 RMSE(均方根误差, 量符号为  $E_{RMS}$ )分别对预测结果进行评估, 各个误差的表达式分别为:

$$E_{AP} = \frac{Y_i - \hat{Y}_i}{Y_i} \quad (8)$$

$$E_{MAP} = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{Y_i} \quad (9)$$

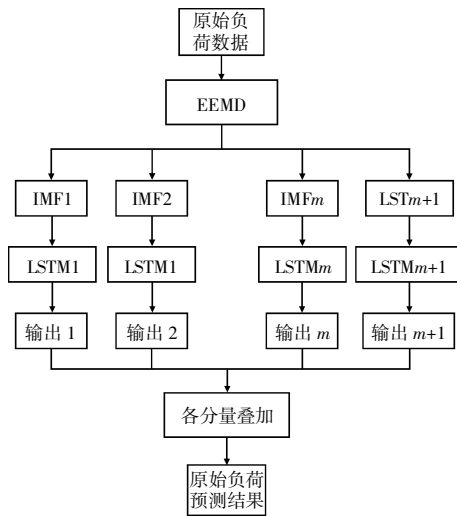


图4 EEMD-LSTM 流程

$$E_{\text{MS}} = \frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2, \quad (10)$$

$$E_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2}, \quad (11)$$

式中： $n$  为数据总个数； $Y_i$  为第  $i$  个实际值； $y_i$  为第  $i$  个预测值。

#### 4 算例分析

本文以某区域能源互联网系统各单元总负荷数据为研究对象，选取 2017 年 1 月 1 日 00:00 至 2017 年 3 月 9 日 24:00，共计 68 个典型日，采样时间间隔为 1 h，共计 1 632 个点。LSTM 的迭代次数设置为 100 次，隐藏层数为 2，训练集和测试集设置比例为 7:3。寻找最优的预测模型之后，把预测样本作为输入值代入最优模型中并得到 2017 年 3 月 9 日的预测值。采用 EEMD 方法对原始负荷序列进行分解，结果如图 5 所示。

通过图 5 各个分量图可以更直观地看出，所有 IMF 分量均反映出频率从高到低的不同尺度的系统负荷序列波动特性。IMF1 的平均振幅小，在所有分量中频率最高，并且规律性不明显，可以判断为负荷序列中的随机分量；IMF2—IMF4 具有较好的规律性，周期性较为明显，而且曲线的幅值随着时间的推移有所增长，可以反映出负荷序列的局部波动情况，因此这些 IMF 分量可判定为负荷序列的细节分量；其余的 IMF 分量和余

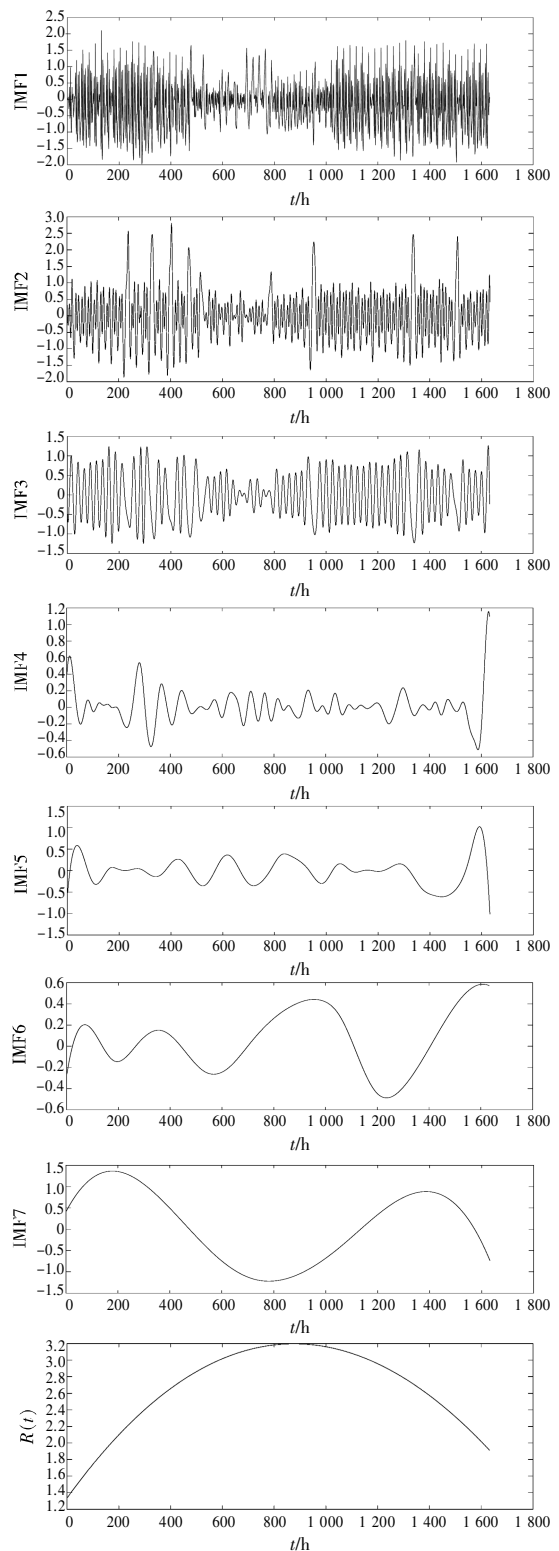


图5 原始负荷数据的 EEMD 结果

量  $R(t)$  的平均振幅大，具有很好的周期性和规律性，尤其是余项反映了负荷曲线的全局变化趋势，因此 IMF5—IMF7 以及余项  $R(t)$  为负荷序列

的趋势分量。

为了验证本文预测模型的可行性, 选取了另外3种模型与其进行对比, 其中一个模型采用单一的 LSTM, 另一个模型采用 EMD-LSTM 算法, 最后一个为其他的单一神经网络模型——Elman 神经网络。选取 2017-01-01 T 00:00—03-09 T 24:00 期间, 时间间隔 1 h 的采样点, 共计 1 632 个点。对后 24 个采样点进行预测, 即对 3 月 9 日 00:00—24:00 采样点数据进行对比分析。预测结果如图 6 所示, 预测相对误差如图 7 所示, 预测误差数据对比见表 1。

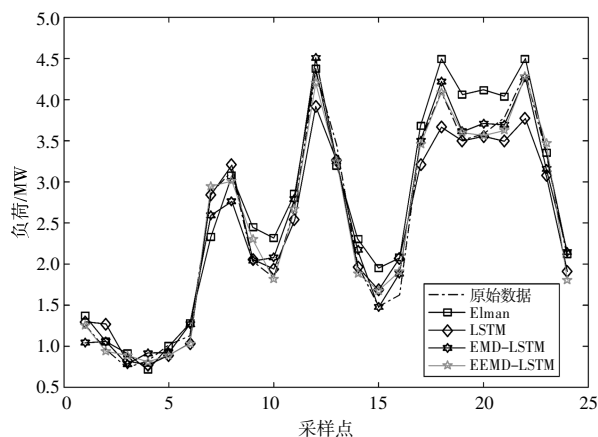


图 6 系统 3 月 9 日负荷预测结果(采样间隔 1 h)

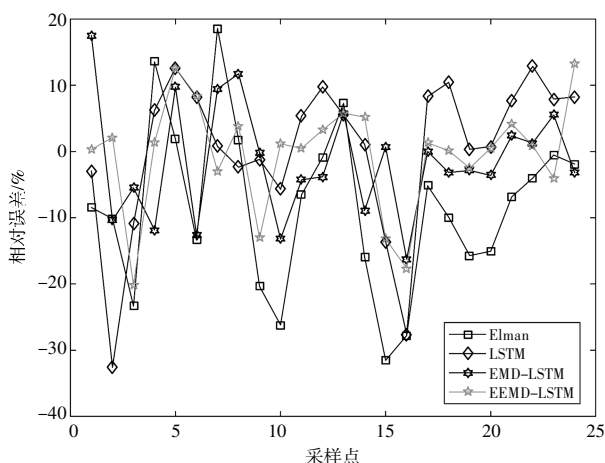


图 7 系统 3 月 9 日负荷预测相对误差对比(采样间隔 1 h)

由图 7 可以看出, EEMD-LSTM 算法的相对误差波动性较小。由表 1 可以看出, 采用 EEMD-LSTM 模型的预测精度最高, 其 MAPE 和 MSE 比

表 1 系统负荷预测误差数据对比(采样间隔 1 h)

预测模型	MAPE/%	MSE	RMSE
Elman	11.95	0.095 2	0.308 6
LSTM	8.45	0.057 8	0.240 4
EMD-LSTM	6.83	0.026 7	0.163 5
EEMD-LSTM	5.76	0.019 3	0.138 9

其他模型的值都要小, 验证了 LSTM 算法的可行性, 同时证明了 EEMD 方法能降低原始序列的非平稳性对预测精度的影响。

由表 1 可以看出, 单一神经网络进行比较时, LSTM 的预测精度比 Elman 要高, 加入白噪声运用 EEMD 方法进行原始负荷数据分解后, 能够克服 EMD 方法分解的不足, 预测精度有明显提高。从整体来看, 基于 EEMD-LSTM 算法的预测结果很少出现相对误差极大的点。而另外两种模型的预测结果中, 往往会出现较多相对误差极大的“失真点”。采用 MAPE, MSE 和 RMSE 作为评价指标, 算例分析得出的值越小, 说明模型的预测精度越高。

为了验证本文提出算法的有效性和适用性, 使用间隔 15 min 采样点进行预测对比。现仍采用 2017-01-01 T 00:00—03-09 T 24:00 期间, 时间间隔为 15 min 的采样点, 共计 6 528 个点。对后 24 个采样点进行预测, 即对 3 月 9 日 19:00—24:00 采样点数据进行对比分析。预测结果如图 8 所示, 为便于与前例对比分析, 误差预测结果如图 9 所示, 预测误差数据对比见表 2。

由表 1 和表 2 对比分析可以得出, 数据量由 1 632 个采样点增大到 6 528 个采样点时, EEMD-LSTM 构造的预测模型的精度最高。同时, 验证了在单一算法中 LSTM 相对于 Elman 的优越性。在组合预测中, EEMD-LSTM 能够比 EMD-LSTM 取得更好的效果, 算法对比表明了 EEMD-LSTM 的优越性。

使用同时段的数据, 采样时间间隔 15 min, 共计 6 528 个采样点, 预测 3 月 9 日全天的负荷预测值, 结果如图 10 所示。训练集与测试集比例不变, 将迭代次数增加至 150 次, EEMD-LSTM 模型的 MAPE 维持在 3.5%左右, 进一步验证了数据量增大时 EEMD-LSTM 算法的优越性。由于预测点增多, 需要扩大预测的维度及增加迭代次数。当需要预测的点增加时, 精度会有所下降,

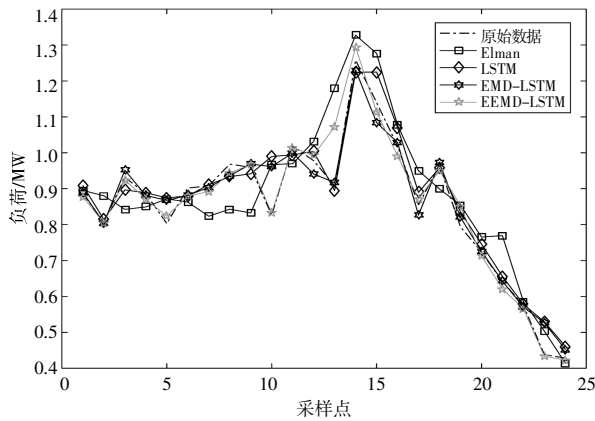


图8 系统3月9日19:00—24:00负荷预测结果  
(采样间隔15 min)

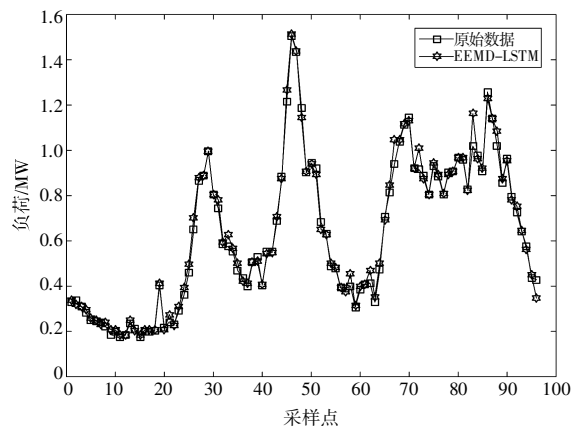


图10 系统3月9日全天负荷预测结果  
(采样间隔15 min)

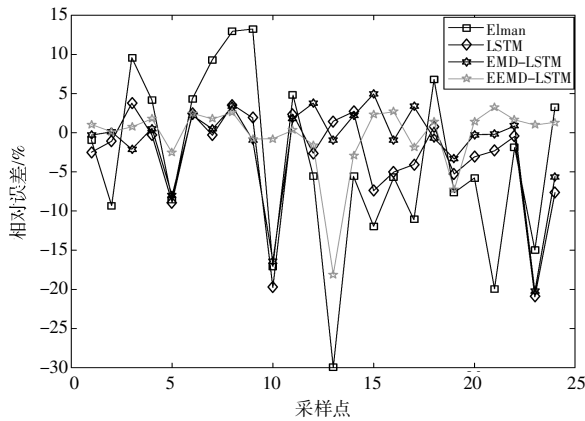


图9 系统3月9日19:00—24:00负荷预测相对误差对比(采样间隔15 min)

表2 预测误差数据对比(采样间隔15 min)

预测模型	MAPE/%	MSE	RMSE
Elman	9.35	0.009 4	0.097 1
LSTM	4.61	0.002 5	0.050 5
EMD-LSTM	3.49	0.001 7	0.041 0
EEMD-LSTM	2.57	0.001 5	0.039 1

计算时间也会增长。但相比于原始数据采样点较少时的预测结果来说,当数据量增大时,EEMD-LSTM的预测结果能够达到更精准的效果。

### 5 结语

针对区域能源互联网发电短期负荷功率预测的准确性问题,本文提出了一种基于EEMD-LSTM算法的区域能源互联网系统小时负荷预测模型。EEMD-LSTM模型能够有效克服原始信号

中的不确定因素,并且利用LSTM存储单元长时间学习和保留电力负荷历史数据中的有用信息,并使用遗忘门将无用信息去除,然后将EEMD得出的每一个分量送入LSTM算法中进行预测,对结果进行叠加处理后得出最终的预测结果。

为便于对比分析,本文使用传统的Elman算法、单一的LSTM算法以及EMD-LSTM构建模型,使用相同的数据采样点建模,对比各个模型预测结果曲线的拟合情况,并对结果进行了误差分析。针对区域能源互联网系统短期负荷序列的无规律波动特性,与EMD相比,EEMD能更有效地降低对预测结果的影响,EEMD-LSTM能够更准确地对负荷进行预测。当需要预测的数据增多时,预测模型的运算时间会稍有延长。当原始数据量增大而需要预测的采样点不增加时,EEMD-LSTM体现出了精度更高的优越性。因此,EEMD-LSTM基本能够满足区域能源互联网短期负荷预测的要求,为能源互联网电力系统的平稳运行提供一定参考。

### 参考文献:

- [1] 董朝阳,赵俊华,文福拴,等.从智能电网到能源互联网:基本概念与研究框架[J].电力系统自动化,2014,38(15):1-11.
- [2] RIFKIN J.The third industrial revolution:how lateral power is transforming energy, the economy, and the world[M]. New York:Palgrave MacMillan,2011:24-71.
- [3] 郭创新,王惠如,张伊宁,等.面向区域能源互联网的“源-网-荷”协同规划综述[J].电网技术,2019,43(9):3071-3080.

- [4] 卢毓东,魏宏,王芙丽,等.能源互联网背景下的客户侧分布式电源综合运营服务平台及其关键技术研究[J].浙江电力,2017,36(10):77-82.
- [5] 张亚军,张大波,许诚昕.神经网络在电力系统短期负荷预测中的应用综述[J].浙江电力,2007,26(2):5-9.
- [6] 吴润泽,包正睿,宋雪莹,等.基于深度学习的电网短期负荷预测方法研究[J].现代电力,2018,35(2):43-48.
- [7] 李龙,魏靖,黎灿兵,等.基于人工神经网络的负荷模型预测[J].电工技术学报,2015,30(8):225-230.
- [8] 程宇也.基于人工神经网络的短期电力负荷预测研究[D].杭州:浙江大学,2017.
- [9] 律方成,刘怡,亓彦珣,等.基于改进遗传算法优化极限学习机的短期电力负荷预测[J].华北电力大学学报(自然科学版),2018,45(6):1-7.
- [10] 黄宇腾,韩翊,赖尚栋.深度神经网络在配电网公变短期负荷预测中的应用研究[J].浙江电力,2018,37(5):1-6.
- [11] 张建寰,吉莹,陈立东.深度学习在电力负荷预测中的应用[J].自动化仪表,2019,40(8):8-12.
- [12] MUZAFFAR S,AFSHARI A.Short-term load forecasts using LSTM networks[J].Energy Procedia,2019,158:2922-2927.
- [13] 张宇帆,艾芊,林琳,等.基于深度长短时记忆网络的区域级超短期负荷预测方法[J].电网技术,2019,43(6):1884-1892.
- [14] 王新,孟玲玲.基于EEMD-LSSVM的超短期负荷预测[J].电力系统保护与控制,2015,43(1):61-66.
- [15] SHRIVASTAVA Y,SINGH B.A comparative study of EMD and EEMD approaches for identifying chatter frequency in CNC turning[J].European Journal of Mechanics-A/Solids,2019,73:381-393.
- [16] 王龙,成天乐,陈宇.基于EMD和ELM的短期负荷预测[J].电力科学与工程,2014,30(6):54-58.
- [17] 钱忆钊,陈良.Elman神经网络在电力负荷预测中的应用[J].电工技术,2019(14):55-56.
- [18] 叶琳,杨滢,洪道鉴,等.深度学习在电力系统中的应用研究综述[J].浙江电力,2019,38(5):83-89.
- [19] 郭威麟,蒋晓艳,罗意,等.基于EEMD-GABP的某地区短期负荷预测研究[J].电力工程技术,2019,38(6):93-98.
- [20] 吴茵,莫东,李秋文,等.基于分类波动性统计的短期负荷改进区间预测方法[J].电力大数据,2019,22(11):14-20.
- [21] 应张驰,陈淑萍,卢旭航.基于多源信息的短期负荷混合预测模型应用研究[J].浙江电力,2019,38(9):100-104.
- [22] 王帅,赵建平,王志远,等.基于云计算平台实现电网短期负荷预测算法的研究[J].四川电力技术,2019,42(1):29-32.
- [23] 刘伟,张锐锋,彭道刚.基于K-Adaboost数据挖掘的配电网负荷预测[J].浙江电力,2019,38(1):104-110.
- [24] 李重春,祝安琪,王烁罡,等.电力大数据下的短期电力负荷预测[J].电力大数据,2019,22(1):66-70.
- [25] 范金骥.基于ARMA与ANN模型组合交叉方法的电网日负荷预测[J].浙江电力,2018,37(8):35-41.

收稿日期:2020-01-09

作者简介:马梦冬(1994),女,在读硕士研究生,主要研究方向为能源互联网。

彭道刚(1977),男,工学博士,教授,研究方向为智能发电自动化、新能源微电网与能源互联网等。(通讯作者)

(本文编辑:童凯)