

基于 3σ -CEEMDAN-LSTM的空间负荷预测方法

肖白¹,高文瑞¹,李道明¹,慕雪松²,阚中锋²

(1. 东北电力大学 电气工程学院,吉林 吉林 132012;2. 国网吉林省电力有限公司吉林供电公司,吉林 吉林 132001)

摘要:为有效利用从配电网采集的海量数据以及改善空间负荷预测效果,提出一种基于 3σ 准则、自适应噪声完备集合经验模态分解(CEEMDAN)和长短期记忆神经网络(LSTM)的空间负荷预测方法。基于 3σ 准则对每个I类元胞的实测负荷数据进行奇异值检测和处理;运用CEEMDAN技术将处理后的I类元胞负荷数据分解为若干个频率和幅值均不同的本征模态函数(IMF);分别对每个IMF分量构建LSTM模型进行预测;将所有IMF分量预测结果进行线性叠加,得到目标年基于I类元胞的空间负荷预测结果,在此基础上使用空间电力负荷网格化技术求得基于II类元胞的空间负荷预测结果。算例分析结果验证了所提方法的正确性和有效性。

关键词:空间负荷预测;元胞;地理信息系统;自适应噪声完备集合经验模态分解;长短期记忆神经网络

中图分类号:TM73

文献标志码:A

DOI:10.16081/j.epae.202208016

0 引言

空间电力负荷预测是预测区域内未来电力负荷大小及其分布的预测,又称为空间负荷预测(spatial load forecasting, SLF)^[1-2]。SLF是城市电网规划的基础,其结果将直接影响电力设备配置的合理性与科学性。

按照预测原理可将现有的SLF方法分为趋势类SLF方法^[3]、多元变量法^[4]、用地仿真类SLF方法^[5]和负荷密度指标类SLF方法^[6-7]四大类。

趋势类SLF方法^[3]是通过外推元胞的历史负荷曲线得到目标年中元胞的空间负荷值。该类方法操作简便,对数据的需求量较小,常用于短期预测,但元胞负荷的非平稳增长、负荷转移和空白小区都会给预测精度带来不利的影响。

多元变量法^[4]是计量经济模型预测法在SLF领域中的应用。该类方法将历史年份元胞的峰值负荷与影响其发展的相关变量相结合,通过建立外推模型预测元胞目标年的峰值负荷。该类方法由于对数据质和量的要求均较高,在20世纪80年代末期已基本被淘汰。

用地仿真类SLF方法^[5]是通过分析规划区域土地利用变化的情况预测规划区域未来的土地利用类型,并在此基础上将土地利用情况转化成空间负荷^[8]。然而,该类方法在将总量负荷预测值根据用

地评分情况进行空间分配的过程中,不但没有计及负荷同时率,而且所需的分类负荷密度也很难精确确定,从而导致出现幅值预测结果偏低的系统性误差以及较大的空间误差。

负荷密度指标类SLF方法^[6-7]是根据元胞内不同功能小区的占地面积及相应的负荷密度来求取元胞的负荷值。该类方法适用于土地使用性质较为明确的场景,对城市规划方案的变化具有较强的适应性。我国政府对未来城市土地的开发和利用通常都提前进行了科学合理的规划,这为负荷密度指标类SLF方法的实现创造了有利条件。同时,智能电表、智能电网调度管理系统等收集到的海量历史负荷数据也为人工智能算法的应用提供了数据基础^[9]。近年来,随着人工智能技术的不断发展,已有许多模型应用于负荷密度指标类SLF方法^[10-14]。文献[10]充分利用元胞负荷和元胞属性间的相关关系,结合支持向量机在有限非线性样本拟合方面的优势来实现SLF。文献[11]利用模糊信息化技术对数据进行预处理,并在此基础上采用支持向量机得到负荷的空间分布。文献[12]从分析元胞负荷特性入手,构建一种最大负荷数据群的双向选取模型,并通过径向基函数(radial basis function, RBF)神经网络完成SLF。文献[13]通过互补集合经验模态分解和游程检验技术确定元胞合理最大值,并结合最小二乘支持向量机进行SLF。文献[14]先利用生成式对抗网络对数据匮乏的小区进行数据增强,再通过循环卷积生成式对抗网络对增强后的样本进行无监督学习,从而实现SLF。然而,文献[10-14]中的SLF方法在进行预测的过程中均未考虑负荷序列的随机波动特性给建模带来的不利影响,且这些方法所使用的人工智能模型均具有一定的局限性,其中:支持向量机的性能主要取决于选取的核函数,而目前核函数

收稿日期:2021-12-19;修回日期:2022-04-10

在线出版日期:2022-08-09

基金项目:国家自然科学基金资助项目(51177009);国家重点研发计划项目(2017YFB0902205)

Project supported by the National Natural Science Foundation of China(51177009) and the National Key R&D Program of China(2017YFB0902205)

及其参数均是根据人为经验得到的,具有一定的主观性;RBF神经网络在处理数据量较大的样本时,其模型结构往往过于复杂;生成式对抗网络在训练过程中常出现模式崩溃、梯度消失等问题。

针对上述问题,本文提出一种基于 3σ 准则(又称为拉伊达准则)、自适应噪声完备集合经验模态分解(complementary ensemble empirical mode decomposition with adaptive noise, CEEMDAN)和长短期记忆神经网络(long-short term memory neural network, LSTM)的SLF方法(简称基于 3σ -CEEMDAN-LSTM的SLF方法),并运用该方法对我国东北某城市的一个行政区域进行SLF来证明其正确性与有效性。

1 3σ -CEEMDAN-LSTM 预测方法基本原理

为了便于后续说明和理解,这里先给出元胞的定义^[15]:在电力地理信息系统(geographic information system, GIS)中,将待预测区域按照一定的原则划分为不规则或规则的供电小区,则每个供电小区称为一个负荷元胞,简称元胞。其中:将待预测区域按照电力设备或设施、供电单位、行政区域的供电范围划分得到的供电小区称为 I 类元胞;将待预测区域按照等大小的规则网格划分得到的供电小区称为 II 类元胞。

实现 3σ -CEEMDAN-LSTM预测方法的步骤如下。

1)对待规划区域内的历史负荷数据、用地信息、网架结构以及10 kV馈线供电范围进行整合。

2)结合收集到的基础数据信息,在电力GIS中建立待预测区域的配准基图。以10 kV馈线供电范围生成 I 类元胞,并建立 I 类元胞图层;选择适当的电力负荷空间分辨率^[15],以等大小的正方形网格生成 II 类元胞,并建立 II 类元胞图层。

3)采用 3σ 准则对所有 I 类元胞的实测历史负荷时间序列进行奇异检测,并对检测出的奇异数据进行修正。分别对每个修正后的 I 类元胞负荷时间序列提取各日负荷的最大值,形成各个日合理最大值时间序列。

4)设定合理的参数,利用CEEMDAN算法分别将各 I 类元胞历史负荷的日合理最大值时间序列分解为多个本征模态函数(intrinsic mode function, IMF)分量。

5)针对分解得到的每个IMF分量,根据其数据特征分别构建合适的LSTM预测模型,将各 I 类元胞的所有IMF分量预测结果进行线性叠加,即将各 I 类元胞所有IMF分量在同一时间点的预测结果直接进行相加,即可得到目标年基于 I 类元胞的SLF结果。

6)采用文献[16]中的空间电力负荷网格化技

术,将基于 I 类元胞的SLF结果转化为基于 II 类元胞的SLF结果。

本文所提基于 3σ -CEEMDAN-LSTM的SLF方法的基本原理图如附录A图A1所示。需要指出的是, 3σ -CEEMDAN-LSTM预测方法适用于用地信息比较明确且历史负荷数据充足的场景,预测结果主要用于一年期的城市电网规划,为电力规划部门在每年对原定的“五年规划”进行调整和修订时提供数据支持。

2 基于 3σ 准则的 I 类元胞日负荷合理最大值的确定

SLF中所使用数据的质量会直接影响预测结果的准确性。一方面, I 类元胞的负荷时间序列多毛刺,基荷小且波动大,而且在数据采集过程中,测量、转换和传输环节均可能使采集到的数据出现奇异数据,如果直接将 I 类元胞的实测值用于SLF,则可能会因为奇异数据的存在而导致预测精度降低。另一方面,面向规划的SLF往往采用最大值进行预测:若使用年最大值进行预测,则可能会因为数据量较少而导致模型出现欠拟合;若使用每小时的最大值进行预测,则会大幅增加模型的训练时间。为此,本文利用 3σ 准则依次对各 I 类元胞历史负荷的实测数据进行奇异值检测和修正,并确定每个 I 类元胞各日负荷的合理最大值。

3σ 准则常用于检测样本中的奇异数据。将包含 s 个数据的样本记为 $[X_1, X_2, \dots, X_s]$ (X_p ($p=1, 2, \dots, s$)为样本中的第 p 个数据),具体检测过程如下。

1)计算整个样本的算数平均值 \bar{X} ,并计算样本中每个数据对应的残余误差,如式(1)所示。

$$V_p = X_p - \bar{X} \quad (1)$$

式中: V_p ($p=1, 2, \dots, s$)为样本中第 p 个数据对应的残余误差。

2)计算整个样本的标准偏差 σ ,如式(2)所示。

$$\sigma = \sqrt{\frac{1}{s} \sum_{p=1}^s (\bar{X} - X_p)^2} \quad (2)$$

3)将每个数据的残余误差取绝对值后依次与 3σ 进行比较,若满足式(3),则认为 X_p 是奇异数据。

$$|V_p| > 3\sigma \quad (3)$$

3 基于CEEMDAN的 I 类元胞负荷时序分解

I 类元胞负荷时间序列具有一定的随机波动性,该特性往往会使建立的预测模型难以实现负荷的准确预测。为降低随机波动性给预测结果带来的不利影响以及提升模型的预测性能,本文采用CEEMDAN算法将各 I 类元胞历史负荷的日合理最大值时间序列分解为不同时间尺度下的多个IMF分

量(最后一个分量称为残余分量)。CEEMDAN是在经验模态分解(empirical mode decomposition, EMD)算法基础上改进的一种算法。

EMD是一种分析非线性、非平稳信号序列的算法,与小波分解有相似之处,但其无需预先设定基函数,因此避免了小波分解时基函数选择缺乏适应性的问题。该算法根据信号本身的特征将复杂的信号分解为若干个IMF分量,每个分量均对应某个时间尺度下的一种局部特征,相较于原始序列具有更强的平稳性和规律性^[17],但分解后的IMF分量中往往会存在模态混叠现象。

针对EMD算法的不足,集合经验模态分解(ensemble empirical mode decomposition, EEMD)被提出,EEMD在每次分解的过程中均向原始信号中加入均值为0、方差固定的白噪声,有效解决了EMD结果中存在的模态混叠问题,但在分解过程中引入的呈正态分布的白噪声会存在一定的残留,导致重构后的误差较大。为此,文献[18]提出CEEMDAN算法,该算法在数据分解的每个阶段均添加自适应的白噪声序列,能在较少的平均次数下提取不同时间尺度上的负荷序列特征信息,并得到重构误差几乎为0的若干个IMF分量,既避免了模态混叠现象的出现,又解决了EEMD重构误差大的问题。

定义: $Y_j(\cdot)$ 为通过EMD获得的第 j 个IMF分量的计算算子; $f(n)$ 为原始负荷时间序列, n 为时间序列中时间点的长度; $b^l(n)$ 为第 l 次加入的具有标准正态分布的白噪声序列; $f^l(n)$ 为第 l 次添加完标准正态分布白噪声后的负荷时间序列; I_k^l 为第 l 次EMD时得到的第 k 个IMF分量; $f_{\text{IMF}_m}(n)$ 为CEEMDAN产生的第 m 个IMF分量; β_d 为求解CEEMDAN第 $d+1$ 个IMF分量时的自适应系数。利用CEEMDAN算法进行负荷时序分解的步骤如下。

1)在原始负荷时间序列 $f(n)$ 中加入 Q 次具有标准正态分布的白噪声序列,形成 Q 个新的负荷时间序列,即:

$$f^l(n)=f(n)+\beta_0 b^l(n) \quad l=1, 2, \dots, Q \quad (4)$$

2)分别对每个新的负荷时间序列进行EMD,计算所有第1个IMF分量的平均值,得到CEEMDAN的第1个IMF分量 $f_{\text{IMF}_1}(n)$ 和第1个残余分量 $r_1(n)$,即:

$$f_{\text{IMF}_1}(n)=\frac{1}{Q} \sum_{l=1}^Q I_1^l \quad (5)$$

$$r_1(n)=f(n)-f_{\text{IMF}_1}(n) \quad (6)$$

3)在残余分量 $r_1(n)$ 中加入 Q 次具有标准正态分布的白噪声序列,构建 Q 个新的序列 $r_1(n)+\beta_1 Y_1(b^l(n))$ ($l=1, 2, \dots, Q$),分别对每个新序列进行EMD,计算所有第1个IMF分量的平均值,进而得到第2个IMF分量 $f_{\text{IMF}_2}(n)$,即:

$$f_{\text{IMF}_2}(n)=\frac{1}{Q} \sum_{l=1}^Q Y_1(r_1(n)+\beta_1 Y_1(b^l(n))) \quad (7)$$

4)对于其余的每个阶段,经过计算可以得到CEEMDAN的第 m 个残余分量 $r_m(n)$ 和第 $m+1$ 个IMF分量 $f_{\text{IMF}_{m+1}}(n)$,即:

$$r_m(n)=r_{m-1}(n)-f_{\text{IMF}_m}(n) \quad (8)$$

$$f_{\text{IMF}_{m+1}}(n)=\frac{1}{Q} \sum_{l=1}^Q Y_1(r_m(n)+\beta_m Y_m(b^l(n))) \quad (9)$$

5)重复执行步骤4),直至所获得的残余分量极值点个数至多不超过2个时,EMD不再分解,算法终止。假设算法终止时CEEMDAN产生 M 个IMF分量,原始信号序列 $f(n)$ 最终被分解为 M 个IMF分量和1个残余分量 $R(n)$,即:

$$f(n)=\sum_{m=1}^M f_{\text{IMF}_m}(n)+R(n) \quad (10)$$

4 基于LSTM的SLF

在以I类元胞为对象进行SLF时,通常采用传统的前馈神经网络构建预测模型,但前馈神经网络中的信息是逐层向前传播的,且传播过程中除了梯度信息外没有其他信息的反馈,这使得前馈神经网络忽略了负荷序列数据的时序关联性。循环神经网络(recurrent neural network, RNN)是具有记忆功能的神经网络,其内部包含的反馈结构使得当前时刻的输出由当前时刻的输入和先前时刻的输入所产生的影响共同决定。LSTM是一种特殊的RNN,其在RNN的基础上加入了3个门结构和1个记忆单元,有效解决了RNN在训练过程中可能出现的梯度弥散和梯度爆炸问题,大幅加快了模型的收敛速度。LSTM具有的长短期记忆能力使其可显著提高在时间序列预测领域的预测精度,这使其非常适合处理和预测时间序列中间隔和延迟相对较长的重要事件^[19]。因此,本文以各I类元胞历史负荷的日合理最大值时间序列为基础,分别对各序列分解得到的IMF分量建立LSTM模型,预测目标年中所有I类元胞各日负荷的最大值,进而得到所有I类元胞的年负荷最大值。

LSTM由输入层、隐含层和输出层3个部分组成,其中隐含层是由若干个记忆单元模块组成的,附录A图A2为记忆单元模块总体结构图。

在 t 时刻,一个记忆单元模块的输入由该时刻的输入向量 x_t 、上一时刻记忆单元的状态 c_{t-1} 和上一时刻隐含层的状态 h_{t-1} 组成,输出由该时刻记忆单元的状态 c_t 和隐含层的状态 h_t 组成。模块内部包含遗忘门、输入门和输出门这3个门,其中遗忘门以一定的概率控制是否遗忘上一时刻的记忆单元细胞信息,输入门决定当前时刻输入的信息有多少保存至当前的单元状态,输出门决定当前单元最终的输出信息,

计算公式如式(11)~(14)所示。

$$f_t = f_{\text{sig}}(W_{hf}h_{t-1} + U_{xf}x_t + B_f) \quad (11)$$

$$i_t = f_{\text{sig}}(W_{hi}h_{t-1} + U_{xi}x_t + B_i) \quad (12)$$

$$v_t = \tanh(W_{hc}h_{t-1} + U_{xc}x_t + B_c) \quad (13)$$

$$o_t = f_{\text{sig}}(W_{ho}h_{t-1} + U_{xo}x_t + B_o) \quad (14)$$

式中： f_t 、 i_t 、 o_t 分别为 t 时刻遗忘门、输入门和输出门的输出结果； v_t 为 t 时刻输入至单元的新信息； f_{sig} 表示Sigmoid激活函数，输入经过激活函数后可将输出控制在 $[0, 1]$ 之间； \tanh 为双曲正切激活函数，将输出缩放到 $(-1, 1)$ 之间； W_{hf} 、 W_{hi} 、 W_{ho} 分别为遗忘门、输入门、输出门对应于 h_{t-1} 的权值矩阵； W_{hc} 、 U_{xc} 分别为 h_{t-1} 、 x_t 与 v_t 的连接权值矩阵； U_{xf} 、 U_{xi} 、 U_{xo} 分别为遗忘门、输入门、输出门对应于 x_t 的权值矩阵； B_f 、 B_i 、 B_o 分别为遗忘门、输入门、输出门的偏置向量； B_c 为新信息的偏置向量。

t 时刻记忆单元模块的外部输出信息 c_t 、 h_t 的计算公式为：

$$c_t = f_t \cdot c_{t-1} + i_t \cdot v_t \quad (15)$$

$$h_t = o_t \cdot \tanh c_t \quad (16)$$

全部I类元胞日负荷合理最大值时间序列的每个IMF分量均有各自的预测模型，每个模型的输入就是对应IMF分量的负荷分量值。训练预测模型的样本容量应不少于4a的历史负荷数据，提前预测时间为1a。

5 算例分析

5.1 电力GIS的建立

选定待预测区域，该区域地理坐标为北纬 $43.77543^\circ \sim 43.87488^\circ$ ，东经 $126.47445^\circ \sim 126.55759^\circ$ 。根据收集到的区域规划方案，将待预测区域的用地类型划分为居民用地、商业用地、文化娱乐用地、工业用地、行政办公用地、绿化用地、市政设施用地和特殊用地8类。将不同类型的区域用不同的颜色进行标记，得到用地信息图层，如附录A图A3(a)所示。根据10kV馈线的供电范围建立对应的I类元胞图层，如附录A图A3(b)所示。

5.2 奇异数据的检测与修正

通过分析I类元胞实测历史负荷数据可知，各10kV馈线关口表的采样周期均为5min，即每个元胞一天会有288个负荷数据值。将各I类元胞2009—2012年的历史负荷数据以288个为1组，采用 3σ 准则对各组奇异数据进行检测。对于检测出的单点奇异值，使用该天的平均值进行替换；对于检测出的连续奇异值，使用该数据所处时间前后时间段的数据以及相似日的数据进行替换。提取修正后的每个I类元胞负荷序列中各日的最大负荷值，形成各个日合理最大负荷时间序列。

5.3 元胞负荷时间序列的分解

采用CEEMDAN算法对5.2节得到的各I类元胞历史负荷的日合理最大值时间序列进行分解，在此过程中，需要先确定自适应系数、最大迭代次数和加入白噪声的次数，由于自适应系数设置在0.01~0.5之间较合适，最大迭代次数和加入白噪声的次数过大会导致运算时间过长，因此，根据数据本身特点，本文将自适应系数设为0.2，最大迭代次数设为500，加入白噪声的次数设为100。

本文以珙春线的I类元胞为例，CEEMDAN的结果如附录A图A4所示。由图可看出，该序列被分解为9个频率由高到低排列的IMF分量和1个残余分量，每个分量的随机波动性均减小。在这些分量中，刻画随机波动特性的负荷表现为高频分量的形式，参考文献[20]，本文将过零率大于0.01的分量看作高频分量，过零率小于0.01的分量看作低频分量。通过计算可知：在珙春线中，IMF₁—IMF₇为高频分量，但这些分量的幅值均较小，预测误差相较于全局误差非常小；而对预测结果影响相对较大的IMF₈、IMF₉分量以及残余分量则更具规律性且较为平稳，有利于LSTM发挥其强大的非线性拟合能力。

对所得的各IMF分量以及残余分量进行重构，并将其与原序列进行比较，得到的重构误差图如图1所示。由图可看出，重构后的负荷序列的重构误差接近于0，可以忽略不计。

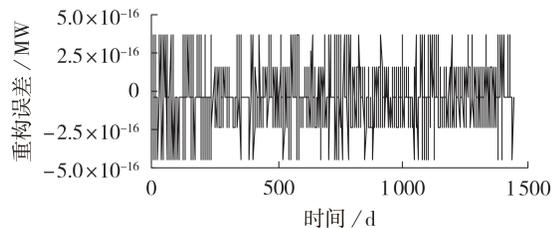


图1 CEEMDAN重构误差图

Fig.1 Diagram of CEEMDAN reconstruction error

5.4 预测模型的构建

本文在使用LSTM模型对2013年各I类元胞日最大负荷时间序列进行预测时，首先分别建立并训练适用于每个I类元胞各IMF分量的LSTM模型，然后将每个I类元胞所有IMF分量的预测结果进行线性叠加，得到2013年I类元胞的预测结果。模型的损失函数采用均方误差，初始学习率设为0.05，最大迭代次数设为250。使用自适应矩估计方法优化内部参数，经过多次实验最终将各模型的超参数确定为：输入层时间步数为7；隐含层层数为1，层中神经元个数为16，输出层神经元个数为1。每个I类元胞各IMF分量的具体预测策略如附录A图A5所示。

另外，为提升模型的收敛速度和预测精度，需对

数据进行归一化和反归一化处理,即先将数据归一化为[-1,1]之间的无量纲数值,再在完成训练和预测后对结果进行反归一化处理,得到带有初始量纲的预测结果。由于本文的预测工作是为规划服务的,因此后续用到的是各I类元胞预测结果的最大值。

5.5 预测结果分析

5.5.1 基于I类元胞的SLF结果

为验证3σ-CEEMDAN-LSTM预测方法的准确性,将该方法的预测结果与采用线性回归法、灰色理论法、指数平滑法、LSTM方法、CEEMDAN-LSTM方法、3σ-EMD-LSTM方法的预测结果进行比较,所得出的I类元胞部分负荷预测结果及其总体预测误差分别如附录A表A1和表1所示。

表1 目标年I类元胞负荷的总体预测误差
Table 1 Overall forecasting error of Class I cell load in target year

| 预测方法 | 平均绝对误差 / MW | 平均相对误差 / % |
|-----------------|-------------|------------|
| 线性回归 | 1.2395 | 36.57 |
| 灰色理论 | 1.2361 | 37.30 |
| 指数平滑 | 1.0476 | 30.73 |
| LSTM | 0.9321 | 21.44 |
| CEEMDAN-LSTM | 0.8277 | 18.47 |
| 3σ-EMD-LSTM | 0.7580 | 16.20 |
| 3σ-CEEMDAN-LSTM | 0.6871 | 15.04 |

由表1可看出,相较于其他6种SLF方法,本文提出的基于3σ-CEEMDAN-LSTM的SLF方法预测出的I类元胞负荷值的平均绝对误差和平均相对误差均为最低。

为了考察元胞的误差分布,分别对每种SLF方法预测结果中各相对误差区间内的I类元胞数进行统计,如图2所示。

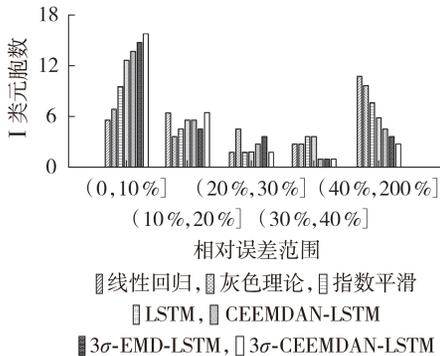


图2 基于I类元胞的各种SLF方法预测结果的误差分布
Fig.2 Error distribution of forecasting results of each SLF method based on Class I cells

由图2可知,基于3σ-CEEMDAN-LSTM的SLF方法的结果中有23个元胞的相对误差小于20%,占全部29个元胞的79.3%,分别优于线性回归法的44.8%、灰色理论法的37.9%、指数平滑法的51.7%、

LSTM方法的65.5%、CEEMDAN-LSTM方法的68.9%、3σ-EMD-LSTM方法的68.9%。

5.5.2 基于II类元胞的SLF结果

在电力GIS中,按照边长为0.3 km的等大小正方形网格生成II类元胞,并且采用空间电力负荷网格化技术^[16]将基于I类元胞的SLF结果转换为基于II类元胞的SLF结果。由于II类元胞数较多,限于篇幅,仅给出部分元胞的SLF结果,如附录A表A2所示。

为了更直观地展示II类元胞的预测结果,本文将全部预测结果在电力GIS中进行显示,给出各种SLF方法的II类元胞负荷平均总体误差,并对各相对误差区间内II类元胞数占待预测区域内元胞总数的比例进行统计,分别如附录A图A6、表2和图3所示。

表2 目标年II类元胞负荷的总体预测误差
Table 2 Overall forecasting error of Class II cell load in target year

| 预测方法 | 平均绝对误差 / MW | 平均相对误差 / % |
|-----------------|-------------|------------|
| 线性回归 | 0.0998 | 28.34 |
| 灰色理论 | 0.0911 | 24.43 |
| 指数平滑 | 0.0900 | 27.96 |
| LSTM | 0.0683 | 16.84 |
| CEEMDAN-LSTM | 0.0596 | 16.68 |
| 3σ-EMD-LSTM | 0.0420 | 10.19 |
| 3σ-CEEMDAN-LSTM | 0.0379 | 9.71 |

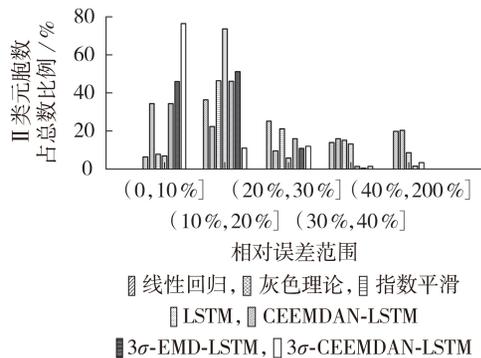


图3 基于II类元胞的各种SLF方法预测结果的误差分布
Fig.3 Error distribution of forecasting results of each SLF method based on Class II cells

由表2可看出,相较于其他6种SLF方法,基于3σ-CEEMDAN-LSTM的SLF方法预测出的II类元胞负荷值的平均绝对误差和平均相对误差均为最低。另外,由图3可知,基于3σ-CEEMDAN-LSTM的SLF方法对所有II类元胞的预测相对误差均小于40%,且其中有77%的II类元胞预测相对误差小于10%,远高于其他6种SLF方法。由此可见,本文所提出的基于3σ-CEEMDAN-LSTM的SLF方法相较于其他6种方法能更好地实现SLF。

6 结论

本文以电力系统规划为背景,提出了一种基于 3σ -CEEMDAN-LSTM 的 SLF 方法,其具有以下特点:

1) 通过使用 3σ 准则能够简便高效地确定每个 I 类元胞各日负荷的合理最大值,且能够有效地避免在 SLF 过程中奇异数据所导致的预测精度降低的问题;

2) 在分析各 I 类元胞负荷时间序列特征的过程中,运用 CEEMDAN 算法能够有效减小随机波动性给 SLF 结果带来的不利影响,为模型预测性能的提升创造了有利条件;

3) 对 I 类元胞的每个 IMF 分量依次构建合适的 LSTM 模型并进行预测,再将各元胞的所有分量预测结果进行叠加,进而实现 SLF,获得了较高的预测精度。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 肖白,周潮,穆钢. 空间电力负荷预测方法综述与展望[J]. 中国电机工程学报,2013,33(25):78-92.
XIAO Bai, ZHOU Chao, MU Gang. Review and prospect of the spatial load forecasting methods[J]. Proceedings of the CSEE, 2013, 33(25): 78-92.
- [2] 康重庆,夏清,刘梅. 电力系统负荷预测[M]. 北京:中国电力出版社,2007:94-97.
- [3] 王雁凌,吴梦凯. 经济新常态下基于偏最小二乘回归的中长期负荷预测模型[J]. 电力自动化设备,2018,38(3):133-139.
WANG Yanling, WU Mengkai. Medium and long term load forecasting model based on partial least-square regression under new normal economy[J]. Electric Power Automation Equipment, 2018, 38(3): 133-139.
- [4] 肖白,杨修宇,穆钢,等. 基于多变量分析的城市电网空间负荷预测方法[J]. 东北电力大学学报,2013,33(增刊1):39-44.
XIAO Bai, YANG Xiuyu, MU Gang, et al. Spatial electric load forecasting of urban power system based on multivariate analysis[J]. Journal of Northeast Electric Power University, 2013, 33(Supplement 1): 39-44.
- [5] 伊桂玲,张焰. 基于用地仿真法的配电系统空间负荷预测[J]. 电力自动化设备,2004,24(2):20-23.
YI Guiling, ZHANG Yan. Land usage-based spatial load forecasting in distribution system[J]. Electric Power Automation Equipment, 2004, 24(2): 20-23.
- [6] 朱凤娟,王主丁,陆俭,等. 考虑小区发展不均衡的空间负荷预测分类分区法[J]. 电力系统自动化,2012,36(12):41-48.
ZHU Fengjuan, WANG Zhuding, LU Jian, et al. Disequilibrium development areas based classification and subarea method for spatial load forecasting[J]. Automation of Electric Power Systems, 2012, 36(12): 41-48.
- [7] 肖白,张婕,姜卓,等. 基于秩次集对分析理论的空间负荷预测方法[J]. 电力自动化设备,2020,40(4):153-158.
XIAO Bai, ZHANG Jie, JIANG Zhuo, et al. Spatial load forecasting method based on rank set pair analysis[J]. Electric Power Automation Equipment, 2020, 40(4): 153-158.
- [8] 马星河,娄晨阳,赵军营,等. 基于人工蜂群算法的空间负荷预测[J]. 电力系统及其自动化学报,2018,30(8):102-107.
MA Xinghe, LOU Chenyang, ZHAO Junying, et al. Spatial load forecasting based on artificial bee colony algorithm[J]. Proceedings of the CSU-EPSA, 2018, 30(8): 102-107.
- [9] 庞昊,高金峰,杜耀恒. 基于神经网络融合的短期负荷预测方法[J]. 电力自动化设备,2020,40(6):37-43.
PANG Hao, GAO Jinfeng, DU Yaoheng. Short-term load forecasting method based on fusion of multiple neural networks[J]. Electric Power Automation Equipment, 2020, 40(6): 37-43.
- [10] 肖白,聂鹏,穆钢,等. 基于多级聚类分析和支持向量机的空间负荷预测方法[J]. 电力系统自动化,2015,39(12):56-61.
XIAO Bai, NIE Peng, MU Gang, et al. A spatial load forecasting method based on multilevel clustering analysis and support vector machine[J]. Automation of Electric Power Systems, 2015, 39(12): 56-61.
- [11] 肖白,赵晓宁,姜卓,等. 利用模糊信息粒化与支持向量机的空间负荷预测方法[J]. 电网技术,2021,45(1):251-260.
XIAO Bai, ZHAO Xiaoning, JIANG Zhuo, et al. Spatial load forecasting method using fuzzy information granulation and support vector machine[J]. Power System Technology, 2021, 45(1): 251-260.
- [12] 肖白,刘庆永,牛强,等. 基于元胞负荷特性分析的RBF神经网络空间负荷预测方法[J]. 电网技术,2018,42(1):301-307.
XIAO Bai, LIU Qingyong, NIU Qiang, et al. A spatial load forecasting method based on RBF neural network and cellular load characteristics analysis[J]. Power System Technology, 2018, 42(1): 301-307.
- [13] 梁雪峰. 基于CEEMD和AFPSO-LSSVM的空间负荷预测方法研究[D]. 吉林:东北电力大学,2020.
LIANG Xuefeng. Research on spatial load forecasting method based on CEEMD and AFPSO-LSSVM [D]. Jilin: Northeast Electric Power University, 2020.
- [14] 肖白,黄钰茹,姜卓,等. 数据匮乏场景下采用生成对抗网络的空间负荷预测方法[J]. 中国电机工程学报,2020,40(24):7990-8001,8236.
XIAO Bai, HUANG Yuru, JIANG Zhuo, et al. The method of spatial load forecasting based on the generative adversarial network for data scarcity scenarios[J]. Proceedings of the CSEE, 2020, 40(24): 7990-8001, 8236.
- [15] 肖白,黎平. 最佳电力负荷空间分辨率的获取方法[J]. 中国电机工程学报,2010,30(34):50-56.
XIAO Bai, LI Ping. Method for acquiring optimum spatial resolution of electric load[J]. Proceedings of the CSEE, 2010, 30(34): 50-56.
- [16] 肖白,杨修宇,穆钢,等. 基于元胞历史负荷数据的负荷密度指标法[J]. 电网技术,2014,38(4):1014-1019.
XIAO Bai, YANG Xiuyu, MU Gang, et al. A load density index method based on historical data of cell load[J]. Power System Technology, 2014, 38(4): 1014-1019.
- [17] 刘云鹏,许自强,董王英,等. 基于经验模态分解和长短期记忆神经网络的变压器油中溶解气体浓度预测方法[J]. 中国电机工程学报,2019,39(13):3998-4008.
LIU Yunpeng, XU Ziqiang, DONG Wangying, et al. Concentration prediction of dissolved gases in transformer oil based on empirical mode decomposition and long short-term memory neural networks[J]. Proceedings of the CSEE, 2019, 39(13): 3998-4008.
- [18] TORRES M E, COLOMINA M A, SCHLOTTHUERG, et al. A complete ensemble empirical mode decomposition with adaptive noise[J]. Brain Research Bulletin, 2011, 125(3): 4144-4147.
- [19] 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(1):135-146,404.

ZHAO Yaxue, WANG Xu, JIANG Chuanwen, et al. A novel short-term electricity price forecasting method based on correlation analysis with the maximal information coefficient and modified multi-hierarchy gated LSTM[J]. Proceedings of the CSEE, 2021, 41(1): 135-146, 404.

[20] 邓带雨,李坚,张真源,等. 基于EEMD-GRU-MLR的短期电力负荷预测[J]. 电网技术,2020,44(2):593-602.

DENG Daiyu, LI Jian, ZHANG Zhenyuan, et al. Short-term electric load forecasting based on EEMD-GRU-MLR [J]. Power System Technology, 2020, 44(2): 593-602.

作者简介:



肖白

肖白(1973—),男,教授,博士,通信作者,主要研究方向为电力系统规划、空间电力负荷预测、多能源互补协调发电、电能质量综合治理等(E-mail: xbxiaobai@126.com);

高文瑞(1996—),男,硕士研究生,研究方向为空间负荷预测(E-mail: 974282332@qq.com)。

(编辑 王锦秀)

Spatial load forecasting method based on 3σ -CEEMDAN-LSTM

XIAO Bai¹, GAO Wenrui¹, LI Daoming¹, QI Xuesong², KAN Zhongfeng²

(1. School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China;

2. Jilin Power Supply Company of State Grid Jilin Electric Power Company Co., Ltd., Jilin 132001, China)

Abstract: In order to effectively use the massive data collected from distribution network and improve the forecasting effect of spatial load, a spatial load forecasting method based on 3σ rule, complementary ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and long-short term memory neural network (LSTM) is proposed. Based on 3σ rule, the singular value detection and disposition for measured load data of each Class I cell are carried out. CEEMDAN technology is applied to decompose the processed load data of Class I cells into several intrinsic mode functions (IMFs) with different frequencies and amplitudes. LSTM model is built for each IMF component for forecasting. The spatial load forecasting based on Class I cells at the target year is obtained by linearly superimposed of the forecasting results of all the IMF components, on this basis, the spatial power load grid technology is used to obtain the spatial load forecasting results based on Class II cells. The correctness and effectiveness of the proposed method are verified by case analysis results.

Key words: spatial load forecasting; cell; geographic information system; CEEMDAN; LSTM

(上接第130页 continued from page 130)

Probability prediction method of equivalent inertia for high wind power penetration grid

BA Wenlan¹, WEN Yunfeng¹, YE Xi², WEN Ming³, HUANG Mingzeng¹, ZHANG Wuqi¹

(1. College of Electrical and Information Engineering, Hunan University, Changsha 410082, China;

2. State Grid Sichuan Electric Power Company, Chengdu 610041, China;

3. Economic & Technology Research Institute of State Grid Hunan Electric Power Co., Ltd., Changsha 410118, China)

Abstract: Driven by virtual inertia control technology, high wind power penetration grid contains different forms of inertia resources, and the system equivalent inertia presents the characteristic of complex nonlinear, time-varying and fluctuation. In order to better explain the uncertainty of system equivalent inertia, a probability prediction method of the system equivalent inertia considering the virtual inertia of wind turbine is proposed. Firstly, the point prediction model of system equivalent inertia is established through data-driven method, and the change trend of the equivalent inertia is predicted. Then, the non-parameter kernel density estimation is used to establish the probability density function of the prediction error for each time period, and the interval in which the system equivalent inertia may fluctuate under a certain confidence level at the prediction moment is obtained. An example is analyzed based on the modified IEEE RTS-79 system, and the results show that the proposed method is more reliable than the traditional parameter estimation method, it can provide useful decision-making information for new power system operation mode arrangement under low inertia scenarios.

Key words: doubly-fed induction generator; system equivalent inertia; inertia estimation; inertia forecasting; uncertainty

附录 A

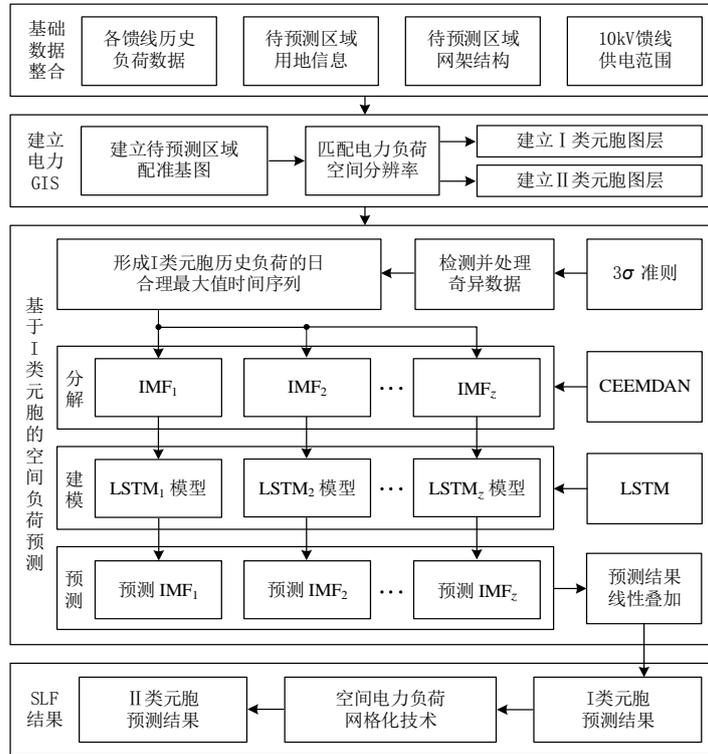


图 A1 3σ -CEEMDAN-LSTM 预测方法基本原理图

Fig.A1 Basic principle diagram of 3σ -CEEMDAN-LSTM forecasting method

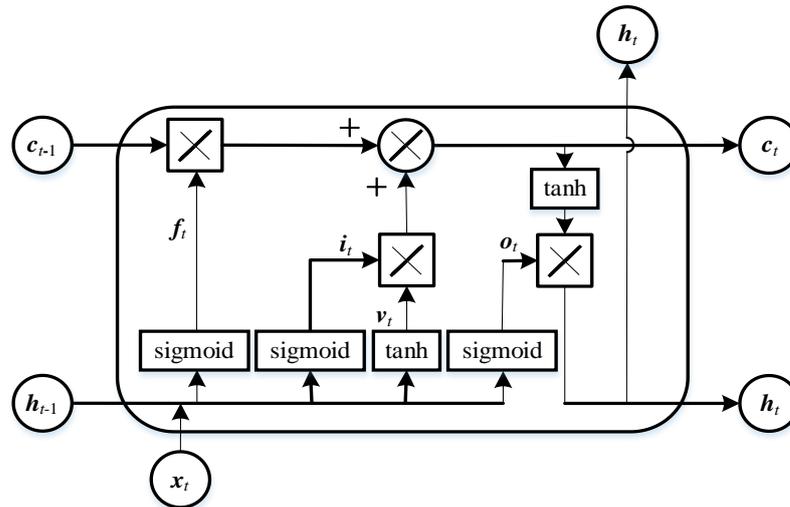
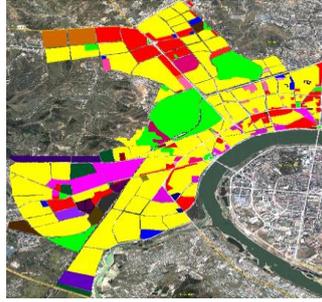
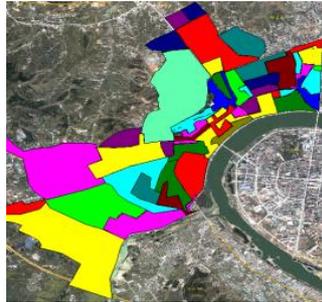


图 A2 LSTM 记忆单元模块结构

Fig.A2 Architecture of LSTM memory unit block



(a)用地信息图层



(b)I 类元胞图层

图 A3 用地信息图层及 I 类元胞图层

Fig.A3 Land use information layer and Class I cell layer

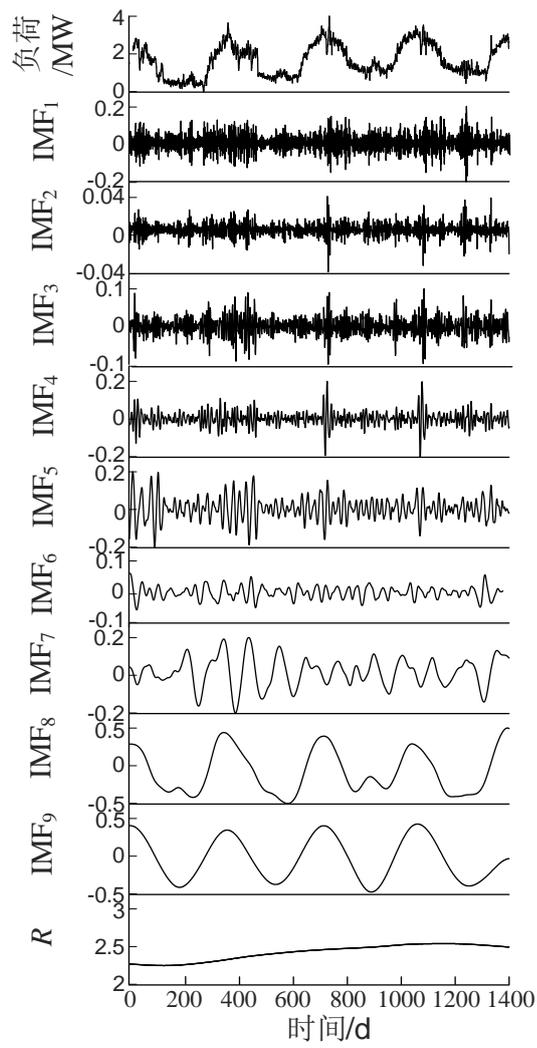
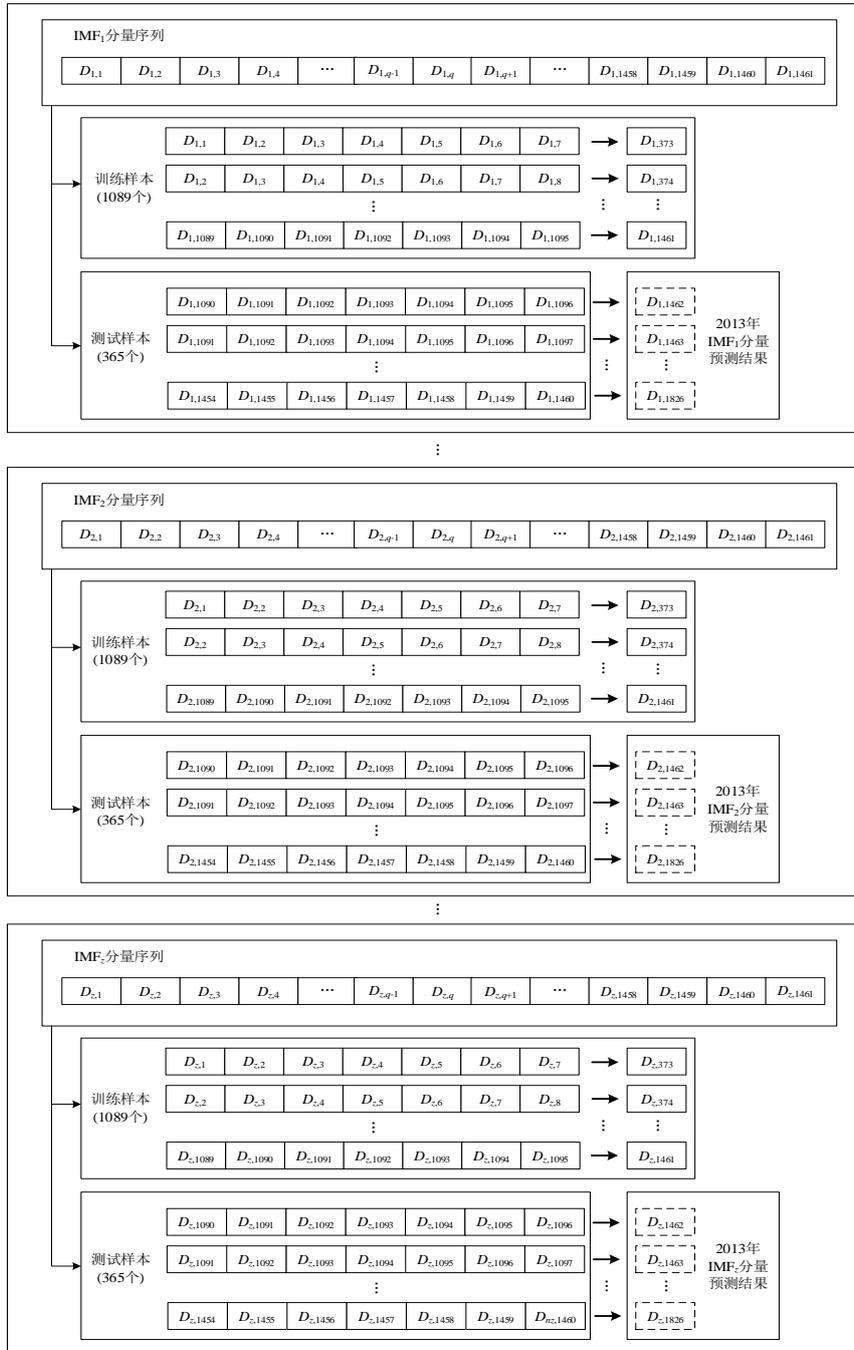


图 A4 珲春线负荷时间序列及 CEEMDAN 结果
 Fig.A4 Hunchun line load time series and CEEMDAN results



注: $D_{z,q}$ 为第 z 个 IMF 分量中的第 q 个负荷分量值。

图 A5 每个 I 类元胞各 IMF 分量的具体预测策略

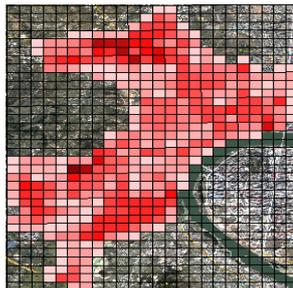
Fig.A5 Specific forecasting strategies for each IMF component of each Class I cell

表 A1 I 类元胞负荷预测结果
Table A1 Load forecasting results of Class I cells

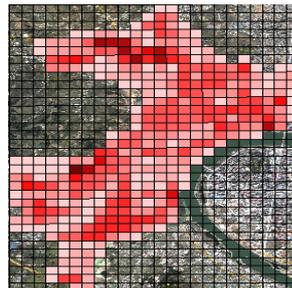
| 元胞名称 | 实际值/MW | 预测结果/MW | | | | | | |
|------|--------|---------|-------|-------|-------|--------------|----------------------|--------------------------|
| | | 线性回归 | 灰色理论 | 指数平滑 | LSTM | CEEMDAN-LSTM | 3 σ -EMD-LSTM | 3 σ -CEEMDAN-LSTM |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 北极线 | 4.393 | 4.916 | 4.352 | 4.818 | 4.209 | 4.394 | 4.445 | 4.501 |
| 城极线 | 2.275 | 2.290 | 2.161 | 2.286 | 2.254 | 2.278 | 2.301 | 2.261 |
| 德胜线 | 2.360 | 3.434 | 4.360 | 3.477 | 1.999 | 2.451 | 2.639 | 2.400 |
| 松北线 | 5.600 | 5.645 | 5.168 | 5.711 | 6.127 | 6.145 | 6.243 | 5.627 |
| 新立线 | 10.398 | 9.883 | 9.556 | 9.484 | 9.520 | 9.630 | 9.631 | 9.970 |
| 长春线 | 4.734 | 5.405 | 5.451 | 5.092 | 4.787 | 4.613 | 4.607 | 4.604 |
| 琿春线 | 3.141 | 3.668 | 3.466 | 3.546 | 3.336 | 3.289 | 3.223 | 3.351 |
| 城松线 | 4.399 | 5.012 | 5.745 | 4.912 | 4.507 | 4.373 | 4.668 | 4.409 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

表 A2 II 类元胞负荷预测结果
Table A2 Load forecasting results of Class II cells

| 元胞名称 | 准实际值/MW | 预测结果/MW | | | | | | |
|----------|---------|---------|---------|---------|---------|--------------|----------------------|--------------------------|
| | | 线性回归 | 灰色理论 | 指数平滑 | LSTM | CEEMDAN-LSTM | 3 σ -EMD-LSTM | 3 σ -CEEMDAN-LSTM |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| C(21,15) | 0.084 2 | 0.100 4 | 0.094 2 | 0.108 3 | 0.072 5 | 0.080 3 | 0.080 1 | 0.080 6 |
| C(22,15) | 0.059 0 | 0.070 3 | 0.064 7 | 0.070 0 | 0.050 4 | 0.052 1 | 0.052 9 | 0.053 7 |
| C(4,16) | 0.005 8 | 0.008 3 | 0.007 8 | 0.006 9 | 0.003 9 | 0.004 4 | 0.004 8 | 0.004 6 |
| C(5,16) | 0.318 5 | 0.198 1 | 0.223 4 | 0.203 7 | 0.270 2 | 0.285 7 | 0.289 1 | 0.295 1 |
| C(6,16) | 0.000 6 | 0.000 7 | 0.001 3 | 0.000 4 | 0.000 5 | 0.0005 | 0.000 5 | 0.000 5 |
| C(7,16) | 0.030 7 | 0.042 2 | 0.039 4 | 0.036 4 | 0.022 3 | 0.024 4 | 0.026 2 | 0.025 3 |
| C(8,16) | 0.387 7 | 0.548 6 | 0.513 3 | 0.459 3 | 0.270 9 | 0.301 5 | 0.327 0 | 0.313 3 |
| C(9,16) | 0.585 0 | 0.307 9 | 0.288 6 | 0.253 2 | 0.463 1 | 0.492 1 | 0.512 6 | 0.508 3 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |



(a) II 类元胞准实测值



(b) 线性回归法

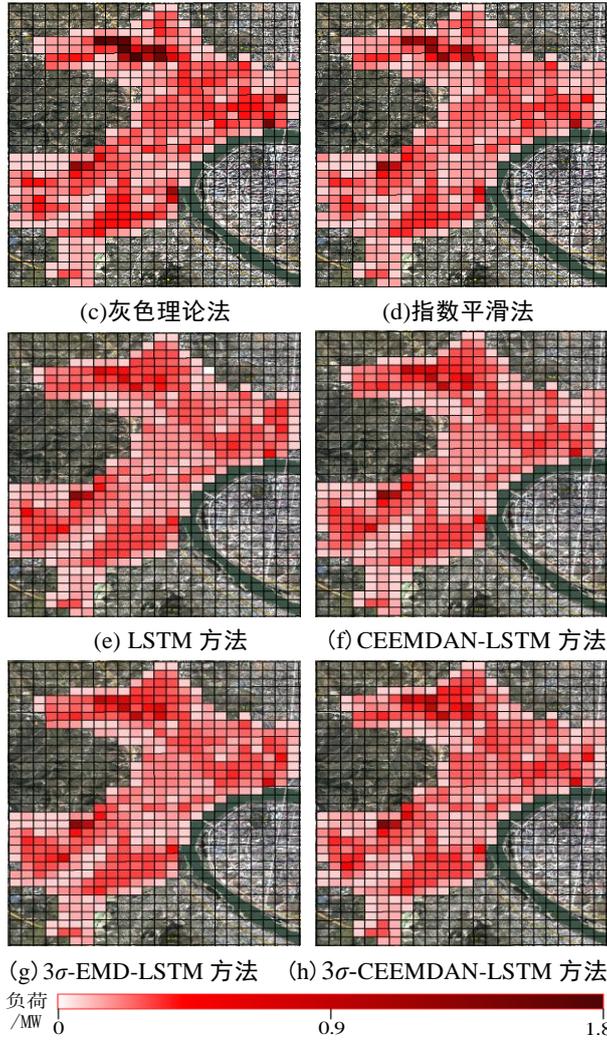


图 A6 II 类元胞负荷准实测值与预测值

Fig.A6 Quasi-measured and forecasted values of Class II cell load