

# 电量预测中的多分辨时序神经网络模型研究

石少青<sup>1</sup>, 郑楷洪<sup>2</sup>, 周尚礼<sup>2</sup>, 杨劲锋<sup>1</sup>, 张英楠<sup>2</sup>, 陈敏娜<sup>2</sup>, 张 健<sup>3</sup>

(1. 中国南方电网有限责任公司, 广东 广州 510663;

2. 南方电网数字电网研究院有限公司, 广东 广州 510663; 3. 中国矿业大学, 江苏 徐州 221116)

**摘要:**精确的电量预测可以为电量生产和供电调度提供可靠的指导,不同行业的用电量数据分布特性不同,用电量的周期性也因行业周期的不同具有不同时间尺度。传统的用电量预测模型难以在拟合用电量数据周期性的同时捕捉数据的波动性,为此本研究提出了多分辨时序神经网络(MTNN)模型,根据用电量数据特性将用电量数据视为近似的周期性数据分量和波动分量的组合,从多时频分辨率的角度对用电量数据进行多尺度分解,在不同时频尺度下预测用电量数据分量,然后通过组合预测分量实现用电量预测任务。具体而言,为捕捉数据周期性,首先使用集成经验模态分解方法(EEMD)对用电量数据分解,得到不同时间尺度的模态分量(IMF),使用 LSTM 预测高阶 IMF 分量;对于低阶 IMF 分量,设计了多阶时频注意力神经网络计算低阶 IMF 分量的预测值,最后组合多个 IMF 分量的预测值得到电量预测结果。实验结果表明,与常用的电量预测模型相比,MTNN 能更好地捕捉电量数据的变化规律,提高电量预测的精度。

**关键词:**电量预测;多分辨分析;时序神经网络;时频神经注意力

**中图分类号:**TM714

**文献标志码:**A

## Multi-resolution time sequence neural net model for electricity consumption forecasting

SHI Shaoqing<sup>1</sup>, ZHENG Kaihong<sup>2</sup>, ZHOU Shangli<sup>2</sup>, YANG Jinfeng<sup>1</sup>,

ZHANG Yingnan<sup>2</sup>, CHEN Minna<sup>2</sup>, Zhang Jian<sup>3</sup>

(1. China Southern Power Grid Co. Ltd., Guangzhou, Guangdong 510663, China;

2. China Southern Power Grid Digital Grid Research Institute Co. Ltd., Guangzhou, Guangdong 510663, China;

3. China University of Mining and Technology, Xuzhou, Jiangsu 221116, China)

**Abstract:** Accurate electricity consumption forecasting provides a reliable guidance for electricity production. However, the characteristics of electricity consumption data in different industries are different. But, rare electricity consumption forecasting models are specially designed to capture the fluctuation and the periodicity of electricity consumption data. To effectively predicting electricity consumption, this paper proposed a multi-resolution time sequence neural network (MTNN), regarding the electricity consumption data as combinations of approximate periodic components and fluctuation components. Specifically, to capture the periodicity of the data, the MTNN decomposed the power consumption data by using the ensemble empirical mode decomposition (EEMD) method to obtain intrinsic mode functions (IMF) in different time scales. Then, LSTM was used to predict the high-order IMF components. For the low-order IMF components, the MTNN designed a multi-order time-frequency attention neural network to calculate the predicted value of the low-order IMF components. The experimental results show that compared with the commonly-used electricity consumption forecasting models, the MTNN can better capture the char-

收稿日期:2020-10-14

基金项目:南方电网科技项目(670000KK52200081)

作者简介:石少青(1965—),男,福建云霄人,高级工程师,主要从事电能计量等方面研究。

张 健(1990—),男,山东泰安人,讲师,博士,主要从事机器学习、数据挖掘等方面研究,本文通信作者。

E-mail:zhangjian10231209@cumt.edu.cn

acteristics of electricity data, thus improving the accuracy of electricity consumption forecasting.

**Key words:** electricity consumption forecasting; multi-resolution analysis; time sequence neural network; time-frequency neural attention

电能稳定供给是社会经济发展和人民正常生活的必要保障,精确的电量预测可以为电能生产和供电调度提供可靠的指导,提升电力系统的供电质量。区域用电量数据的来源不同且不同来源的数据具有不同的统计特性,导致用电量数据在总体上呈现出一定周期性和波动性,为用电量预测带来了挑战。图1为南方电网某区域6、7月份的总电量和部分行业用电量。

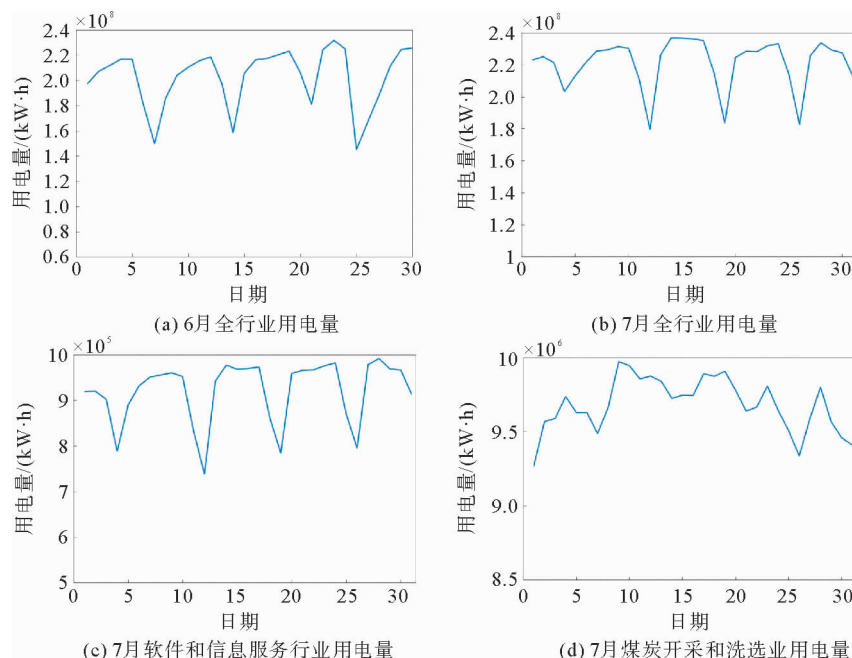


图1 南方电网某区域用电量变化曲线

Fig. 1 Electricity consumption curves of an area in China Southern Power Grid

从图1可以看出,区域用电量在时间尺度上呈明显的周期性,其中工作日以工业用电和商业办公用电为主,用电量除了部分波动外相对平稳;周末或节假日由于部分行业的用电量下降导致该时段用电量下降。这说明时间是电量预测中重要的影响因素,区域用电量和行业用电量在呈现出部分波动性的同时具有近似周期性,但是不同行业之间又表现出不同的统计特性。为了进行自动、精确的用电量预测,多种预测模型被应用到该领域中,传统的电量预测模型包括回归分析<sup>[1]</sup>、马尔可夫模型<sup>[2]</sup>、支持向量回归机<sup>[3]</sup>、时间序列分析模型<sup>[4-5]</sup>等。随着人工智能技术的发展,孪生(多生)支持向量回归机<sup>[6]</sup>、神经网络<sup>[7-9]</sup>等机器学习模型被应用于电量数据的处理中,其中长短时记忆(long short-term memory, LSTM)网络应用尤为广泛<sup>[10-13]</sup>。文献[11-12]将LSTM应用到短期电量、负荷预测中。文献[14]设计了多尺度LSTM,从时域出发,利用不同的时间步长提取数据中时序特征。文献[15]从时频分析的角度出发,将小波分析理论引入到时序数据的处理中,提取时频特征预测用电量,但是小波函数的构建比较复杂且使用同类小波拟合不同统计特性的数据时需要做更深尺度的分解,这会带来额外的计算负担并影响预测的准确性。文献[16]从信号分解的角度对用电量序列数据进行分解后从不同时间尺度进行预测。然而,由于用电量存在一定的波动性,且不同行业自身周期性的时间尺度不同,使得行业用电量的近似周期性体现在不同的时间尺度上,因此传统的用电量预测模型难以在拟合此类用电量数据周期性的同时捕捉数据的波动性。

为缓解此问题,本研究提出组合了集成经验模态分解(ensemble empirical mode decomposition, EEMD)方法、傅里叶变换和小波分析方法的多分辨时序神经网络(multiresolution time-sequence neural net,

MTNN)模型,使用 EEMD 从不同的时间分辨率上建模用电量数据的周期性特征,根据傅里叶变换和小波分析从频域和不同的时域分辨率上建模低阶 IMF 分量的周期性特征和波动性特征,然后组合 IMF 分量预测值得到预测结果。首先,使用 EEMD 方法从不同周期时长的角度对用电量序列数据进行分解,得到不同时间尺度下的本征模态函数(intrinsic mode function,IMF)分量。然后,对于高阶 IMF 分量和残余项,由于周期的时间尺度较大,比较容易预测其变化趋势,故使用 LSTM 模型来预测高阶 IMF 分量和残余项。其次,对于低阶 IMF 分量,考虑到其周期的时间尺度小,波动性更显著,故设计一种多阶时频注意力神经网络(multi-order time-frequency attention neural net,MTFANN)作为 MTNN 预测低阶 IMF 分量的子模型,该网络包含一种融合了小波分析和傅里叶变换的时频神经注意力神经网络和一种多阶特征提取方法,根据 IMF 低频分量及其时频变换特征来建模低阶 IMF 分量的周期性和波动性。最后,为了验证 MTNN 模型的有效性,本研究在人造的数据集上验证了 MTNN 能够在拟合数据周期性的同时捕捉数据的波动变化,然后在南方电网某区域 7 个地区的用电量数据集上验证了 MTNN 在用电量预测任务中的有效性。

## 1 小波分析、傅里叶变换和 EEMD 分解

### 1.1 小波分析

小波分析是一种多分辨分析方法,通过对信号进行多尺度分解,从不同分辨率上对信号进行分析,通过对信号进行小波分解可以在更细的分辨率上获取信号的变化特征,从而有效地建模信号的波动特性<sup>[15]</sup>。常用的小波分析分为连续小波变换和离散小波变换(discrete wavelet transform,DWT)。直观上看信号空间有不同正交小波构成的基底值的和组成:

$$\lim_{j \rightarrow -\infty} L^2(R) = \sum_j \oplus W_j, \quad (1)$$

式中, $W_j$  是由正交小波  $2^{j/2}\psi(2^j t)$  及其平移变换组构成的基底。此时,DWT 可以表示为:

$$W_f(2^{-j}, 2^{-j}k) = \int_{-\infty}^{\infty} f(t) \bar{\psi}_{j,k}(t) dt, \quad (2)$$

式中, $\bar{\psi}$  是  $\psi$  的复共轭,表示为:

$$\{\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), (j, k) \in \mathbf{Z}^2\}. \quad (3)$$

从多分辨分析的角度出发,信号空间还可以写为如下形式:

$$L^2(R) = V_{j_0} \oplus \sum_{j=j_0}^{\infty} \oplus W_j, \quad (4)$$

式中, $V_{j_0}$  可以看作是尺度函数空间,尺度基函数和小波基函数共同张成了信号空间。实际上,对于任何的尺度  $j_0$ ,都可以在信号空间找到一组标准正交基,该基底通过组合尺度  $j_0$  上的尺度函数及所有在尺度  $j (j > j_0)$  上的小波得到。基于该标准正交基,所有信号空间中的信号都可以写成基底中函数的线性组合:

$$f(n) = \sum_k c_{j_0,k} \varphi_{j_0,k}(n) + \sum_{j>j_0} \sum_k d_{j,k} \psi_{j,k}(n). \quad (5)$$

将数据序列进行小波分解的过程可以看作:每一层分解的结果是上次分解得到的低频信号再分解成低频和高频两个部分,如此经过  $N$  层分解后源信号  $f$  被分解为  $f = D_1 + D_2 + \dots + D_N + A_N$ ,其中  $D_1, D_2, \dots, D_N$  分别为第 1 层,第 2 层, ..., 第  $N$  层分解得到的高频信号,  $A_N$  为第  $N$  层分解得到的低频信号。由于小波变换的多尺度特性,显然适合随时间波动信号的建模。

### 1.2 傅里叶变换

傅里叶变换是常用的信号时频分析手段,当信号表现出一定周期性时,其傅里叶变换在频域上具有显著的能量分布曲线,能够将信号有效映射到频域,实现滤波处理,是提取信号周期性频域特征的重要方法。离散傅里叶变换可以表示为:

$$F(k) = \sum_{n=1}^N x(n) \exp\left(-j \frac{2\pi}{N} kn\right), k = 1, 2, \dots, N. \quad (6)$$

### 1.3 经验模态分解与集成经验模态分解

经验模态分解(empirical mode decomposition,EMD)是依据数据自身的时间尺度特征进行信号分解,

可将复杂信号分解为有限个本征模态函数 IMF, 分解出来的各 IMF 分量包含了原信号在不同时间尺度的局部特征信号。分解过程为:

- 1) 找到信号  $x(t)$  所有的极值点;
- 2) 用 3 次样条曲线拟合出上下极值点的包络线  $e_{\max}(t)$  和  $e_{\min}(t)$ , 并求出上下包络线的平均值  $m(t)$ , 求差值  $h(t) = x(t) - m(t)$ ;
- 3) 根据预设判据判断  $h(t)$  是否为 IMF 分量, 如果不是, 则以  $h(t)$  代替  $x(t)$ , 重复以上步骤直到  $h(t)$  满足判据, 则  $h(t)$  就是需要提取的 IMF 分量  $C_j(t)$ ;
- 4) 每得到一阶 IMF, 即将其从原信号中扣除, 重复以上步骤直到信号最后剩余部分  $r_n$  是单调序列或者常值序列。

这样, 经过 EMD 方法分解就将原始信号  $x(t)$  分解成一系列 IMF 以及剩余部分的线性叠加

$$x(t) = \sum_{j=1}^J C_j(t) + r_J(t) \quad (7)$$

但是该方法存在两方面问题, 一是用 EMD 分解得到的 IMF 存在模态混叠现象, 二是末端效应影响分解效果。为了抑制 EMD 的模态混叠现象, 基于噪声辅助分析的改进 EMD 方法即集成经验模态分解 EE-MD 方法被提出<sup>[17]</sup>, 其本质上是一种叠加高斯白噪声的多次经验模式分解方法。

## 2 多分辨时序神经网络模型

### 2.1 多分辨时序神经网络模型架构

MTNN 首先用 EEMD 方法对用电量数据进行分解和特征提取得到 IMF 分量, 旨在使用不同阶的 IMF 分量从不同的时间分辨率上建模用电量数据的近似周期性特征, 从而在不同时间尺度上捕捉用电量数据的近似周期性。然后针对不同阶的 IMF 分量设计神经网络模型预测下一时间段的 IMF 分量, 其中 LSTM 用于预测高阶 IMF 分量和残差项, 多阶时频注意力神经网络用于预测低阶 IMF 分量。MTNN 模型架构如图 2 所示。

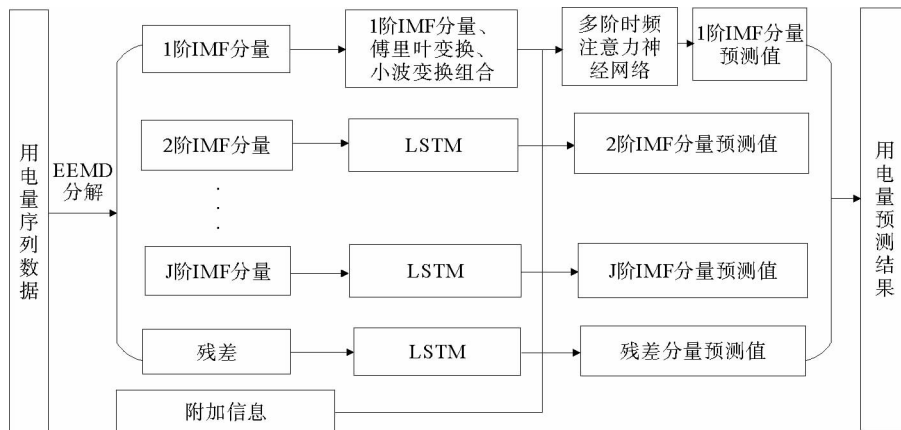


图 2 MTNN 模型架构图

Fig. 2 Structure of the MTNN model

如图 2 所示, 首先将用电量数据 EEMD 分解的过程表示为:

$$x(t) = \sum_{j=1}^J C_j(t) + r_J(t) \quad (8)$$

式中:  $J$  是 IMF 分量的残差项索引,  $r_J(t)$  表示残差项,  $C_j(t)$  表示 IMF 的第  $j$  阶分量。然后综合附加的天气信息和节假日信息, 将高阶 IMF 分量和残差项作为 LSTM 模型的输入, 计算高阶分量和残差项的预测值:

$$\begin{aligned} y_{\text{IMF}_j} &= \text{LSTM}_j([C_j(t), \text{add}(t)]), \\ y_{r_J} &= \text{LSTM}_J([r_J(t), \text{add}(t)]) \end{aligned} \quad (9)$$

其中:  $add(t)$  表示在  $t$  时间点的天气信息和节假日信息的编码;  $y_{IMF_j}$  表示 IMF 的第  $j$  阶分量的预测值, 满足  $j > 1$ 。对于低阶的 IMF 分量, 使用傅里叶变换和小波分析方法计算 IMF 分量的时频特征, 综合附加的天气信息和节假日信息作为多阶时频注意力神经网络(MTFANN)的输入, 计算低阶 IMF 分量的预测值

$$y_{IMF_1} = MTFANN([C_1(t), FFT(C_1(t)), DWT(C_1(t)), add(t)]), \quad (10)$$

式中:  $C_1(t)$  表示 1 阶 IMF 分量, FFT 表示快速傅里叶变换, DWT 表示离散小波变换。本研究采用 Coiflets 母小波, 小波变换层数为 3。最终, 综合 IMF 分量和残差的预测结果, 得到用电量数据的预测值

$$y = y_{IMF_1} + \sum_{j=2}^J y_{IMF_j} + y_{r_j}。 \quad (11)$$

## 2.2 多阶时频注意力神经网络

在低阶 IMF 分量的预测过程中, 为了建立低阶 IMF 分量的波动特性和周期性模型, 构建多阶时频注意力神经网络。首先对低阶 IMF 分量使用小波变换提取不同尺度的时频表达  $T_{IMF}$ , 使用傅里叶变换提取不同的频域表达  $F_{IMF}$ ; 再组合 IMF 分量、 $T_{IMF}$ 、 $F_{IMF}$  和附加的天气、节假日信息编码输入  $C_{IMF}$ ; 最后, 使用多阶时频注意力神经网络提取  $C_{IMF}$  的时频特征。多阶时频注意力神经网络包括多阶特征提取方法<sup>[18]</sup>和时频神经注意力神经网络的构建, 其中时频神经注意力神经网络包括时频神经相似函数、注意系数和注意特征的构建。

把多阶特征提取思想从自然语言处理引入到时序数据分析中。首先逐个遍历节点, 提取输入数据本身的特征(1 阶特征提取), 然后对节点进行成对遍历, 提取成对输入数据的组合特征(2 阶特征提取)。以此类推, 至  $N$  阶遍历提取  $N$  阶特征。由此, 多阶特征提取方法在考虑单个时间节点特征的基础上, 以多个时间节点的数据为一个整体, 考虑各个整体之间的关联关系,  $N$  阶特征提取可以用大小为  $N$  的卷积核实现。最后, 对每阶特征构建时频注意力神经网络提取其时频特征来预测低阶 IMF 分量。

为了表达输入之间由于相似性不同产生的不同影响, 定义了时频神经相似函数。时频神经相似函数通过测量两个包含有时频特征的输入向量之间基于神经网络的距离测度, 来度量两者的相似性。以  $C_{IMF}$  的特征提取过程为例, 其相似度为:

$$Similarity_{ij} = Neural(Dot(C_{IMF_i}, Cosine(W \times C_{IMF_i}, W \times C_{IMF_j})))。 \quad (12)$$

式中:  $W$  为可学习权值,  $C_{IMF_i}$  为第  $i$  个组合输入向量,  $C_{IMF_j}$  为第  $j$  个组合输入向量,  $Dot$  为点乘函数,  $Neural()$  为单层神经网络。时频神经相似函数计算其他向量与当前向量的夹角, 利用该夹角点乘当前向量, 计算这些向量对当前向量产生的影响。

由时频神经相似函数, 定义组合输入向量  $i$  和  $j$  之间的注意力系数为:

$$\alpha_{ij} = \text{softmax}(Similarity_{ij}) = \frac{\exp(Similarity_{ij})}{\sum_j \exp(Similarity_{ij})}。 \quad (13)$$

当前节点  $i$  的注意力特征  $h_i$  可以表示为:

$$h_i = \text{sigmoid}(\sum_j \alpha_{ij} V \times C_{IMF_j})。 \quad (14)$$

式中  $V$  为可学习权值。

由此构建一个由时频神经相似函数构成的时频神经注意力层, 多个时频神经注意力层堆叠得到时频神经注意力神经网络, 综合多阶特征提取方法, 可以构建多阶时频注意力神经网络。以阶数  $N=2$  为例, 一个二阶时频注意力神经网络的模型结构如图 3 所示。

由图 3 所示的二阶时频注意力神经网络可以预测一阶 IMF 分量值, 而高阶 IMF 分量和残差可以通过 LSTM 预测, 组合 IMF 各分量的预测值可以得到用电量的预测结果。

## 3 实验分析

实验分为 3 部分, 第 1 部分是数据集和基准模型介绍; 第 2 部分通过人造的包含多个时间尺度的周期信号序列和噪声的叠加信号序列数据, 验证 MTNN 能够提取周期特征和波动特征, 并验证 MTNN 在南方电网某区域用电量预测的有效性; 第 3 部分通过消融实验验证本研究所提出的时频处理方法、时频神经相似函

数、多阶时频注意力神经网络的有效性。

### 3.1 数据集和基准模型

#### 3.1.1 数据集

为了验证提出的 MTNN 模型能够拟合数据不同尺度的周期性和波动性,构建包含多个时间尺度的周期信号序列和噪声的叠加信号序列数据的人造数据集,该信号序列在不同周期长度的正弦函数上增加了服从高斯分布的随机噪声:

$$S(n) = 3 \times \sin(4 \times T) + 3 \times \cos(9 \times T) + \sin(8.12 \times T + 1.2) + 1.5 \times (i \sim \text{Gaussian}(0, 1)).$$

式中:  $i$  表示随机变量,  $n$  表示样本。本实验  $n = 2\,000$ , 使用窗口尺寸为 50 的滑动窗口遍历序列

长度为  $n$  的人造数据,得到 1 951 条序列数据,每条序列数据的长度为 50,根据本窗口内的序列数据来预测下一个样本点的数据值。

第二个数据集为南方电网某区域 7 个地区 2020 年 1 月 1 日到 2021 年 11 月 1 日的用电量数据,每 24 h 记录一次总用电量  $A$ ,第一产业  $A_1$ 、第二产业  $A_2$ 、第三产业  $A_3$  的地区用电总量以及 9 个不同行业的用电总量,得到 670 天的区域总用电量数据,总计 56 280 条初始数据。

为了剔除无效数据,首先进行数据清洗,即对数据中包含的离群值(主要是异常值)进行识别和修正,去除离群点并使用邻接点的平均值取代,然后进行数据归一化处理。使用过去 14 天的电量值、天气和节假日信息预测下一天的用电量,对数据集做窗口滑动处理得到长度为 14 的序列数据(作为输入),对应的预测值为下一天的用电量(作为标签)。在 MTNN 中,将 1 阶 IMF 分量扩展,其中每天的数据包括当天用电量的 1 阶 IMF 分量、1 阶 IMF 分量的快速傅里叶变换值、小波变换值、对应的天气数据和节假日信息,以向量的形式拼接为  $14 \times 5$  的向量。实验采用十折交叉验证测试模型效果,采用最小均方误差(MSE)衡量预测效果,MSE 值越小表明模型预测值越接近真实值。

#### 3.1.2 基准模型

采用三类基准模型与本研究模型进行对比。第一类为普通时序预测模型:LSTM<sup>[8]</sup>、SVR<sup>[3]</sup>;第二类为基于时频分析的时序预测模型:基于小波变换和用电量序列数据的 DWT-LSTM<sup>[13]</sup>、基于傅里叶变换和用电量序列数据的 FTT-LSTM<sup>[12]</sup>、多尺度 LSTM 模型 MSD-LSTM<sup>[14]</sup>;第三类为基于信号分解的时序预测模型:基于 EMD 分解和优化 SVR 的模型 EMD-SVRCKH<sup>[16]</sup>、基于 EEMD 分解和 LSTM 的 EEMD-LSTM 模型<sup>[19]</sup>。

### 3.2 MTNN 预测有效性验证

用人造数据集验证 MTNN 模型捕捉数据近似周期性和波动性特性的能力,数据集归一化的采样序列和 EEMD 分量如图 4 所示。

图 4 中 Original sequence 表示规则化后的数据,图 4(a)横坐标为  $[0, 2\pi]$  范围内 50 个数据采样点,纵坐标为横坐标对应的函数值;图 4(b)横坐标表示天数,共记录了 100 天的数据,纵坐标表示当天规则化后的用电量。可见,不管是人造数据集还是真实的用电量数据集,随着 IMF 阶数的增加,IMF 分量趋于平滑,对于 1 阶 IMF 分量,其波动性比较明显。MTNN 在 2 阶及以上 IMF 分量上使用 LSTM,在 1 阶 IMF 上使用多阶时频注意力神经网络进行预测。

首先,在人造数据集上验证 MTNN 的预测能力。取 LSTM 层数为 3,隐层节点数为 100,注意力层数为 3,阶数为 2;离散小波变换级数为 3,每层节点数为 100, batch\_size 为 16,梯度更新方法选择 Adam,初始学习率为  $10^{-3}$ ,采用 dropout 方法防止过拟合,dropout rate 选为 0.2。在对比实验中,MSD-LSTM 每层的跳跃尺度分别设为 2、4、6<sup>[14]</sup>,MTNN 在人造数据集上的预测结果的测试 MSE 如表 1 所示。由表 1 可见,基于信号分解的预测模型的测试 MSE 优于基于时频分析的 LSTM 模型,而 MTNN 取得了最优的测试 MSE,表

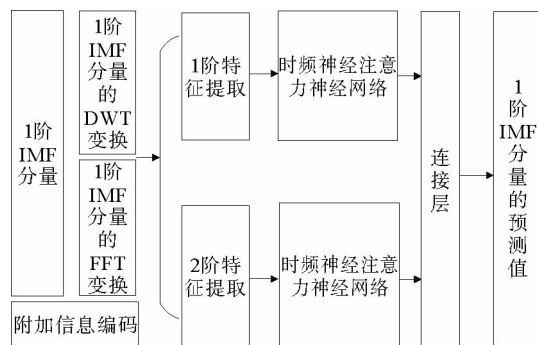


图3 二阶时频注意力神经网络结构图

Fig. 3 Structure of the two-order time-frequency attention neural net

明针对包含多个时间尺度的周期信号序列和噪声的叠加信号序列,MTNN 能够更有效地捕捉周期性特征和波动性特征,并完成预测。

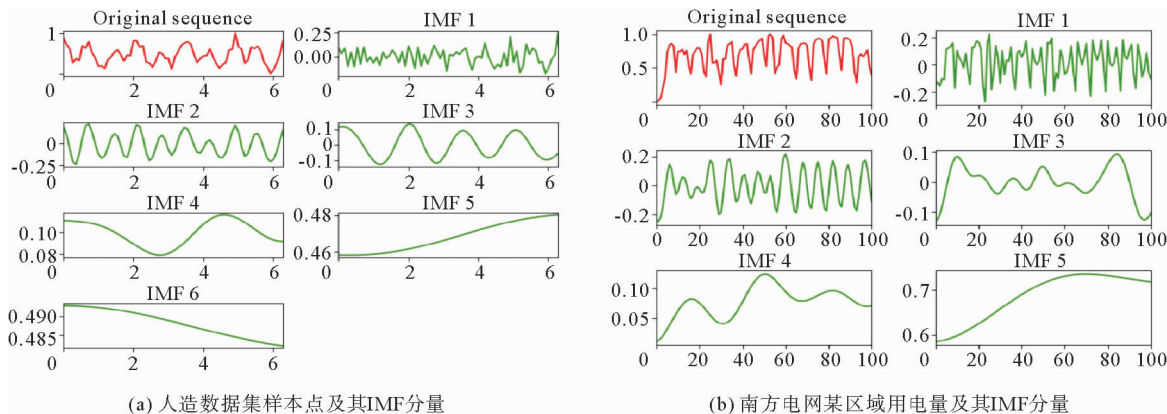


图 4 实验所采用的数据采样序列及其 IMF 分量

Fig. 4 Sample sequences and their IMFs of used data in the experiments

表 1 人造数据上的测试 MSE

Tab. 1 Test MSE on manufactured dataset

	MTNN	LSTM	SVR	DWT-LSTM	FFT-LSTM	MSD-LSTM	EMD-SVRCKH	EEMD-LSTM
MSE	<b>0.023</b>	0.035	0.043	0.031	0.029	0.030	0.026	0.027

然后,验证 MTNN 模型在用电量预测上的有效性。取 MTNN 的 LSTM 层数为 3,隐层节点数为 64,注意力层数为 2,阶数为 2;离散小波变换级数为 3,每层节点数为 64,batch\_size 为 16,梯度更新方法选择 Adam,初始学习率为  $10^{-4}$ ,采用 dropout 方法防止过拟合,dropout rate 选为 0.2。在对比实验中,MSD-LSTM 每层的跳跃尺度分别设为 1、2、3<sup>[14]</sup>。测试 MTNN 模型在 7 个不同地区用电量的预测效果,验证模型对不同地区用电量的适用性,实验结果如表 2 所示。由表 2 可以看出,模型取得的预测 MSE 与人造数据集的结果基本一致,且 MTNN 取得的结果最优,表明 MTNN 能够对不同尺度周期性和波动性的序列数据建立模型并有效预测用电量。一方面说明从不同时间尺度的周期性和波动性的角度,分析用电量数据的特征并进行用电量预测的思路框架是有效的,另一方面说明 MTNN 模型预测用电量是有效的。尽管预测的 MSE 不同,但是在这 7 个不同的区域上,本研究提出的 MTNN 模型取得了最优的测试 MSE,相对于基于 LSTM 的模型和基于 EMD、EEMD 分解的预测模型有一定的提升。

表 2 各地区电量测试 MSE

Tab. 2 Test MSE on total electricity consumption in each area

数据集	MTNN	LSTM	SVR	DWT-LSTM	FFT-LSTM	MSD-LSTM	EMD-SVRCKH	EEMD-LSTM
地区 1	<b>0.037 9</b>	0.041 1	0.041 9	0.039 9	0.040 8	0.040 2	0.038 2	0.038 1
地区 2	<b>0.037 5</b>	0.039 1	0.041 1	0.039 0	0.039 3	0.039 0	0.037 9	0.038 1
地区 3	<b>0.050 1</b>	0.060 3	0.080 9	0.059 8	0.060 0	0.059 4	0.051 6	0.050 2
地区 4	<b>0.029 3</b>	0.031 5	0.042 0	0.030 7	0.031 1	0.030 3	0.030 1	0.029 5
地区 5	<b>0.037 7</b>	0.041 3	0.051 1	0.040 2	0.042 4	0.039 8	0.039 9	0.038 0
地区 6	<b>0.029 9</b>	0.030 7	0.041 3	0.030 7	0.030 9	0.030 0	0.030 2	0.030 4
地区 7	<b>0.029 6</b>	0.032 3	0.039 4	0.030 9	0.032 1	0.030 2	0.030 7	0.030 3

为了进一步测试模型对具有不同周期性和分布特性的不同行业用电量的适应性,本研究对地区 1 的三大产业和工业、制造业、城镇居民用电量分别进行了预测,测试结果如表 3 所示。

表 3 地区 1 用电量数据测试 MSE

Tab. 3 Test MSE on electricity consumption in No.1 area

数据集	MTNN	LSTM	SVR	DWT-LSTM	FFT-LSTM	MSD-LSTM	EMD-SVRCKH	EEMD-LSTM
第一产业	<b>0.034 1</b>	0.037 9	0.041 3	0.036 5	0.036 9	0.036 2	0.035 1	0.034 7
第二产业	<b>0.037 6</b>	0.038 2	0.042 7	0.038 5	0.038 7	0.038 3	<b>0.037 6</b>	0.037 7
第三产业	<b>0.037 8</b>	0.040 3	0.047 9	0.039 4	0.039 2	0.040 1	0.038 5	0.038 0
工业	<b>0.039 1</b>	0.040 5	0.047 7	0.040 7	0.040 8	0.039 3	0.039 2	0.039 3
制造业	<b>0.037 7</b>	0.041 3	0.042 0	0.040 2	0.040 5	0.039 9	0.038 4	0.037 9
城镇居民	<b>0.042 5</b>	0.047 7	0.059 4	0.044 4	0.044 9	0.044 3	0.043 3	0.042 9

不同行业的用电量数据具有不同的分布特性。从表 3 可知,在不同行业用电量数据的预测上,MTNN 的表现比较稳定,可以有效地预测近似周期、周期性不明显的用电量数据以及数据中的瞬时变化特性,并且 MTNN 在大部分数据集上取得的效果要优于其他的对比算法。

### 3.2 MTNN 消融实验

设计消融实验,来验证 MTNN 中时频处理方法、时频神经相似函数、多阶时频注意力神经网络在用电量预测任务上的有效性。

#### 1) 时频处理方法的有效性验证

在 MTNN 中去掉 EEMD 后的主体模型记为多阶时频注意力神经网络 MTFANN 模型,在 MTNN 中使用 LSTM 替换 MTFANN 后得到的模型记为 EEMD-LSTM,在 MTNN 中的 MTFANN 中仅使用傅里叶变换而不使用小波变换的模型记为 EEMD-MFANN,在 MTNN 中的 MTFANN 中仅使用小波变换而不使用傅里叶变换的模型记为 EEMD-MTANN,模型在 7 个地区用电量的测试 MSE 如表 4 所示。

表 4 各地区用电量测试 MSE

Tab. 4 Test MSE on electricity consumption in each area

数据集	MTNN	MTFANN	EEMD-LSTM	EEMD-MFANN	EEMD-MTANN
地区 1	0.037 9	0.040 1	0.038 1	<b>0.037 6</b>	0.037 8
地区 2	<b>0.037 5</b>	0.038 9	0.038 1	0.037 8	0.037 7
地区 3	0.050 1	0.058 8	0.050 2	0.050 5	0.050 0
地区 4	<b>0.029 3</b>	0.030 3	0.029 5	0.029 6	0.029 4
地区 5	<b>0.037 7</b>	0.039 3	0.038 0	0.037 9	0.038 0
地区 6	0.029 9	0.030 1	0.030 4	0.029 6	0.029 3
地区 7	<b>0.029 6</b>	0.030 2	0.030 3	0.029 7	0.029 9

如表 4 所示,综合 EEMD、小波分析、傅里叶变换的 MTNN 模型在 4 个地区的用电量预测上取得最优的测试 MSE,说明本研究使用的时频处理方法对用电量预测任务是有帮助的。

#### 2) 时频神经相似性函数的有效性验证

将时频神经相似性函数分别替换为 Conise 相似性函数得到 MTNN-Conise 模型、Dot 相似性函数得到 MTNN-Dot 模型、神经相似性函数(使用神经网络计算相似性)得到 MTNN-Neural 模型、使用连接函数得到 MTNN-Concat 模型,各模型在 7 个地区用电量的测试 MSE 如表 5 所示。



表 5 各地区用电量测试 MSE  
Tab. 5 Test MSE on electricity consumption in each area

数据集	MTNN	MTNN-Conise	MTNN-Dot	MTNN-Neural	MTNN-Concat
地区 1	0.037 9	<b>0.037 7</b>	0.038 1	0.037 9	0.038 0
地区 2	<b>0.037 5</b>	0.037 9	0.037 7	0.037 9	0.037 9
地区 3	<b>0.050 1</b>	0.050 3	0.050 5	0.050 3	0.050 7
地区 4	<b>0.029 3</b>	0.029 4	0.029 5	0.029 4	0.029 9
地区 5	0.037 7	0.037 9	0.038 2	<b>0.037 4</b>	0.037 8
地区 6	<b>0.029 9</b>	0.030 3	0.030 2	0.030 3	0.030 6
地区 7	<b>0.029 6</b>	0.029 7	0.030 2	0.029 7	0.030 1

由表 5 可见,相比于常用的几种相似性度量函数,本研究提出的基于时频神经相似函数的 MTNN 模型在 5 个地区的用电量预测上取得了最优的测试 MSE,说明时频神经相似函数能够更好地度量 IMF 低阶分量的变化特性。

3) 多阶特征提取方法中阶数  $N$  的影响

$N=1$  时,相当于只考虑单个时间点的数据特征,使用注意力机制建模时间点之间的关联性; $N=2$  时,相当于将两个相邻节点的特征作为一个整体特征来考虑。 $N$  取不同值时,模型在 7 个地区用电量的测试 MSE 如表 6 所示。

表 6 各地区用电量的测试 MSE  
Tab. 6 Test MSE on electricity consumption in each area

数据集	MTNN	MTNN- $N=1$	MTNN- $N=3$	MTNN- $N=4$	MTNN- $N=5$
地区 1	0.037 9	0.038 0	<b>0.037 7</b>	<b>0.037 7</b>	0.037 8
地区 2	0.037 5	<b>0.037 4</b>	<b>0.037 4</b>	0.037 5	0.037 5
地区 3	<b>0.050 1</b>	<b>0.050 1</b>	0.050 2	0.050 2	0.050 2
地区 4	<b>0.029 3</b>	0.029 4	0.029 4	0.029 5	0.029 5
地区 5	<b>0.037 7</b>	0.037 9	<b>0.037 7</b>	<b>0.037 7</b>	<b>0.037 7</b>
地区 6	0.029 9	0.030 1	<b>0.029 8</b>	<b>0.029 8</b>	<b>0.029 8</b>
地区 7	<b>0.029 6</b>	<b>0.029 6</b>	0.029 8	0.029 9	0.029 9

如表 6 所示, $N=3$  时,MTNN 的预测效果开始停止增长,当  $N>3$  时,预测的 MSE 出现下降,可能是因为训练数据的数量上相对不足导致模型出现了过拟合,也可能是实现多阶机制的卷积操作使得多个节点的特征出现了平均效应,降低了预测的效果。

4 总结与展望

本研究提出一种多分辨时序神经网络模型 MTNN,根据用电量数据特性将用电量数据看作近似的周期性数据分量和波动分量的组合,使用 EEMD 方法从时域的角度对用电量序列数据进行分解,然后从时频多分辨率的角度对低阶 IMF 分量进行多尺度分解,构建多阶时频注意力神经网络预测低阶 IMF 分量,使用 LSTM 预测高阶 IMF 分量和残差,组合预测结果预测用电量。通过人造数据集、南方电网某区域 7 个地区的用电量数据以及消融实验验证了 MTNN 的有效性。

后续将利用更加丰富的数据信息完善 MTNN 模型,一方面利用更细粒度的模型预测实时用电量,另一方面将实时的附加信息以先验形式引入到模型中,进一步提高模型的预测效果。

参考文献:

[1]BIANCO V, MANCA O, NARDINI S.Linear regression models to forecast electricity consumption in Italy[J].Energy

- Sources, Part B, Economics Planning & Policy, 2013, 8(1/2/3/4): 86-93.
- [2] WANG X, MENG M. Forecasting electricity demand using Grey-Markov model[C]// International Conference on Machine Learning & Cybernetics. Kuming, Jul. 12-15, 2008: 1244-1248.
- [3] 齐德昱, 葛超, 葛韧. 混合核支持向量回归及对社会用电量的预测[J]. 重庆工学院学报, 2009(10): 50-52.  
QI Deyu, GE Chao, GE Ren. SVR using mixtures of kernels and its application in forecasting of electricity consumption of society[J]. Journal of Chongqing Institute of Technology, 2009(10): 50-52.
- [4] HOU X C, ZHU L J. Scenario analysis of city residential electricity consumption based on the time-series forecast: Taking Quanzhou City as an example[J]. Applied Mechanics & Materials, 2012, 178-181: 184-188.
- [5] LI D, CHANG C, CHEN C, et al. Forecasting short-term electricity consumption using the adaptive grey-based approach: An Asian case[J]. Omega, 2012, 40(6): 767-773.
- [6] 彭显刚, 王洪森. 基于竞争 ISPO 双胞胎支持向量回归短期负荷预测[J]. 电力系统及其自动化学报, 2014, 26(10): 46-50.  
PENG Xiangang, WANG Hongsen. Short-term load forecasting based on competitive ISPO and twin support vector regression[J]. Proceedings of the CSU-EPSA, 2014, 26(10): 46-50.
- [7] LI G. Electricity consumption forecast based on wavelet neural network[C]// International Conference on Information System & Artificial Intelligence. Hong Kong, Jun. 1, 2016: 361-364.
- [8] ZHENG K, QIAN B, LI S, et al. Long-short term echo state network for time series prediction[J]. IEEE Access, 2020, 8: 91961-91974.
- [9] MARVUGLIA A, MESSINEO A. Using recurrent artificial neural networks to forecast household electricity consumption [J]. Energy Procedia, 2012, 14(14): 45-55.
- [10] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [11] KIM N, KIM M, CHOI J. LSTM based short-term electricity consumption forecast with daily load profile sequences[C/OL]// Global Conference on Consumer Electronics. Nara, 2018. DOI: 10.1109/GCCE.2018.8574484.
- [12] SHAHID F, ZAMEER A, MUNEEB M. A novel genetic LSTM model for wind power forecast[J/OL]. Energy, 2021. DOI: 10.1016/j.energy.2021.120069.
- [13] LIU Y, GUAN L, HOU C, et al. Wind power short-term prediction based on LSTM and discrete wavelet transform[J/OL]. Applied Sciences, 2019, 9(6): 1108. DOI: 10.3390/app9061108.
- [14] 肖勇, 郑楷洪, 郑镇境, 等. 基于多尺度跳跃深度长短期记忆网络的短期多变量负荷预测[J]. 计算机应用, 2021, 41(1): 231-236.  
XIAO Yong, ZHENG Kaihong, ZHENG Zhenjing, et al. Multi-scale skip deep long short-term memory network for short-term multivariate load forecasting[J]. Journal of Computer Applications, 2021, 41(1): 231-236.
- [15] IMANI M, GHASSEMIAN H. Lagged load wavelet decomposition and LSTM networks for short-term load forecasting [C/OL]// International Conference on Pattern Recognition and Image Analysis (IPRIA 2019). Tehran, Mar. 6-7, 2019. DOI: 10.1109/PRIA.2019.8785972.
- [16] ZHANG Z, DING S, SUN Y. A support vector regression model hybridized with chaotic krill herd algorithm and empirical mode decomposition for regression task [J]. Neurocomputing, 2020, 410(13): 185-201.
- [17] WU Z, HUANG N. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in adaptive data analysis, 2009, 1(1): 1-41.
- [18] 杜鹏, 丁世飞. 基于混合词向量深度学习模型的 DGA 域名检测方法[J]. 计算机研究与发展, 2020, 57(2): 433-446.  
DU Peng, DING Shifei. A DGA domain name detection method based on deep learning models with mixed word embedding [J]. Journal of Computer Research and Development, 2020, 57(2): 433-446.
- [19] HUANG Y, ZHANG X, YANG L. Short-term wind speed prediction based on EEMD-LSTM[J/OL]. Journal of Physics Conference Series, 2019. DOI: 10.1088/1742-6596/1314/1/012105.

(责任编辑: 齐敏华)