

文章编号: 2095-2163(2021)03-0033-06

中图分类号: TP18

文献标志码: A

基于注意力机制和 EMD-GRU 模型的电力负荷预测

黄元行, 刘晓波, 范津玮, 熊焱, 叶琦

(贵州大学 电气工程学院, 贵阳 550025)

摘要: 为提高非平稳性电力负荷的预测精度以及充分挖掘历史负荷数据的时序特征, 本文提出了基于注意力(Attention)机制和经验模态分解(EMD)以及门控循环神经网络(GRU)组合的负荷预测方法。首先使用 EMD 对原始负荷数据进行 EMD 分解, 继而得到有限个具有本征模态函数(IMF)的分量; 然后考虑到各分量间的相关性, 使用多层 GRU 网络对 IMF 分量进行多输入多输出预测, 同时引入注意力机制, 深入挖掘历史负荷数据的时序相关性特征; 最后对多层神经网络输出的有限个分量预测结果进行重构得到最终的负荷预测值。算例分析采用贵州电网某地实际负荷数据, 经过与不同模型进行预测误差分析和对比, 本文所提方法能够有效提高预测精度。

关键词: 电力负荷预测; 经验模态分解; 门控循环神经网络; 注意力机制

Power load forecasting based on Attention mechanism and EMD-GRU model

HUANG Yuanxing, LIU Xiaobo, FAN Jinwei, XIONG Yan, YE Qi

(The Electrical Engineering College, Guizhou University, Guiyang 550025, China)

[Abstract] In order to improve the prediction accuracy of non-stationary electric load and fully exploit the time-series characteristics of historical load data, this paper proposes a method of load prediction based on a combination of Attention mechanism and empirical modal decomposition (EMD) and gated recurrent unit (GRU). Firstly, EMD is used to decompose the original load data, and then some components with intrinsic modal functions (IMFs) are obtained; subsequently, considering the correlation among these components, a multi-layer GRU network is used to predict the IMF components with multiple inputs and multiple outputs, and the Attention mechanism is also introduced to deeply explore the time-series correlation characteristics of the historical load data; finally, the results of load prediction are obtained by reconstructing the finite components from the multilayer neural network. The actual load data of a place in Guizhou power grid are used to the simulative analysis, and the prediction errors of different models are compared, which shows the proposed method could effectively improve the prediction accuracy.

[Key words] power load forecasting; empirical modal decomposition; gated recurrent unit; Attention mechanism

0 引言

随着智能电网的发展, 配电侧用电负荷呈现多元化, 电网负荷波动性和尖峰性较为明显, 其对电网的调度和运行规划提出了更高的要求。因此, 准确而实时的负荷预测有助于电网的经济和可靠运行。同时随着调度系统对负荷的采集频率和数据准确性也不断提升, 为负荷特性分析和负荷预测提供了高质量、海量化的数据基础^[1]。

过去由于预测算法的限制以及计算机并行处理的能力和电力负荷大数据的不足, 负荷预测通常采用经典的预测方法如回归分析法^[2]、时间序列法^[3]和传统的负荷求导法、指数平滑法^[4]、卡尔曼滤波法^[5]以及灰色预测法^[6]等。近年来, 随着计算机和

智能算法的发展, 以小波分析法、支持向量机、专家系统以及人工神经网络为代表的智能预测方法使得负荷预测的精度越来越高。文献[7]通过循环神经网络(Recurrent Neural Network, RNN)将上一时刻的神经元状态值输入当前神经元, 实现了时间序列数据的挖掘, 但是循环神经网络无法对长时序数据进行特征提取。文献[8]通过长短期记忆神经网络(Long Short-Term Memory, LSTM)网络有效解决了 RNN 梯度消失问题且可以有效地处理和挖掘长时间序列的有效信息, 但是单一的 LSTM 网络在数据特征提取上效果不明显。文献[9]先通过卷积神经网络(Convolutional Neural Networks, CNN)提取特征向量, 将特征向量以时序序列方式构造并作为门控循环神经网络(Gated Recurrent Unit, GRU)输入

作者简介: 黄元行(1993-), 男, 硕士研究生, 主要研究方向: 电力系统规划; 刘晓波(1964-), 女, 副教授, 主要研究方向: 电力系统运行与控制; 范津玮(1995-), 男, 硕士研究生, 主要研究方向: 配电网故障辨识与诊断; 熊焱(1994-), 男, 硕士研究生, 主要研究方向: 电力系统运行与控制; 叶琦(1992-), 男, 硕士研究生, 主要研究方向: 配电网网格化规划。

通讯作者: 黄元行 Email: 1083004291@qq.com

收稿日期: 2020-12-08

数据从而提高预测精度,但是 CNN 对单变量的负荷数据特征提取不明显。考虑到负荷数据的非平稳性特点,文献[10]通过经验模态分解(Empirical Mode Decomposition, EMD)对原始数据进行分解得到有限个本征模态(Intrinsic Mode Function, IMF)分量,然后再对每个分量进行 LSTM 模型预测,但是单独的分量重构未考虑到分量间的相关性。文献[11]使用了注意力机制与 LSTM 结合,利用注意力机制对历史数据进行侧重点的学习,增强网络的关注点。以上文献算法较于单一,在处理非平稳性负荷数据时有所不足。

为了减小直接对非平稳性时序数据直接预测所带来的误差,本文首先使用 EMD 对数据进行分解,得到有限个平稳性 IMF 分量;为了减少超参数设置,提高训练效率,同时考虑到由原始负荷数据进行 EMD 分解得到的每个 IMF 分量之间具有一定相关性,使用 GRU 网络对所有 IMF 分量进行多输入多输出预测;为了提高 GRU 网络对历史数据特征的学习能力,在 GRU 网络中引入注意力机制重点关注历史数据的时序相关性特征;最后对输出的 IMF 分量预测结果进行重构,得到最终的负荷预测结果。本文选取了 GRU 模型、EMD-GRU 模型以及本文提出的含注意力机制的 EMD-GRU 模型进行实验误差分析,经过算例分析表明,本文所提含注意力机制的 EMD-GRU 模型能够有效提高负荷预测精度。

1 原理分析

1.1 EMD 基本原理

经验模态分解(EMD)是一种新的自适应处理非平稳信号分析方法,适用于处理非线性、非平稳性的时间序列数据。短期的电力负荷数据常呈现非线性、非平稳性等特点,受影响因素较多,波动性较大,将其看作是标准的非平稳性时间序列数据,使用 EMD 对其进行分解可以得到不同时间尺度的平稳性分量。对时间序列数据 $x(t)$ 进行经验模态分解包含有以下几个步骤:

(1) 找出原始数据 $x(t)$ 的所有极大值点,并使用三次样条插值函数拟合出原始数据的上包络线 $U_s(t)$,并使用同样的方法找出原始数据的下包络线 $V_s(t)$,并运算求出均值包络线 m_1 。具体公式为:

$$m_1 = \frac{U_s(t) + V_s(t)}{2}, \quad (1)$$

(2) 使用原始信号减去均值包络线得到中间信号 $c_1(t)$ 。具体公式为:

$$c_1(t) = x(t) - m(t), \quad (2)$$

(3) 判断该中间信号是否满足 IMF 的 2 个条件,即:

① 在整个数据段内,极值点的个数和过零点的个数必须相等或相差最多不能超过一个。

② 在任意时刻,由局部极大值点形成的上包络线和由局部极小值点形成的下包络线的平均值为零。

如果满足这两个条件,该信号就是一个 IMF 分量;如果不满足,则以该信号为基础,重新做(1)~(3)的分析。

(4) 当中间信号满足 2 个条件时,令 $c_1(t) = IMF_1$,用原始信号减 IMF_1 ,作为新的原始信号。具体公式为:

$$r_1(t) = x(t) - IMF_1, \quad (3)$$

(5) 重复通过(1)~(4)的分析,可以得到 IMF_2 ,以此类推,完成 EMD 分解。最终得到 n 个 IMF 分量和一个残余分量 r_n 。具体公式为:

$$x(t) = \sum_i^n IMF_i + r_n. \quad (4)$$

1.2 GRU 基本原理

为了解决传统循环神经网络在训练长序列数据时产生的梯度消失以及梯度爆炸问题,长短期记忆神经网络(LSTM)通过在循环神经网络的神经元中加入输入门、输出门和遗忘门来控制存储单元中的信息,从而提高网络的记忆能力^[12]。但是在面对大量数据时,LSTM 需要设置的参数较多,训练时间长,收敛速度慢。

门控循环神经网络(GRU)是 LSTM 的一个变体^[13]。相较于具有 3 个门的 LSTM,GRU 只有 2 个门:更新门和重置门。使用更新门代替了 LSTM 中的输入门和遗忘门。这样的设计结构使得 GRU 参数设置减少,收敛速度更快,有效提高了训练效率^[14]。GRU 网络结构如图 1 所示。对图 1 中的结构阐释,拟做研究分述如下。

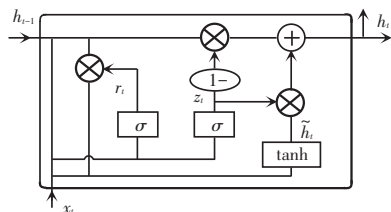


图 1 GRU 网络结构

Fig. 1 GRU network structure

(1) 更新门:决定了前一时刻有多少状态历史

信息被保留到当前状态中,同时决定了保留多少当前时刻的候选状态信息。此时需用到的公式为:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z), \quad (5)$$

其中, σ 为激活函数; W_z 、 b_z 为更新门的权重矩阵、偏置; h_{t-1} 、 x_t 为前一隐藏层的输出和当前的输入。

(2)重置门:决定了当前时刻的状态对上一时刻网络状态的依赖程度,此时需用到的公式为:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r), \quad (6)$$

其中, W_r 、 b_r 为重置门的权重矩阵、偏置。

在此基础上,研究推导可得候选隐藏状态的数学公式为:

$$\tilde{h}_t = \tanh(W_c \cdot [r_t \times h_{t-1}, x_t] + b_c), \quad (7)$$

其中, \tanh 为激活函数; W_c 、 b_c 为隐藏状态单元的权重矩阵、偏置。

最终,可得当前时刻的网络状态输出为:

$$h_t = z_t \times h_{t-1} + (1 - z_t) \times \tilde{h}_t. \quad (8)$$

1.3 Attention 机制

Attention 机制的本质是模仿人类视觉注意力的资源分配机制,特定时刻人的大脑只会将注意力集中在某一重点关注的区域^[15],从而减少对其他区域的关注。这样的关注方式可以减少资源损耗,从重点需要的区域获取更有用的特征信息,提高任务效率。从网络结构上来说,注意力机制通过对数据输入进行关注,学习出一个对数据特征的权重分布,再把学习到的权重施加到原来的数据特征上,为后续对数据的处理能够提供一个突出的特征影响,使重点特征得到更大的关注。Attention 机制的模型结构如图 2 所示。

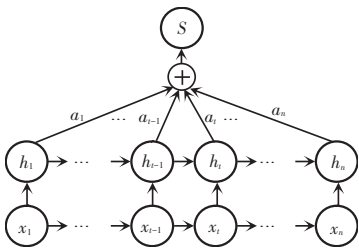


图2 Attention 机制网络结构

Fig. 2 Attention mechanism network structure

Attention 的计算过程是:首先根据 t 时刻 GRU 层的输出向量 h_t 计算得到注意力概率分布值 e_t ,此时的数学公式可写为:

$$e_t = \text{utanh}(u h_t + b), \quad (9)$$

其中, u 、 w 为权重系数; b 为偏置系数。

其次,采用 softmax 函数对注意力得分值进行

归一化得到 a_t 。这里的数学公式可写为:

$$a_t = \text{softmax}(e_t) = \frac{\exp(e_t)}{\sum_{j=1}^t \exp(e_j)}, \quad (10)$$

最后,再将计算得到的得分值 a_t 与向量 h_t 进行加权求和,得到最终输出的隐藏层状态值 s_t 。这里的数学公式可写为:

$$s_t = \sum_{i=1}^i a_i h_i. \quad (11)$$

2 含注意力机制的 EMD-GRU 预测模型

2.1 输入数据处理

为了便于模型的训练,本文先对输入数据进行处理,由于所使用的数据无缺失、异常等问题,现只需对输入数据进行 MinMaxScaler 归一化处理,公式为:

$$x_s = \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \quad (12)$$

其中, x_s 为归一化后的数据; x 为原始数据; x_{\min} 、 x_{\max} 为负荷数据的最小值和最大值。

2.2 含注意力机制的 GRU 网络结构设计

本文设计使用了 2 层 GRU 网络,并在 GRU 层后加入注意力层,而后通过一个全连接层输出全部分量的预测结果。含注意力机制的多层 GRU 网络结构如图 3 所示。

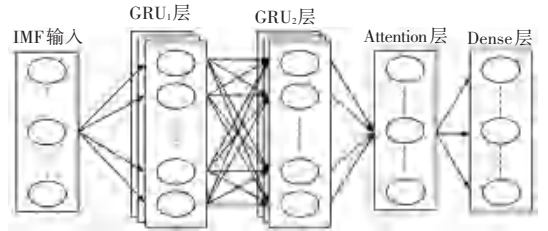


图3 含注意力机制的多层 GRU 网络结构

Fig. 3 Multi-layer GRU network structure with Attention mechanism

基于 keras 的 GRU 模型参数设置为:2 层 GRU 网络,每层网络神经元数目为 128 个,激活函数使用修正线性单元(ReLU),并采用 100 次 epochs 迭代训练,损失函数采用均方误差 MSE ,梯度优化函数使用 $Adam$ 可以更快收敛,并提高准确率。

2.3 含注意力机制的 EMD-GRU 模型预测流程

本文采用含注意力机制的 EMD-GRU 模型,EMD 部分对数据进行分解,把非平稳性数据分解成多个平稳性分量;GRU 对负荷数据进行多输入多输出短期预测,同时使用 Attention 机制进行历史数据学习,选择负荷数据的历史时序状态特征,增加网络对历史时序特征的关注。对此预测流程,可做分述如下。

(1) 先对原始数据进行归一化,使数据分布在 $[0,1]$ 区间,以便后续模型的训练。

(2) 对归一化的数据进行 EMD 分解,分解成有限个不同频率的 IMF 分量和一个趋势分量。

(3) 本文在大量的历史负荷数据的基础上,同时考虑到每个分量之间的相关性,对所有分量建立了 2 层 GRU 神经网络模型。为了深入挖掘历史负荷曲线特征,在 GRU 模型中加入 Attention 机制进行历史数据的时序状态特征学习,获取历史时序状态权重,为预测提供关键信息。

(4) 根据得到的每个 IMF 分量预测数据进行重构获得最终的负荷预测结果。

含注意力机制的 EMD-GRU 预测模型流程图如图 4 所示。

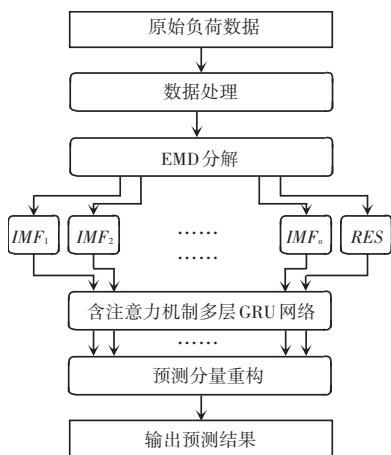


图 4 含注意力机制的 EMD-GRU 模型流程图

Fig. 4 Flow chart of EMD-GRU model with Attention mechanism

2.4 评价指标

本文选取均方根误差 (R_{MSE})、平均绝对百分误差 (M_{APE}) 来评价模型的预测准确性。其公式分别为:

$$M_{APE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\%, \quad (13)$$

$$R_{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (14)$$

其中, y_i 、 \hat{y}_i 、 n 分别为实际值、预测值、数据集长度。

3 算例分析

选取贵州某地电网 2019 年 10 月 27 日至 2020 年 10 月 25 日电力负荷数据作为数据集,采样时间间隔为 10 min,总计 364 天,有 52 560 个数据。同时划分 2019 年 10 月 27 日至 2020 年 10 月 18 日的

12 个月数据为训练集,用于模型的训练,调整模型的权重。以 2020 年 10 月 19 ~ 25 日一个星期数据作为测试集,用于评估模型的预测精度。基于此则在同一数据集下,分别采用 GRU、EMD-GRU 和含注意力机制的 EMD-GRU 三个模型进行负荷预测,并对 3 个模型进行评价。

3.1 数据展示

算例分析使用的 2019 年 10 月 27 日到 2020 年 10 月 25 日实际负荷数据如图 5 所示。从图 5 中可以看出,该地区 2020 年负荷基本稳定在均值 800 MW。在一年的负荷趋势中,11 月和 12 月具有最大值,尖峰性较为明显;在 2 月中达到最小值,这是由于 2 月处于春节假期,并且 2020 年由于疫情影响,企业还未复工复产,因此总体负荷偏小;到 3 月以后疫情控制有利,企业逐渐开始有序复工复产,总体负荷逐渐增加;4~10 月期间总体负荷较为稳定。

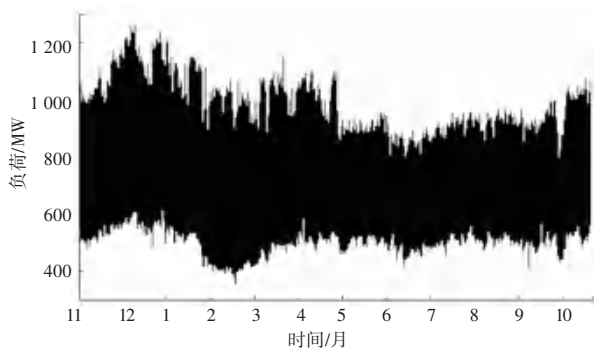


图 5 原始电力负荷数据图

Fig. 5 Raw power load data graph

3.2 EMD 分解结果

负荷数据经过 EMD 分解以后得到 13 个 IMF 分量和一个趋势 RES 分量,各分量如图 6 所示。从图 6 中可以看出每个 IMF 分量频率都是从高频到低频分布,同时高频分量的周期性比较明显,而低频分量具有一定的趋势性。在 RES 分量中,分量变化趋势从高到低,这与原始负荷数据整体负荷从高到低变化趋势相符。

3.3 含注意力机制的 EMD-GRU 模型预测结果

为了更好地体现出模型预测的效果,把作为比较的 GRU 和不含注意力机制的 EMD-GRU 模型预测结果也进行叠加得到不同模型的预测曲线图,其中含有注意力机制的 EMD-GRU 模型缩写为 AT-EMD-GRU。3 个模型 6 月 1 日一天的预测结果见图 7。由图 7 中可以看出,含有注意力机制的 EMD-GRU 模型预测的结果与实际值更接近,更能拟合真实值,具有更好的预测精度。

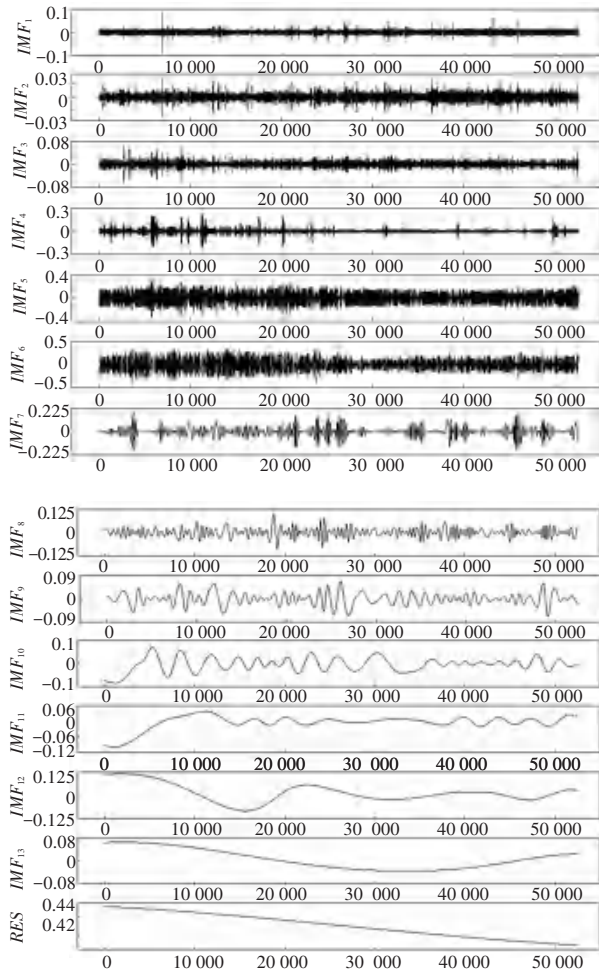


图6 负荷数据 EMD 分解结果图

Fig. 6 EMD decomposition result diagram of load data

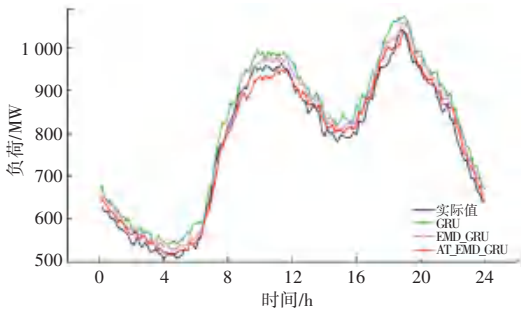


图7 不同模型一天预测结果曲线图

Fig. 7 Curves of one-day forecast results of different models

不同模型预测七天的指标对比结果见表1。从指标对比可以看出,含有注意力机制的 EMD-GRU 模型(AT-EMD-GRU)预测误差比其他模型更小。在7日均误差中,含注意力机制的 EMD-GRU 模型在 R_{MSE} 、 M_{APE} 误差指标中相对于 GRU 模型分别降低了 2.64 MW 和 0.22%;相对于 EMD-GRU 模型分别降低了 1.37 MW 和 0.05%。同时从每日的误差指标中也可以看出,使用 EMD-GRU 模型预测的误差要比直接使用 GRU 预测的误差有很大降低,加上注意

力机制的 EMD-GRU 模型预测误差又比 EMD-GRU 模型预测误差有了小幅的降低,这表明注意力机制在重要时间特征上发挥了作用。

表1 不同模型预测误差

Tab. 1 Forecast errors of different models

预测日期	指标	GRU	EMD-GRU	AT-EMD-GRU
星期一	R_{MSE} / MW	18.80	16.56	16.02
	M_{APE} / %	1.94	1.69	1.66
星期二	R_{MSE} / MW	19.67	16.78	16.35
	M_{APE} / %	1.96	1.72	1.67
星期三	R_{MSE} / MW	19.28	16.69	16.57
	M_{APE} / %	1.89	1.68	1.64
星期四	R_{MSE} / MW	17.94	16.62	16.12
	M_{APE} / %	1.81	1.74	1.73
星期五	R_{MSE} / MW	18.88	16.71	16.64
	M_{APE} / %	1.82	1.71	1.66
星期六	R_{MSE} / MW	17.88	14.12	13.64
	M_{APE} / %	1.82	1.45	1.34
星期日	R_{MSE} / MW	20.45	19.32	18.89
	M_{APE} / %	2.04	2.06	2.07
七日均误差	R_{MSE} / MW	19.20	17.53	16.16
	M_{APE} / %	1.90	1.73	1.68

图8为含注意力机制的 EMD-GRU 模型预测一个星期的误差。从图8中可以看出,一周内最大 R_{MSE} 为 18.89 MW,最小 R_{MSE} 为 13.6 MW,最大 M_{APE} 为 2.07%,最小 M_{APE} 为 1.3%。最大误差出现在星期天,最小误差出现在星期六。另外,周一至周五预测误差相差不大,而周末预测误差相对较大。这与实际负荷变化中工作日负荷变化规律性较强而休息日负荷波动较大的情况相吻合。

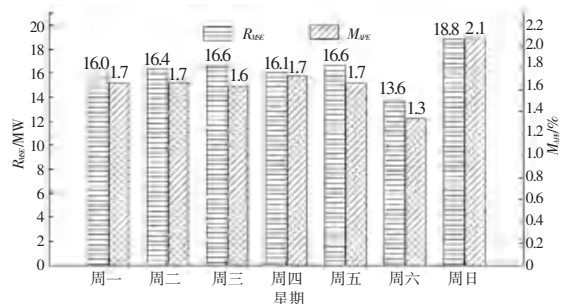


图8 含注意力机制的 EMD-GRU 模型预测七日误差柱状图

Fig. 8 Histogram of EMD-GRU model with Attention mechanism seven-day error prediction

4 结束语

为了有效处理非平稳性负荷数据以及进一步提
(下转第 43 页)