

DOI: 10.19783/j.cnki.pspc.231329

## 考虑天气特征与多变量相关性的配电网短期负荷预测

于越<sup>1</sup>, 葛磊蛟<sup>2</sup>, 金朝阳<sup>1</sup>, 王玥<sup>1</sup>, 丁磊<sup>1</sup>

(1. 电网智能化调度与控制教育部重点实验室(山东大学), 山东 济南 250061;

2. 智能电网教育部重点实验室(天津大学), 天津 300072)

**摘要:** 针对配电网短期负荷预测受到众多复杂天气特征等随机不确定性因素影响, 以及传统预测模型难以有效分析不同特征序列之间的相关性等问题, 提出一种考虑天气特征与多变量相关性的配电网短期负荷预测方法。首先, 提出多变量快速最大信息系数(multi-variable rapid maximal information coefficient, MVRapidMIC)提取相关性高的天气特征序列。其次, 引入探索性因子分析法(exploratory factor analysis, EFA), 对高相关性特征序列进行降维处理。最后, 将维度分段(dimension-segment-wise, DSW)机制和两阶段注意力(two-stage attention, TSA)机制与 Informer 模型结合, 提高预测模型对不同特征序列相关性的分析能力。通过 DTU 7K 47 节点实际配电网的历史负荷数据开展仿真测试, 验证所提方法的预测精度、鲁棒性和时效性。

**关键词:** 配电网; 短期负荷预测; 天气特征; 最大信息系数; Informer 框架

### Short-term load prediction method of distribution networks considering weather features and multivariate correlations

YU Yue<sup>1</sup>, GE Leijiao<sup>2</sup>, JIN Zhaoyang<sup>1</sup>, WANG Yue<sup>1</sup>, DING Lei<sup>1</sup>

(1. Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University), Ministry of Education, Jinan 250061, China; 2. Key Laboratory of Smart Grid of Ministry of Education (Tianjin University), Tianjin 300072, China)

**Abstract:** To address challenges in short-term load forecasting for distribution networks, challenges such as the impact of complex weather features and the difficulty in analyzing correlations between different feature sequences using traditional models, a method considering those issues is proposed. First, the method presents a multi-variable rapid maximal information coefficient (MVRapidMIC) to extract highly correlated weather feature sequences. Exploratory factor analysis (EFA) is then employed for dimensionality reduction on these sequences. Finally, the dimension-segment-wise (DSW) and two-stage attention (TSA) mechanisms are integrated with the Informer model to enhance the model's ability to analyze correlations between different feature sequences. Simulation tests using historical load data from the DTU 7K 47-bus distribution system validate the forecasting accuracy, robustness, and timeliness of the method.

This work is supported by the National Natural Science Foundation of China (No. U22B20101).

**Key words:** distribution network; short-term electricity load forecasting; weather features; maximal information coefficient; Informer framework

## 0 引言

负荷预测是提高电网调度水平的重要部分<sup>[1]</sup>, 对配电网进行快速、准确地短期负荷预测有利于保障电网安全稳定运行、降低运营成本以及提高能源的利用率<sup>[2]</sup>。

配电网中存在着海量且种类繁多的主动时变敏感负荷<sup>[3]</sup>, 它们的功率变化均与天气特征变化密切相关<sup>[4]</sup>。这些特征的不确定性导致短期配电网负荷序列具有波动性大、随机性强等特点<sup>[5]</sup>, 给短期负荷预测的精度、鲁棒性以及运算效率带来巨大挑战, 难以满足新型电力系统的运行与控制需求<sup>[6]</sup>。为了解决这个问题, 专家学者对特征提取和降维技术进行研究。在特征提取方面, 常用的算法有皮尔逊系数、斯皮尔曼系数和最大信息系数(maximum information

基金项目: 国家自然科学基金项目资助(U22B20101)

coefficient, MIC)<sup>[7]</sup>等。以上模型虽然能够准确评估两个变量之间的相关性,但运算速度较慢。文献[8]则提出了一种两变量快速最大信息系数(RapidMIC),以提高两变量MIC的运算速度。但某些天气特征(如累积温度效应)是由多个特征变量组成,包括预测日和历史相似日不同时刻的温度。这些不同组成部分之间存在着较强的时序和空间相关性,在进行预测时会产生一定的共线性问题<sup>[9]</sup>。然而,现有的两变量RapidMIC算法仅能分别考虑单个特征部分与负荷序列之间的相关性,却忽视了不同组成部分之间的时序和空间相关性,导致特征之间的信息重叠度较高,降低了模型的预测能力。因此,需要引入一种可以同时考虑多自变量与单因变量相关性的快速分析方法,对复杂特征序列进行深度提取。在特征降维方面,常用的算法有主成分分析法(principal component analysis, PCA)<sup>[10]</sup>、独立成分分析法(independent component analysis, ICA)<sup>[11]</sup>和因子分析法(factor analysis, FA)<sup>[12]</sup>等。其中,FA将不同特征之间的相关性建模为公共因子,可以在降低维度的同时,准确地捕捉数据中的重要信息。但是,FA的公共因子通常以线性组合的形式呈现,容易出现冗余,从而对预测性能造成负面影响。

为进一步提高预测性能,国内外专家学者对考虑天气特征的配电网短期负荷预测模型进行了大量研究,主要分为数学统计方法和人工智能方法。由于人工智能方法具有强大的数据分析和预测能力<sup>[13]</sup>,已成为短期负荷预测领域未来发展的重要方向。典型的人工智能方法有循环神经网络(recurrent neural network, RNN)<sup>[14]</sup>、长短期记忆网络(long short-term memory, LSTM)<sup>[15]</sup>和门控循环单元(gated recurrent unit, GRU)<sup>[16]</sup>等,以上模型虽然可以对时间序列进行较为准确的预测,但是它们依赖于逐步迭代的循环过程,在处理长时间序列时容易丢失特征信息<sup>[17]</sup>。为此,Facebook开发了Prophet预测模型,其无须依赖复杂的循环结构,即可捕捉长期时间依赖性<sup>[18]</sup>。但是,Prophet是基于加法模型的预测模型,对复杂的季节性模式和非线性趋势没有足够的灵活性<sup>[19]</sup>。而基于自注意力机制的Transformer模型采用多层神经网络结构,具有出色的长时间序列预测能力,且能够更好地处理复杂的季节性模式和非线性趋势<sup>[20]</sup>。在Transformer的基础上,文献[21]提出了Informer预测模型,该模型使用了多头概率稀疏自注意力(multi-head probsparse attention, MPSA)和蒸馏(distilling)机制<sup>[22]</sup>,显著降低了计算复杂度并提高了时序预测的精度,同时增强了对输入和输出之间的远程关联性的捕捉能力<sup>[23]</sup>。此外,Informer会将

所有特征序列的数据点同时嵌入到一个特征向量中,并提取不同时间序列之间的相关性。但是,不同特征序列之间的相关性尚未被Informer的机制充分考虑,这会限制其预测能力<sup>[24]</sup>。

基于此,本文针对配电网的特征序列具有高复杂性和高维度等特点,提出一种考虑多变量相关性的改进Informer配电网短期负荷预测模型。首先,提出利用多变量快速最大信息系数(multi-variable RapidMIC, MVRapidMIC),对众多天气特征与配电网负荷序列进行相关性分析,使预测模型更好地适应多维度的输入数据;其次,采用探索性因子分析法(exploratory FA, EFA)<sup>[25]</sup>对提取后的高相关性天气特征序列进行降维处理;最后,将维度分段(dimension-segment-wise, DSW)嵌入机制和两阶段注意力(two-stage attention, TSA)机制<sup>[24]</sup>与Informer模型结合,提高预测模型对不同特征序列相关性的分析能力。利用配电网历史负荷数据集进行离线训练,建立基于MVRapidMIC-EFA的改进Informer预测模型。通过基于DTU 7K 47节点真实配电系统<sup>[26]</sup>的仿真分析,验证了本文方法的精度、鲁棒性和时效性。

## 1 特征处理方法

### 1.1 基于MVRapidMIC的特征提取方法

#### 1.1.1 RapidMIC

为分析配电网短期负荷序列与复杂特征序列之间的相关性,需要引入一种可以同时考虑多特征序列的相关性分析算法。

MIC是一种由Reshef在2011年的Science上提出的相关性分析方法<sup>[7]</sup>。其基本原理是在两个变量构成的散点图上划分多个网格,从而封装变量之间的关系。然而,在处理大规模数据集时,MIC会出现计算时间过长等缺点。为此,文献[8]提出了一种名为RapidMIC的相关性算法。该方法采用POSIX并行策略,通过对两个变量进行动态规划和均衡化,将复杂问题分解为多个简单任务进行并行处理,从而提高了计算效率。

#### 1.1.2 MVRapidMIC

为了使现有的两变量RapidMIC算法具有计算多自变量(例如累积温度效应)与单因变量之间相关性的能力。本文将多变量MIC<sup>[27]</sup>的概念引入到两变量RapidMIC算法中,提出一种MVRapidMIC算法。

MVRapidMIC的具体定义如下:设定多变量数据集 $\sigma = \{L, (H_1, H_2, \dots, H_n)\}$ ,  $(H_1, H_2, \dots, H_n)$ 为区域天气特征(自变量), $L$ 为配电网负荷序列(因变量)。采用POSIX并行策略,分别将 $(H_1, H_2, \dots, H_n)$

划分到  $X$  轴,  $L$  划分到  $Y$  轴。通过对  $X$ 、 $Y$  轴进行网格划分, 每个网格表示数据点的一个子集。将落入相应网格中的点的数目与总点数的比例定义为该网格的近似概率密度, 进而可以得到  $(H_1, H_2, \dots, H_n)$  与  $L$  之间的互信息  $I(L; H_1, H_2, \dots, H_n)$ ; 之后进行归一化处理, 从而得到在数据集  $\sigma$  下,  $(H_1, H_2, \dots, H_n)$  与  $L$  之间的 MVRapidMIC 值。其求取步骤如下。

1) 对于给定多变量数据集  $\sigma$  和正整数  $r, s$ , 其中  $r, s \geq 2$ ,  $\sigma = \{L, (H_1, H_2, \dots, H_n)\}$ , 当  $L$  和  $(H_1, H_2, \dots, H_n)$  分别被网格划分为  $r, s$  块时, 本文定义自变量  $(H_1, H_2, \dots, H_n)$  和因变量  $L$  之间的互信息为

$$I(\sigma, r, s, L, (H_1, H_2, \dots, H_n)) = \sum_{i=1}^r \sum_{j=1}^s P(ij) \log_2 \frac{P(ij)}{P_{\text{ind}}(i)P_{\text{sd}}(j)} \quad (1)$$

式中:  $P_{\text{ind}}(i)$  表示  $L$  被划分在  $r$  个网格下的边缘概率密度;  $P_{\text{sd}}(j)$  表示  $(H_1, H_2, \dots, H_n)$  在  $s$  个网格划分下的边缘概率密度;  $P(ij)$  为数据集  $\sigma$  被划分在  $r \times s$  个网格下的联合概率密度。

2) 本文定义  $L$  和  $(H_1, H_2, \dots, H_n)$  之间的 MVRapidMIC 值为

$$\text{MVRapidMIC}(L; (H_1, H_2, \dots, H_n)) = \max_{r, s < G(n)} \frac{I(\sigma, r, s, L; (H_1, H_2, \dots, H_n))}{\log_2(\min(r, s))} \quad (2)$$

式中:  $I(\sigma, r, s, L; (H_1, H_2, \dots, H_n))$  为三者之间的互信息;  $r, s < G(n)$ ,  $G(n)$  为最大网格划分数目;  $\log_2(\min(r, s))$  为  $r, s$  的归一化处理方式。其中, MVRapidMIC 值越大, 表示相应的区域天气特征变量与配电网负荷序列的相关性越强。该算法通过计算多自变量与单因变量之间的 MVRapidMIC 值, 提高模型对多维信息的感知能力, 进而提升预测精度和运算效率。

## 1.2 EFA

为减少配电网负荷预测模型的输入特征维数, 本文使用 EFA 算法对输入的天气特征序列进行降维处理。EFA 是一种降维方法<sup>[25]</sup>, 用于将众多特征合并为几个代表性的公共因子, 以提取特征之间的潜在因素, 其主要步骤如下。

1) 通过计算各天气特征序列的特征值大小和累计方差贡献率(cumulative variance contribution rate, CVCR)等指标, 以确定公共因子数量。其中,  $C_{\text{VCR}}$  为前  $\phi$  个公共因子对总方差贡献的累积比例, 在因子分析中用于评估数据的变异程度。当公共因子的  $C_{\text{VCR}}$  高于 85% 时, 视为可以较好地解释数据变化。 $C_{\text{VCR}}$  的计算公式为

$$C_{\text{VCR}} = \frac{\sum \tau_\phi}{\sum \tau} \quad (3)$$

式中,  $\tau$  表示每个公共因子的特征值。

2) 将观测变量(区域天气特征变量)转化为相关矩阵, 分析变量之间的相关性, 相关矩阵  $\mathbf{R}$  表示为

$$\mathbf{R} = \text{cor}(\boldsymbol{\kappa}) \quad (4)$$

式中,  $\boldsymbol{\kappa}$  为包含观测变量的矩阵。

3) 从相关矩阵中提取公共因子, 通过正交旋转法进行因子旋转, 并计算因子载荷矩阵  $\boldsymbol{\lambda}$ 。因子载荷矩阵  $\boldsymbol{\lambda}$  由各个变量与各个因子之间的相关系数构成, 表示为

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1F} \\ \lambda_{21} & \dots & \lambda_{2F} \\ \vdots & & \vdots \\ \lambda_{E1} & \dots & \lambda_{EF} \end{bmatrix} \quad (5)$$

式中:  $E$  为观测变量的数量;  $F$  为因子数量;  $\lambda_{EF}$  为第  $E$  个观测变量与第  $F$  个因子之间的载荷系数。因子载荷矩阵中的值越大, 表示观测变量与对应因子之间的相关性越强。

4) 计算因子得分矩阵。因子得分矩阵是观测变量映射到公共因子空间的结果。它表示每个变量在每个公共因子上的得分, 并使用回归法进行计算, 如式(6)所示。

$$\boldsymbol{\delta}_{\text{score-EFA}} = \boldsymbol{\kappa}(\boldsymbol{\lambda}^T \boldsymbol{\lambda})^{-1} \boldsymbol{\lambda}^T \boldsymbol{\kappa}^T \quad (6)$$

式中,  $\boldsymbol{\delta}_{\text{score-EFA}}$  为因子得分矩阵。因子得分矩阵中的元素值越大, 表示该因子所代表的特征在数据集中的表现越显著。

## 2 改进的 Informer 预测模型

### 2.1 Informer 预测模型

Informer 主要由改进的编码器和解码器等模块构成。相较于以往的 Transformer 模型, Informer 在其编码器层中采用了 MPSA 机制和蒸馏机制, 以替代传统的自注意力机制。

#### 2.1.1 编码器

传统自注意力机制主要由 Query、Key 和 Value 组成, 通过比较 Key 与 Query 的相似程度, 赋予 Key 相应的 Value。传统自注意力机制公式为

$$A(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Soft max} \left( \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}} \right) \mathbf{V} \quad (7)$$

式中:  $\mathbf{Q} \in \mathbb{R}^{\ell_q \times d}$ ;  $\mathbf{K} \in \mathbb{R}^{\ell_k \times d}$ ;  $\mathbf{V} \in \mathbb{R}^{\ell_v \times d}$ ;  $d$  为输入维度。

在编码器部分, Informer 引入了 MPSA 机制。令 Key 只关注重要性更强的 Query, MPSA 公式为

$$A(\bar{Q}, K, V) = \text{Soft max} \left\{ \frac{\bar{Q}K^T}{\sqrt{d}} \right\} V \quad (8)$$

式中,  $\bar{Q}$  为通过  $Q$  的概率稀疏得到的矩阵, 其使计算复杂度由  $O(\mathcal{G}^2)$  降到了  $O(\mathcal{G} \ln \mathcal{G})$ , 有效地提高了运算效率。

为进一步减少计算时间, Informer 引入了蒸馏机制, 逐次将序列长度减半。对第  $\varphi$  层序列  $\mathbf{x}'_{\varphi}$  进行蒸馏操作, 得到第  $\varphi+1$  层序列  $\mathbf{x}'_{\varphi+1}$ 。其过程如式(9)所示。

$$\mathbf{x}'_{\varphi+1} = \text{MaxPool}(\text{ELU}(\text{Convld}([\mathbf{x}'_{\varphi}]_{\text{AB}}))) \quad (9)$$

式中: Max-Pool( $\cdot$ )为最大池化层; ELU( $\cdot$ )为激活函数; Convld( $\cdot$ )为一维卷积操作;  $[\cdot]_{\text{AB}}$  为多头注意力机制。

### 2.1.2 解码器

在解码器部分, Informer 引入了生成式译码方法, 能够快速生成所有的预测结果。生成式译码的方式如式(10)所示。

$$\mathbf{x}'_{\text{feed\_de}} = \text{Concat}(\mathbf{x}'_{\text{token}}, \mathbf{x}'_0) \in \mathbb{R}^{(\mathcal{G}_{\text{token}} + \mathcal{G}_y) \times d} \quad (10)$$

式中: Concat( $\cdot$ )表示拼接操作;  $\mathbf{x}'_{\text{feed\_de}}$  为输入的时间序列;  $\mathbf{x}'_{\text{token}}$  为目标序列之前部分已知序列;  $\mathbf{x}'_0$  为要一次性预测输出的目标序列, 初始值设为 0;  $\mathcal{G}_{\text{token}}$ 、 $\mathcal{G}_y$  分别为  $\mathbf{x}'_{\text{token}}$  和  $\mathbf{x}'_0$  的长度。

### 2.1.3 Informer 预测模型存在的问题

虽然 Informer 模型具有较高的计算效率, 但是由于自身结构的限制, 多特征序列之间的相关性未被充分考虑, 这会影响 Informer 的预测能力。

## 2.2 基于 DSW 嵌入和 TSA 机制的 Informer 模型

为更有效地捕获不同特征序列(或称之为“特征维度”)之间的相关性, 本文将 DSW 嵌入机制和 TSA 机制<sup>[24]</sup>引入到 Informer 模型中, 提出一种改进的 Informer 预测模型(以下简称为“DSWInformer”)

### 2.2.1 DSW 嵌入机制

DSW 嵌入机制的实现方式为: 将每个特征维度的数据分成指定长度的段, 每段的长度用  $\mathcal{G}_{\text{seq}}$  表示, 然后进行嵌入操作, 其公式为

$$\mathbf{x}'_{1:T} = \left\{ \mathbf{x}'_{\theta,d} \mid 1 \leq \theta \leq \frac{T}{\mathcal{G}_{\text{seq}}}, 1 \leq d \leq D \right\} \quad (11)$$

式中:  $\mathbf{x}'_{1:T}$  为归一化后的输入序列;  $T$  为时间滑动窗口;  $\mathbf{x}'_{\theta,d} \in \mathbb{R}^{\mathcal{G}_{\text{seq}}}$  表示特征维度  $d$  中时间步为  $\theta$  且长度为  $\mathcal{G}_{\text{seq}}$  的序列段;  $D$  为总输入维度数。

然后, 将每个分段嵌入到一个向量中, 同时融合位置信息, 其公式为

$$\mathbf{h}_{\theta,d} = \mathbf{E}\mathbf{x}'_{\theta,d} + \mathbf{E}_{\theta,d}^{(\text{pos})} \quad (12)$$

式中:  $\mathbf{h}_{\theta,d}$  表示在特征维度  $d$  上嵌入第  $\theta$  个时间步后得到的向量;  $\mathbf{E} \in \mathbb{R}^{d_{\text{target}} \times \mathcal{G}_{\text{seq}}}$  为一个可学习参数矩阵,  $d_{\text{target}}$  为向量嵌入后的特征维度;  $\mathbf{E}_{\theta,d}^{(\text{pos})}$  为用于位置嵌入的可学习参数矩阵。

通过 DSW, 得到一个二维数据集  $\zeta$ , 两个维度分别对应时间序列和特征序列, 其表达式为

$$\zeta = \left\{ \mathbf{h}_{\theta,d} \mid 1 \leq \theta \leq \frac{T}{\mathcal{G}_{\text{seq}}}, 1 \leq d \leq D \right\} \quad (13)$$

### 2.2.2 TSA 机制

TSA 机制分为交叉时间阶段和交叉特征阶段。

1) 交叉时间阶段。给定一个二维向量  $\mathbf{Z} \in \mathbb{R}^{(T/\mathcal{G}_{\text{seq}}) \times D \times d_{\text{target}}}$  作为 TSA 层的输入(此处  $\mathbf{Z}$  为 DSW 嵌入层的输出)。在每个特征维度上沿着时间序列应用 MPSA 机制, 以捕捉不同时间序列之间的相关性。在这个阶段, 时间复杂度为  $O(D^2(T^2/\mathcal{G}_{\text{seq}}^2))$ , 其主要由每个维度中的段数决定。具体表示为

$$\hat{\mathbf{Z}}'_{:,d} = \text{LayerNorm}(\mathbf{Z}_{:,d} + M'_{\text{PSA}}(\bar{\mathbf{Z}}_{:,d}, \mathbf{Z}_{:,d}, \mathbf{Z}_{:,d})) \quad (14)$$

$$\mathbf{Z}' = \text{LayerNorm}(\hat{\mathbf{Z}}'_{:,d} + \text{MLP}(\mathbf{Z}^{t-1})) \quad (15)$$

式中:  $\mathbf{Z}_{:,d}$  为维度  $d$  中所有时间步的向量;  $\bar{\mathbf{Z}}_{:,d}$  为维度  $d$  中所有时间步的随机向量; LayerNorm 为层归一化操作; MLP 为双层前馈网络;  $M'_{\text{PSA}}$  为时间序列  $t$  时刻的 MPSA 层操作;  $\hat{\mathbf{Z}}'_{:,d}$  为 MPSA 层  $t$  时刻的输出;  $\mathbf{Z}'$ 、 $\mathbf{Z}^{t-1}$  分别为 MLP 层  $t$  和  $t-1$  时刻的输出。

2) 交叉特征阶段。为捕捉不同特征序列之间的相关性, 引入一种中间路由机制: 在每个时间步  $\theta$ , 设置一个固定维数的可学习向量作为中间路由向量  $\boldsymbol{\eta}_{\theta,:}$ , 用于汇总每个维度  $d$  的信息。然后, 从中间路由中分离信息, 以捕捉各个特征序列之间的相关性, 得到  $\mathbf{Z}_{\theta,:}^d$ , 具体表示为

$$\boldsymbol{\eta}_{\theta,:} = M_{\text{PSA1}}^d(\mathbf{R}_{\theta,:}, \mathbf{Z}_{\theta,:}, \mathbf{Z}_{\theta,:}), 1 \leq \theta \leq \frac{T}{\mathcal{G}_{\text{seq}}} \quad (16)$$

$$\bar{\mathbf{Z}}_{\theta,:}^d = M_{\text{PSA2}}^d(\mathbf{Z}_{\theta,:}, \boldsymbol{\eta}_{\theta,:}, \boldsymbol{\eta}_{\theta,:}), 1 \leq \theta \leq \frac{T}{\mathcal{G}_{\text{seq}}} \quad (17)$$

式中:  $\mathbf{R}_{\theta,:} \in \mathbb{R}^{(T/\mathcal{G}_{\text{seq}}) \times C \times d_{\text{target}}}$  为中间路由的可学习参数矩阵;  $\mathbf{Z}_{\theta,:}$  为时间步  $\theta$  的二维向量;  $M_{\text{PSA1}}^d$ 、 $M_{\text{PSA2}}^d$  分别为特征序列的两层 MPSA 操作;  $\bar{\mathbf{Z}}_{\theta,:}^d$  为时间步  $\theta$  时中间路由机制的输出。

在 TSA 机制之后, DWSInformer 能够同时捕获跨时间序列和跨特征序列的相关性, 并将时间复杂度降低  $O(2cD(T/\mathcal{G}_{\text{seq}}))$ 。DWSInformer 预测模型的结构如图 1 所示。

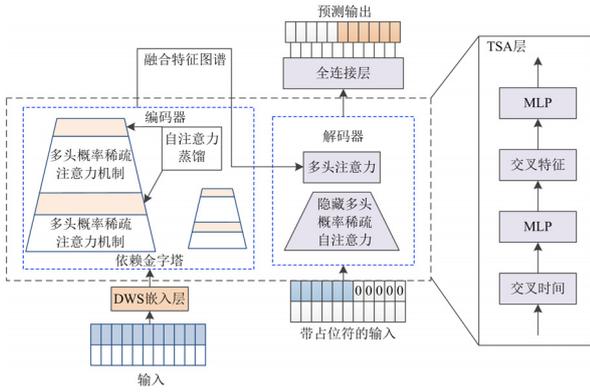


图1 DWSInformer 结构示意图

Fig. 1 Structure of DWSInformer

在本文中, DWSInformer 模型的最终短期负荷预测方程为

$$v_k = f_{\text{DWSInformer}}(v_{k-1}, v_{k-2}, \dots, v_{k-\rho}) \quad (18)$$

式中:  $f_{\text{DWSInformer}}$  为 DWSInformer 负荷预测模型;  $v_k$  为  $k$  时刻的负荷值; 变量  $\rho$  为时间步的历史窗口大小, 表示用于预测当前时间步的输入  $v_k$  的过去  $\rho$  个时间步的历史信息。配电网负荷预测值可由式(18)得到。

### 3 整体预测模型的训练流程

对于配电网短期负荷预测, 本文利用以 15 min 为采样间隔的历史数据集, 预测未来 1 h 内的配电网负荷序列, 构建输出层步长为 4 的 Informer 模型。在预测模型输入特征的选取上, 由于数值天气预报的时间分辨率较长且精度不足, 故仅将历史气象数据作为天气特征输入, 并将其与配电网历史负荷序列结合, 共同构成输入矩阵。MVRapidMIC-EFA 的 DWSInformer 预测模型的具体预测步骤如下。

- 1) 数据补全: 对于配电网负荷数据集中的缺失, 本文选取采样点前后 6 个数据的平均值进行补全。
- 2) 归一化: 本文使用 Z-score 算法进行归一化处理, 以确保数据分布均衡。
- 3) 时间戳对齐: 本文将使用样条插值法进行数据集的时间戳对齐操作。
- 4) 特征筛选: 通过对天气、历史和时间特征与配电网负荷序列之间的相关性进行分析, 提取与预测任务密切相关的特征。设定预测时间为  $k$ , 输入特征序列包括与天气特征相关的 5 个具有相似天气条件的历史负荷、近 5 天的历史负荷, 以及温度、累积温度效应、大气压、风速、风向、相对湿度、降水量和时间。输入特征序列的定义如表 1 所示。

对特征序列  $H_1 - H_5$ 、 $H_6 - H_{10}$  和  $H_{12} - H_{21}$  采用 MVRapidMIC 算法的多变量模式进行相关性分析, 其他特征则采用 MVRapidMIC 算法的单变量模式。

5) 特征降维: 在保留原始特征基本信息的基础上, 采用 EFA 对提取的特征序列进行降维处理。

6) 离线训练: 将处理好的特征与配电网负荷序列构建 DWSInformer 输入矩阵。根据 DWSInformer 模型的训练方法, 将处理后的数据集按照 7:2:1 的比例划分为训练集、验证集和测试集。对 DWSInformer 模型进行离线训练, 得到预测模型。

表 1 输入特征序列的定义

Table 1 Definition of input feature sequences

输入特征序列	变量索引	输入特征序列	变量索引
5 个具有相似天气条件的历史负荷	$H_1, H_2, H_3, H_4, H_5$	风速	$H_{23}$
近 5 天的历史负荷	$H_6, H_7, H_8, H_9, H_{10}$	风向	$H_{24}$
温度	$H_{11}$	相对湿度	$H_{25}$
累积温度效应(选择预测日的前 5 天, 分别取它们 $k$ 时刻的历史温度以及预测日当天 $k-2$ 、 $k-4$ 、 $k-6$ 、 $k-8$ 和 $k-10$ 时刻的温度)	$H_{12}, H_{13}, H_{14}, H_{15}, H_{16}, H_{17}, H_{18}, H_{19}, H_{20}, H_{21}$	降水量	$H_{26}$
大气压	$H_{22}$	时间	$H_{27}$

本文所提预测模型的训练流程如图 2 所示。

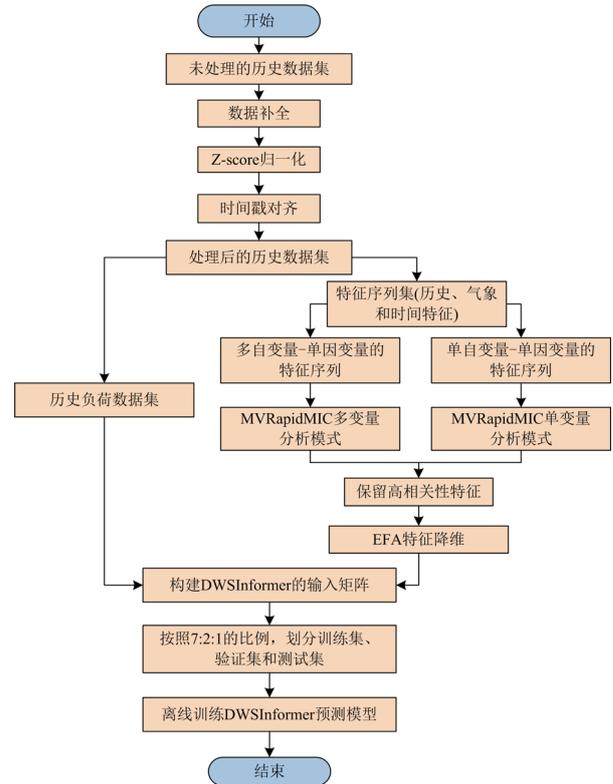


图 2 基于 MVRapidMIC-EFA 的 DWSInformer 预测模型的训练流程图

Fig. 2 Training flow chart of the DWSInformer prediction model based on MVRapidMIC-EFA

## 4 算例分析

### 4.1 仿真设定

本文采用 DTU 7K 47 节点配电网进行仿真测试, 其结构如图 3 所示。DTU 7K 47 节点配电网的历史负荷数据集来自官方提供, 本文使用 2014 年 1 月 1 日到 2015 年 1 月 1 日的负荷数据, 采样频率为 15 min, 共 35 040 个样本。随机选取 24 528 个样本作为训练集, 7008 个样本作为验证集, 3504 个样本作为测试集, 输入时间序列长度设为 96。

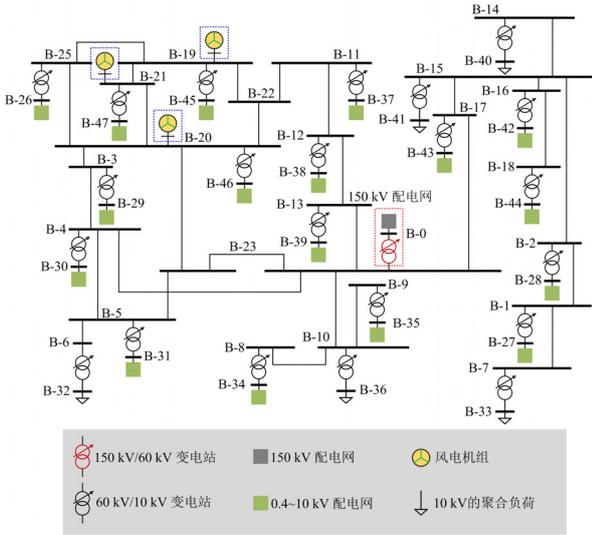


图 3 DTU 7K 47 节点配电网

Fig. 3 DTU 7K 47-bus distribution system

由图 3 可以看出: 该系统有 3 个风电场, 由第四代可控风机组成, 装机容量分别为 12 MW、15 MW 和 15 MW。因为 DTU 7K 47 节点配电网位于丹麦境内, 所以气象数据来源于丹麦气象局。所选取的天气特征包括温度、累积温度效应、大气压、风速、风向、相对湿度以及降水量。具体的预测时间为 2014 年的任意一天, 设为  $k$ 。

本文模型均使用 Python 软件进行训练。DWSInformer 预测模型选取 Pytorch 工具箱进行训练, 其参数设置见表 2。

计算机配置为 Win10 系统, 处理器为 Intel Core i7-8700k 系列, 主频为 3.0 GHz, 内存为 16 G。显卡为英伟达 GTX-1660TI, 显存为 4 GB。本文采用均方根误差(root mean squared error, RMSE)和平均绝对误差(mean absolute error, MAE)评估本文所提方法的性能。

$$v_{\text{RMSE}} = \left( \sum_{k=1}^{\gamma} (\hat{v} - v)^2 / \gamma \right)^{\frac{1}{2}} \quad (19)$$

$$v_{\text{MAE}} = \left( \sum_{k=1}^{\gamma} |\hat{v} - v| \right) / \gamma \quad (20)$$

式中:  $v_{\text{RMSE}}$  为负荷的 RMSE;  $v_{\text{MAE}}$  为负荷的 MAE;  $v$  为负荷真实值;  $\hat{v}$  为负荷预测值;  $\gamma$  为测试集中元素的数量。

表 2 DWSInformer 预测模型的参数设置

Table 2 Parameter settings for the DWSInformer prediction model

参数名称	参数设置	参数名称	参数设置
features	MS	e_layers	2
freq	15min	d_layers	1
seq_len	96	s_layers	3,2,1
label_len	48	enc_in	3
pred_len	4	dec_in	3
d_model	512	c_out	3
n_heads	8	d_ff	2048
factor	5	dropout	0.05
padding	0	attn	prob
distil	True	embed	timeF
activation	gelu	itr	3
output_attention	False	train_epochs	100
do_predict	False	learning_rate	0.0001
mix	True	lradj	type1

### 4.2 仿真案例 1: 特征处理方法测试

#### 4.2.1 特征提取方法测试

选择 2014 年的输入特征序列(包括天气、历史和时间)与配电网负荷序列, 通过 MVRapidMIC 算法进行相关性分析, 具体结果如表 3 所示。

表 3 各输入特征序列(包括天气、历史和时间)与配电网负荷序列的 MVRapidMIC 值

Table 3 MVRapidMIC value of input feature sequences (including weather, historical and time) and distribution network load sequences

算法	输入特征序列	MVRapidMIC 值	时间/s
MVRapidMIC	$H_1-H_5$	0.76	1.8435
	$H_6-H_{10}$	0.72	
	$H_{11}$	0.88	
	$H_{12}-H_{21}$	<b>0.93</b>	
	$H_{22}$	0.21	
	$H_{23}$	0.89	
	$H_{24}$	0.29	
	$H_{25}$	0.42	
	$H_{26}$	0.81	
	$H_{27}$	0.69	

本仿真设定 RapidMIC 算法的置信度为 95%, 且表示 MVRapidMIC 值出现极端情况的概率  $p$  为 0.1%, 并参考两变量 RapidMIC 的阈值<sup>[8]</sup>, 设定当某个输入特征序列的 MVRapidMIC 值大于等于 0.51 时,

表明该输入特征序列和配电网负荷序列之间存在密切关系。

由表 3 可以看出: 累积温度效应、温度、降水量、风速、5 个具有相似天气条件的历史负荷以及近 5 天的历史负荷等特征与配电网负荷序列高度相关; 而大气压、风向和相对湿度的 MVRapidMIC 值均低于设定的阈值, 表明它们与配电网负荷序列之间的关联性较弱。因此, 在后续分析中不考虑大气压、风向和相对湿度等特征, 保留其余 24 个输入特征, 以确保预测模型的准确性和鲁棒性。

为证明本文所提特征提取算法的合理性, 将全部特征序列(all features, AF)、MIC 选取的特征序列(其 MIC 值见表 4, 运算时间为 7.9234 s)、RapidMIC 选取的特征序列(其 RapidMIC 值同见表 4, 运算时间为 1.9809 s)、相关性最大的特征序列——累积温度效应(cumulative temperature effect, CTE)分别作为 DWSInformer 预测模型的特征序列输入, 与 MVRapidMIC-DWSInformer 的预测结果进行比较(以

上对照组的预测步长均为 4), 结果如表 5 所示。其中, AF-DWSInformer 模型、MIC-DWSInformer 模型、RapidMIC-DWSInformer 模型、CTE-DWSInformer 模型和 MVRapidMIC-DWSInformer 模型的运行时间分别为 168.85 s、154.53 s、38.63 s、19.09 s 和 35.77 s。表 4 各输入特征序列(包括天气、历史和时间)与配电网负荷序列的 MIC/RapidMIC 值

Table 4 MIC/RapidMIC value of input feature sequences (including weather, historical and time) and distribution network load sequences

输入特征序列	MIC 值	RapidMIC 值
$H_1-H_5$	0.77	0.77
$H_6-H_{10}$	0.71	0.71
$H_{11}$	0.87	0.87
$H_{12}-H_{21}$	0.90	0.90
$H_{22}$	0.21	0.21
$H_{23}$	0.89	0.89
$H_{24}$	0.29	0.29
$H_{25}$	0.41	0.41

表 5 不同特征提取方法的负荷预测精度比较

Table 5 Comparison of load prediction accuracy by different feature extraction methods

预测模型	MW									
	春		夏		秋		冬		平均	
	RMSE	MAE								
AF-DWSInformer	9.461	8.312	9.453	8.306	9.722	8.647	9.518	8.473	9.538	8.435
MIC-DWSInformer	8.171	7.036	8.043	6.935	8.364	7.178	8.417	7.229	8.248	7.095
RapidMIC-DWSInformer	8.171	7.036	8.043	6.935	8.364	7.178	8.417	7.229	8.248	7.095
CTE-DWSInformer	8.256	7.147	8.285	7.159	11.963	10.492	12.375	11.451	10.219	9.062
MVRapidMIC-DWSInformer	<b>8.019</b>	<b>6.886</b>	<b>7.938</b>	<b>6.765</b>	<b>8.278</b>	<b>7.015</b>	<b>8.306</b>	<b>7.173</b>	<b>8.135</b>	<b>6.959</b>

由表 5 可以看出: 若将全部天气特征作为预测模型的特征输入, 信息冗余会导致预测精度降低以及训练时间过长。若使用 MIC 算法进行特征提取, 则可有效提高预测精度(RMSE 和 MAE 分别降低 15.64%、18.88%), 但是训练时间仍然较长。使用 RapidMIC 算法则可以在保证精度不变的情况下, 训练时间减少了 75%。若仅选择相关性最大的累积温度效应作为预测模型的特征输入, 虽然可以实现较好的预测精度和高效的训练效率, 但是其鲁棒性较差。这是因为丹麦为温带海洋性气候, 降水和大风集中在秋冬季, 且会出现暴风雪等极端天气。若仅考虑累积温度效应, 则不能全面反映天气特征对负荷预测的影响。因此, 该对照组在春夏季表现良好, 但在秋冬季的预测误差较大。

相比以上 3 种只能进行单自变量-单因变量分析的特征提取方法, 本文提出的 MVRapidMIC-DWSInformer 方法具有最好的预测精度和鲁棒性。

同时, 其训练时间相比 RapidMIC-DWSInformer 算法, 又减少了 7.40%, 具有较好的训练效率。

#### 4.2.2 特征降维方法测试

对各项输入特征序列进行特征降维, 可有效选取对配电网负荷预测重要的信息, 在确保预测精度的同时又不失鲁棒性。本文采用 EFA 算法, 对提取出的 24 个输入特征进行降维处理。得到的方差分析和因子得分矩阵结果见表 6 和表 7。

由表 6 和表 7 可以看出: 本文从输入特征中降维得到了天气、历史和时间等公共因子, 这 3 个公

表 6 方差分析

Table 6 Variance analysis

公共因子	特征值	方差贡献率/%	累积方差贡献率/%
公共因子 1	7.020	54.040	54.040
公共因子 2	3.768	29.003	83.043
公共因子 3	1.023	7.876	90.919

表 7 因子得分矩阵  
Table 7 Factor score matrix

索引	公共	公共	公共	索引	公共	公共	公共
	因子 1	因子 2	因子 3		因子 1	因子 2	因子 3
$H_1$	-0.013	0.173	-0.021	$H_{13}$	0.167	-0.041	0.060
$H_2$	-0.014	0.174	-0.020	$H_{14}$	0.166	-0.040	0.060
$H_3$	-0.014	0.171	-0.019	$H_{15}$	0.165	-0.040	0.061
$H_4$	-0.015	0.172	-0.023	$H_{16}$	0.164	-0.040	0.061
$H_5$	-0.015	0.179	-0.024	$H_{17}$	0.163	-0.039	0.062
$H_6$	-0.015	0.180	0.033	$H_{18}$	0.162	-0.039	0.062
$H_7$	-0.016	0.181	0.032	$H_{19}$	0.161	-0.038	0.062
$H_8$	-0.016	0.182	0.031	$H_{20}$	0.160	-0.037	0.063
$H_9$	-0.017	0.183	0.029	$H_{21}$	0.159	-0.036	0.063
$H_{10}$	-0.017	0.185	0.028	$H_{23}$	0.140	-0.012	-0.081
$H_{11}$	0.169	-0.042	-0.068	$H_{26}$	0.128	-0.015	-0.079

共因子的累积方差贡献率为 90.919%(> 85%)，表明以上 3 个公共因子涵盖了 24 个输入特征序列中的绝大部分有效信息，这对于解释原数据变化具有重要意义。

为验证 EFA 算法的有效性，将其分别与 PCA、ICA 和 FA 等特征降维方法进行对比(以上对照组的预测步长均为 4，输入的特征均来自于 MVRapidMIC 算法的提取结果)。预测结果对比如表 8 所示。其中，PCA-DWSInformer 模型、ICA-DWSInformer 模型、FA-DWSInformer 模型和 EFA-DWSInformer 模型的运行时间分别为 22.46 s、24.94 s、30.19 s 和 27.08 s。

表 8 不同特征降维方法的负荷预测精度的比较

Table 8 Comparison of load prediction accuracy by different feature dimensionality reduction methods

预测模型	MW									
	春		夏		秋		冬		平均	
	RMSE	MAE								
PCA-DWSInformer	8.034	6.953	7.926	6.687	8.265	7.154	8.206	7.109	8.108	6.976
ICA-DWSInformer	7.192	6.077	7.425	6.395	7.649	6.588	7.617	6.539	7.471	6.399
FA-DWSInformer	6.752	5.641	6.756	5.679	6.898	5.739	6.921	5.874	6.382	5.733
EFA-DWSInformer	<b>6.307</b>	<b>5.285</b>	<b>6.437</b>	<b>5.336</b>	<b>6.475</b>	<b>5.383</b>	<b>6.492</b>	<b>5.396</b>	<b>6.427</b>	<b>5.350</b>

由表 8 可以看出：在气候变化平稳的春夏季，4 种预测模型都可以对配电网负荷序列进行较为准确的预测，且 EFA-DWSInformer 模型的预测精度最接近真实值，其 RMSE 和 MAE 均低于其他 3 种预测模型。而在气候变化复杂的秋冬季，4 种预测模型的预测精度均有所下降。不过，EFA-DWSInformer 模型的预测精度仍然最高。相比于春夏季，EFA-DWSInformer 模型的 RMSE 仅增加了 1.75%，MAE 增加了 1.49%。而另外 3 个预测模型的 RMSE 分别增加了 3.20%、4.44%和 2.30%，MAE 分别增加了 4.56%、5.25%和 2.59%。在训练效率方面，虽然 EFA-DWSInformer 算法相对于 PCA-DWSInformer 算法和 ICA-DWSInformer 算法分别慢了 4.62 s 和 2.14 s，但是其平均 RMSE 分别降低了 26.16%、16.24%，平均 MAE 分别降低了 30.39%、19.61%。

以上仿真结果证明本文所提特征处理方法对输入特征序列进行提取与降维后，具有较高的预测精度、鲁棒性和训练效率。

### 4.3 仿真案例 2：不同负荷预测模型精度对比测试

4.3.1 与基础预测模型对比测试  
为了验证 DWSInformer 预测模型的精度，将本文所提模型分别与 RNN、LSTM、GRU、Prophet、Transformer 和 Informer 等 6 种模型进行比较(以上对照组的预测步长均为 4，RNN、LSTM、GRU、

Transformer 和 Informer 的训练迭代次数均为 100 次，Prophet 则使用马尔科夫链蒙特卡洛方法，计算置信区间时设定采样次数为 1500 次，所有预测模型使用的输入特征序列均为 MVRapidMIC-EFA 算法处理后的 3 个公共因子)，其预测结果见表 9。

表 9 不同模型预测性能的比较

Table 9 Comparison of predictive performance among different models

预测模型	评价指标	
	RMSE	MAE
RNN	16.353	15.759
LSTM	15.618	14.256
GRU	13.592	12.098
Prophet	12.218	10.951
Transformer	10.375	8.786
Informer	8.146	6.837
DWSInformer	<b>6.427</b>	<b>5.350</b>

由表 9 可以看出：本文所提模型在预测步长为 4 时，其 RMSE 和 MAE 均优于各个基础模型。具体而言，DWSInformer 模型的 RMSE 值与 Informer 的预测结果相比降低了 26.75%，与 Transformer、Prophet、GRU、LSTM 和 RNN 相比，RMSE 值分别降低了 38.05%、47.40%、52.71%、58.85%和 60.69%。DWSInformer 模型的 MAE 值与 6 种对照

模型相比, 分别降低了 27.79%、39.11%、51.30%、55.78%、62.47%和 66.05%。

为展现预测结果, 本文随机选择 2014 年某一天 0 时作为预测的起始点, 将不同预测模型的预测结果与真实数据进行比较, 具体结果如图 4 所示。

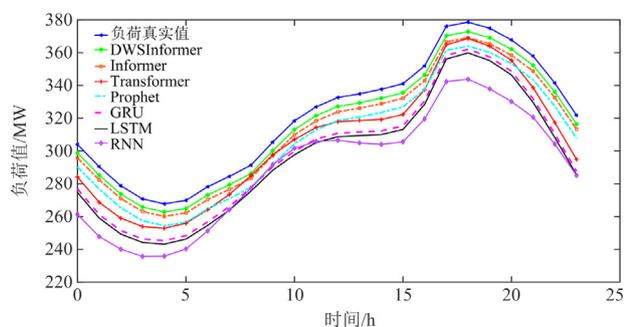


图 4 DWSInformer 模型与其他模型对比

Fig. 4 Comparison of the DWSInformer model with other models

从图 4 可以看出: 与其他 6 种模型相比, DWSInformer 模型可以较好地拟合真实负荷数据。因为利用 DWS 嵌入机制和 TSA 机制可以提取多变量之间的相关性, 以全面考虑与负荷数据相关的特征序列的复杂性。此外, DWSInformer 的生成式结构避免了误差累积问题。这些特点使得 DWSInformer 预测模型能更好地适应真实负荷数据, 有效提高了预测精度。

#### 4.3.2 不同预测步长对比测试

为测试不同预测步长对 DWSInformer 模型预测性能的影响, 利用测试集数据对预测步长为 8、16 和 32 情况下的模型预测效果进行对比(所有预测模型使用的输入特征均为 MVRapidMIC-EFA 算法处理后的 3 个公共因子), 预测结果如表 10 所示。

由表 10 可以看出, 在不同预测步长上本文所提模型的预测精度最高。随着预测步长的增加, 7 种模型的误差均随着时间尺度的增加而增加。但 DWSInformer 模型的 RMSE 和 MAE 值变化范围最小。这是由于 DWSInformer 模型能够通过 DWS 嵌入机制和 TSA 机制更准确地建模特征与负荷数据之间的相关性, 并通过自身的生成式结构和自注意力机制有效提高预测的鲁棒性, 从而在不同预测步长下均表现出较好的预测性能。

#### 4.4 仿真案例 3: 在连续日下的负荷预测分析

为验证 DWSInformer 模型在连续日下的负荷预测中同样具有较好的精度和鲁棒性, 本文将 7 种预测模型的预测情况进行对比(所有预测模型使用的输入特征序列均为 MVRapidMIC-EFA 算法处理

表 10 不同步长的模型预测性能的比较

Table 10 Comparison of predictive performance among different models with different time steps

预测模型	步长	评价指标	
		RMSE	MAE
RNN	8	26.525	25.227
	16	35.001	33.779
	32	56.080	54.928
LSTM	8	24.224	23.188
	16	34.578	32.671
	32	43.191	41.591
GRU	8	20.959	19.133
	16	30.764	27.933
	32	39.297	37.769
Prophet	8	18.250	16.010
	16	26.250	24.301
	32	30.883	28.297
Transformer	8	15.393	13.379
	16	22.120	19.786
	32	29.830	27.211
Informer	8	11.944	9.992
	16	16.539	13.505
	32	22.051	18.638
DWSInformer	8	<b>9.424</b>	<b>7.819</b>
	16	<b>13.049</b>	<b>10.568</b>
	32	<b>17.397</b>	<b>15.585</b>

后的 3 个公共因子), 从 2014 年中随机选择 5 个连续日进行预测(共 120 h), 预测结果如图 5 所示。

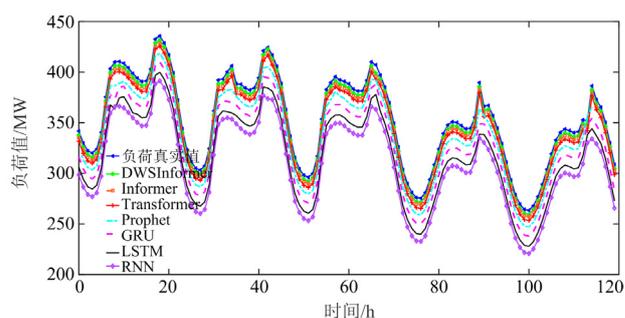


图 5 连续日下负荷预测模型预测结果对比

Fig. 5 Comparison of forecasting results of load models in continuous days

由图 5 可知, 负荷值具有一定的周期性和趋势性, 而 7 种预测模型在预测整体趋势方面表现相似。但是, 在连续日下的负荷预测中, RNN、LSTM、GRU 和 Prophet 等预测模型会出现某些时段误差增大的情况, 其预测的鲁棒性较差。而 DWSInformer、Informer 和 Transformer 模型在负荷突变处预测性能

良好。当突变范围较小时,这两种模型都能够较为准确地进行预测,但前者误差更小且拟合能力更强。当突变范围较大,特别是在一些突变极值点处,DWSInformer模型相对于传统Transformer模型表现得更加稳定。

综上,在连续日下进行负荷预测时,相比于传统方法,DWSInformer模型的精度和鲁棒性更优。

## 5 结论

为了在复杂多变的天气特征下,实现对配电网准确且快速的短期负荷预测,本文提出了一种考虑天气特征与多变量相关性的配电网短期负荷预测方法。该预测方法由MVRapidMIC、EFA和DWSInformer三个主要部分组成。采用DTU 7K 47节点真实配电网系统的历史负荷数据对本文所提方法进行验证,得出以下结论。

1) 本文提出的MVRapidMIC在进一步提升运算速度的同时,解决了传统RapidMIC算法无法分析多自变量与单变量之间相关性的问题。仿真案例1的结果证明,预测模型通过合理的特征提取和降维,使训练时间减少了10.30%,预测精度提升了14.71%。

2) 本文提出的DWSInformer预测模型在一定程度上弥补了忽略不同特征序列之间的依赖性所带来的缺陷。与RNN、LSTM、GRU、Prophet、Transformer和Informer等预测模型相比,在不同预测步长下RMSE和MAE值的变化范围最小,具有较高的预测精度。

3) 本文提出的DWSInformer模型能够高效地完成在连续日下的完整模型训练和负荷预测任务。相较于其他传统的基于机器学习的预测模型,DWSInformer模型具有更高的精度和鲁棒性,在配电网短期负荷预测方面具有较好的应用价值。

本文未来的工作集中于将所提模型和方法在更多的实际配电系统中进行测试,并为配电网预测辅助状态估计提供高精度的数据基础,以提高其实际应用价值。

## 参考文献

- [1] 秦博宇,李恒毅,张哲,等. 地下空间支撑下的电力能源系统:构想、挑战与展望[J]. 中国电机工程学报, 2022, 42(4): 1321-1331.
- [2] 林盛振,谢敏,黄彬彬,等. 基于贝叶斯正则优化NARX神经网络的电力负荷预测[J]. 供用电, 2022, 39(9): 51-60.
- [3] 朱琼锋,李家腾,乔骥,等. 人工智能技术在新能源功率预测的应用及展望[J]. 中国电机工程学报, 2023, 43(8): 3027-3048.
- [4] ZHU Qiongfeng, LI Jiateng, QIAO Ji, et al. Application and outlook of artificial intelligence technology in new energy power forecasting[J]. Proceedings of the CSEE, 2023, 43(8): 3027-3048.
- [5] ZHOU Z, XU Y, REN C. A transfer learning method for forecasting masked-load with behind-the-meter distributed energy resources[J]. IEEE Transactions on Smart Grid, 2022, 13(6): 4961-4964.
- [6] 马伟,乔颖,鲁宗相,等. 基于敏感气象特征因子筛选与优化组合的短期风电功率预测[J]. 电网技术, 2023, 47(7): 2897-2908.
- [7] MA Wei, QIAO Ying, LU Zongxiang, et al. Short-term wind power prediction based on combination of screening and optimizing sensitive meteorological characteristics[J]. Power System Technology, 2023, 47(7): 2897-2908.
- [8] 葛磊蛟,刘航旭,孙永辉,等. 智能配电网多元电力用户群体特性精准感知技术综述[J/OL]. 电力系统自动化: 1-18[2023-06-22]. <http://kns.cnki.net/kcms/detail/32.1180.TP.20230525.1634.004.html>.
- [9] GE Leijiao, LIU Hangxu, SUN Yonghui, et al. Review on the technology of accurate awareness of the characteristics of multi-power users in smart distribution network[J/OL]. Automation of Electric Power Systems: 1-18[2023-06-22]. <http://kns.cnki.net/kcms/detail/32.1180.TP.20230525.1634.004.html>.
- [10] RESHEF D N, RESHEF Y A, FINUCANE H. K, et al. Detecting novel associations in large data sets[J]. Science, 2011, 334: 1518-1524.
- [11] TANG D, WANG M, ZHENG W, et al. RapidMic: rapid computation of the maximal information coefficient[J]. Evolutionary Bioinformatics, 2014.
- [12] 李忠红,何乐生,汪静,等. 基于气象信息充分挖掘的多尺度光伏功率预测研究[J]. 中国测试, 2022, 48(12): 111-117.
- [13] LI Zhonghong, HE Lesheng, WANG Jing, et al. Research on multi-scale photovoltaic power prediction based on full mining of meteorological information[J]. China Measurement & Test, 2022, 48(12): 111-117.
- [14] ATTOURI K, MANSOURI M, HAJJI M, et al. Wind power converter fault diagnosis using reduced kernel PCA-based BiLSTM[J]. Sustainability, 2023, 15(4): 3191.
- [15] ZHANG G, DARAZ A, KHAN I A, et al. Driver training based optimized fractional order PI-PDF controller for frequency stabilization of diverse hybrid power system[J].

- Fractal and Fractional, 2023, 7(4): 315.
- [12] ZHANG L, YAN Y, XU W, et al. Carbon emission calculation and influencing factor analysis based on industrial big data in the “double carbon” era[J]. Computational Intelligence and Neuroscience, 2022.
- [13] YU Y, GUO J, JIN Z. Optimal extreme random forest ensemble for active distribution network forecasting-aided state estimation based on maximum average energy concentration VMD state decomposition[J]. Energies, 2023, 16(15): 5659.
- [14] KUMAR V, PRASAD U, MOHANTY S R. Entirely coupled recurrent neural network-based backstepping control for global stability of power system networks[J]. IEEE Transactions on Automation Science and Engineering, 2023: 1-14.
- [15] 穆晨宇, 薛文斌, 穆羨瑛, 等. 基于 VMD-LSTM-Attention 模型的短期负荷预测研究[J]. 现代电子技术, 2023, 46(17): 174-178.
- MU Chenyu, XUE Wenbin, MU Xianying, et al. Research on short-term load forecasting based on VMD-LSTM-Attention model[J]. Modern Electronics Technique, 2023, 46(17): 174-178.
- [16] 王淞瑶, 张智晟. 基于量子加权多层次 GRU 神经网络的综合能源系统多元负荷短期预测[J]. 电力系统保护与控制, 2022, 50(23): 85-93.
- WANG Songyao, ZHANG Zhisheng. Short-term multivariate load forecasting of an integrated energy system based on a quantum weighted multi-hierarchy gated recurrent unit neural network[J]. Power System Protection and Control, 2022, 50(23): 85-93.
- [17] 欧阳福莲, 王俊, 周杭霞. 基于改进迁移学习和多尺度 CNN-BiLSTM-Attention 的短期电力负荷预测方法[J]. 电力系统保护与控制, 2023, 51(2): 132-140.
- OUYANG Fulian, WANG Jun, ZHOU Hangxia. Short-term power load forecasting method based on improved hierarchical transfer learning and multi-scale CNN-BiLSTM-Attention[J]. Power System Protection and Control, 2023, 51(2): 132-140.
- [18] SAEED N, NGUYEN S, CULLINANE K, et al. Forecasting container freight rates using the Prophet forecasting method[J]. Transport Policy, 2023, 133: 86-107.
- [19] 于越, 丁磊, 金朝阳. 考虑小样本不平衡的主动配电网预测辅助鲁棒状态估计[J/OL]. 高电压技术: 1-12[2023-06-23]. DOI: 10.13336/j.1003-6520.hve.20230785.
- YU Yue, DING Lei, JIN Zhaoyang. Robust forecasting-aided state estimation method of active distribution network considering small sample imbalance[J]. High Voltage Engineering: 1-12[2023-06-23]. DOI: 10.13336/j.1003-6520.hve.20230785.
- [20] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv, 2017.
- [21] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: beyond efficient transformer for long sequence time-series forecasting[C]// 35th AAAI Conference on Artificial Intelligence: Association for the Advancement of Artificial Intelligence, 2021: 11106-11111.
- [22] 张倩, 蒙飞, 李涛, 等. 基于周期信息增强的 Informer 光伏发电功率预测[J]. 中国电力, 2023, 56(10): 186-193.
- ZHANG Qian, MENG Fei, LI Tao, et al. Informer photovoltaic power generation forecasting based on cycle information enhancement[J]. Electric Power, 2023, 56(10): 186-193.
- [23] 齐结红, 钱虹, 吴文军. 超短期热负荷预测在发电机组厂级供热调度的应用[J]. 电力系统保护与控制, 2023, 51(18): 117-124.
- QI Jiehong, QIAN Hong, WU Wenjun. Application of ultra-short-term heat load forecasting in power plant level heat supply dispatching[J]. Power System Protection and Control, 2023, 51(18): 117-124.
- [24] ZHOU N, ZHENG Z, ZHOU J. Prediction of the RUL of PEMFC based on multivariate time series forecasting model[C]// 2023 3rd International Symposium on Computer Technology and Information Science (ISCTIS), 2023, Chengdu, China: 87-92.
- [25] 赵昕迪. 基于 EFA-GA-XGBoost 组合预测模型的绝缘子表面污秽程度预测方法[J]. 电子测试, 2022, 36(5): 68-70.
- ZHAO Xindi. Insulator surface pollution degree prediction method based on EFA-GA-XGBoost combined forecasting model[J]. Electronic Testing, 2022, 36(5): 68-70.
- [26] 王月. 最大信息系数的算法分析及改进[D]. 西安: 西安电子科技大学, 2020.
- [27] BAVISKAR A, HANSEN A D, DAS K, et al. Open-source active distribution grid model with a large share of res-features, and studies[C]// 2021 9th IEEE International Conference on Power Systems (ICPS), December 16-18, 2021, Kharagpur, India: 1-6.

收稿日期: 2023-10-16; 修回日期: 2023-11-24

作者简介:

于 越(1998—), 男, 硕士研究生, 研究方向为基于机器学习的电力系统状态估计; E-mail: yuyue\_98@163.com

葛磊蛟(1984—), 男, 博士, 副教授, 研究方向为智能配电网态势感知、新能源并网优化控制和智能配用电大数据云计算技术; E-mail: legendgj99@tju.edu.cn

金朝阳(1991—), 男, 通信作者, 博士, 副教授, 研究方向为低惯量电力系统关键参数估计和电力系统状态估计。E-mail: zhaoyang.jin@sdu.edu.cn

(编辑 许 威)