

基于条件增强扩散模型的短期负荷概率预测

刘锦翔¹, 张江丰², 董山玲¹, 刘妹琴^{1,3}, 张森林^{1,4}

(1. 浙江大学电气工程学院, 浙江省杭州市 310027; 2. 国网浙江省电力有限公司电力科学研究院, 浙江省杭州市 310027; 3. 西安交通大学人工智能与机器人研究所, 陕西省西安市 710049; 4. 浙江大学金华研究院, 浙江省金华市 321036)

摘要: 负荷概率预测可为电网规划提供指导,条件生成模型通过挖掘历史相似日信息,有效提高了预测性能。然而,以往研究未关注曲线形态信息和无条件模型的噪声解析作用,使得生成曲线的不确定性较高。为此,提出基于条件增强扩散模型的短期负荷概率预测方法。首先,构建改进的iTransformer日负荷预测模型,对相邻日负荷数据进行预测。其次,采用无条件噪声估计和条件噪声估计相结合的损失函数,构建结合多头自注意力机制和U型网络的扩散模型。然后,将日负荷预测结果和温度等特征作为条件输入,通过条件增强引导的反向扩散过程,将多组随机噪声进行去噪,生成多条负荷曲线进行概率密度分析。最后,基于中国某地区公开数据集展开算例分析,与多种模型进行对比,结果表明所提方法具有更高的预测精度。

关键词: 概率预测; 负荷预测; 扩散模型; 注意力机制

0 引言

负荷概率预测是应对负荷不确定性影响的一项关键技术手段,可为电力系统的调度运行、风险管控、经济优化等提供必要的基础信息。

常用的负荷概率预测方法包括统计学方法和人工智能方法等。自回归移动平均模型^[1]、半参数方法^[2]等统计学方法难以充分挖掘数据中的非线性特征。长短期记忆(long short-term memory, LSTM)网络^[3-4]等人工智能方法可以学习序列的时间关系。近年来,Transformer^[5-6]等算法通过注意力机制^[7-8]捕捉序列中的全局信息,有效提高了预测的精度。

然而,传统的概率预测方法^[9-10]使用滑动窗口的方式构建序列特征,未能充分利用数据集中存在的相似日特征。相似日算法^[11-13]通过考虑气象等影响因素,挖掘历史数据集中的强相关序列,降低了预测误差。但为了构建相似日数据集,使用Pearson等相关性算法寻找相似日的步骤较为烦琐,人工调整参数较多,往往存在数据集过拟合。

条件生成模型^[14]能够基于历史数据集特征,根据输入条件生成内容。文献[15]提出基于条件生成对抗网络(conditional generative adversarial network, CGAN)的负荷概率预测方法,构建日负荷

关键值(日负荷最大值和最小值)到全天负荷的非线性关系,有效挖掘了历史数据集中的相似日信息。但CGAN存在难以训练、模式崩溃等问题,而扩散模型^[16-18]通过加噪和还原数据的方式进行内容生成,仅需要训练单个网络进行噪声估计,并在生成准确性和多样性方面优于基于生成对抗网络(GAN)的模型^[19-21]。文献[22-23]提出基于条件扩散模型的风电概率预测方法,以风速、风向等环境影响因素作为条件输入,准确预测了风电出力的概率区间。

上述研究有效提升了概率预测性能,但仍存在以下问题:1)CGAN方法输入的条件较为简单,未能考虑负荷曲线的形态信息;2)扩散模型方法在反向扩散时仅依赖条件模型进行数据生成,忽略了无条件模型对噪声的有效解析能力,导致预测区间较宽,预测精度存在进一步提升的空间。

为此,本文提出基于条件增强扩散模型(conditional enhanced diffusion model, CEDM)的短期负荷概率预测方法。首先,构建适应单序列预测的改进iTransformer日负荷预测模型进行日前负荷预测。然后,构建基于多头注意力机制和U型网络的扩散模型,提出无条件噪声估计和条件噪声估计相结合的损失函数,使模型充分学习噪声解析能力及条件引导能力。最后,将包含负荷形态信息的日负荷预测结果与环境影响因素作为条件,通过条件增强引导的反向扩散过程,对多组随机噪声进行去噪,生成多条负荷曲线,并对结果进行核密度估计,得到负荷预测区间。

收稿日期: 2024-01-09; 修回日期: 2024-04-23。

上网日期: 2024-07-08。

浙江省自然科学基金资助项目(LZ22F030006)。

1 基于改进 iTransformer 的日负荷预测

扩散模型的输入由日负荷数据和负荷影响因素组成,日负荷数据能否准确预测将影响曲线的生成效果。因此,本文对时序预测领域中表现出色的 iTransformer 模型进行改进,以进行日负荷预测,其对 Transformer 的结构进行调整以适应负荷、温度等多序列输入输出的任务,能够充分挖掘日间电力负荷的时序特征。如图 1 所示, iTransformer 主要包括输入模块、特征提取模块和输出模块。该模型的初始结构已经在多个数据集中取得了优越的表现^[6],本文在保留其大部分结构的同时,在输出模块中添加并行的一维卷积神经网络(1D-CNN),进一步提升模型在单序列负荷预测任务中的性能。

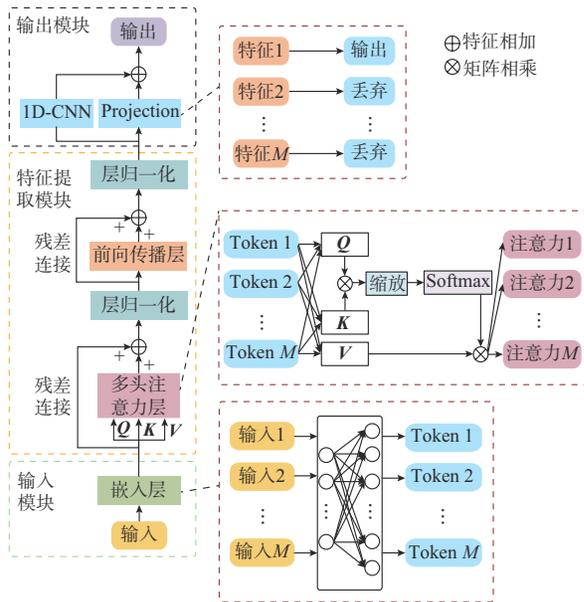


图 1 改进 iTransformer 结构
Fig. 1 Structure of improved iTransformer

1.1 iTransformer 的输入模块

如图 1 输入模块所示,为了保持不同输入特征的独立性, iTransformer 的嵌入层将每个变量分别编码为高维特征 Token,扩充了每个 Token 的感受野,有效挖掘了时序信息。此外,层归一化(LayerNorm)对每个 Token 分别处理,可以有效避免不同特征可能存在的数值量级差异,如式(1)所示。

$$\text{LayerNorm}(H) = \left\{ \frac{h_n - E(h_n)}{\sqrt{D(h_n)}} \mid n = 1, 2, \dots, N \right\} \quad (1)$$

式中: H 为待归一化的矩阵; h_n 为矩阵 H 中待归一化的第 n 个序列; $E(\cdot)$ 为均值函数,对输入序列计算

均值; $D(\cdot)$ 为方差函数,对输入序列计算方差; N 为矩阵 H 中待归一化的序列数。

1.2 iTransformer 的特征提取模块

特征提取模块由多头注意力层和前向传播层构成。多头注意力层可以有效提取序列特征,而前向传播层为全连接层,用于丰富网络特征。注意力机制可以挖掘输入数据不同部分之间的相关性,在处理时序数据上有较为出色的性能表现。

如图 1 所示,注意力机制(Attention)结构可分为 3 个分支,分别为查询向量 Q 、关键向量 K 和值向量 V ,三者关系如式(2)所示。其中,Softmax 用于数据归一化。

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

式中: d_k 为输入 K 的维度。

当输入注意力机制的 Q 、 K 和 V 为同一特征时,注意力机制转化为自注意力机制。自注意力机制可以关注同一序列内部特征之间的相关性。设计多个自注意力机制模块对同一序列进行处理,即转化为多头自注意力机制,既能增加神经网络可训练参数,提高数据挖掘能力,又能够提取序列内部多对特征之间的互信息。

1.3 改进 iTransformer 的输出模块

如图 1 所示, iTransformer 模型在结构设计上适应了多序列输入和多序列输出的任务,其预测层(Projection)模块会选择对应数量的序列进行输出,而其他特征将被抛弃,这在单序列负荷预测的任务中会丢失特征,未能充分利用网络提取所得到的多维特征。

本文构造的输出模块在原 Projection 模块的基础上,以跳跃连接的方式加入并行的 1D-CNN 层,将多维特征映射为单维特征,有效利用了 Projection 模块丢弃的特征信息,表达式如下:

$$M_{\text{Output}}(F) = M_{\text{Projection}}(F) + M_{\text{CNN}}(F) = f_n + M_{\text{CNN}}(F) \quad (3)$$

式中: F 为特征矩阵; f_n 为特征矩阵 F 的第 n 个序列特征; M_{Output} 为输出模块; $M_{\text{Projection}}$ 为 Projection 模块; M_{CNN} 为 1D-CNN 模块。

2 条件增强扩散模型

扩散模型是新兴的生成模型,分为前向扩散与反向扩散两个过程。如图 2 所示,在前向扩散过程中,模型逐步对输入数据添加高斯噪声,逐渐破坏原始数据;在反向扩散过程中,生成模型的任务是通过学习噪声数据、扩散步长和条件三者的关系,逆转扩散过程,从加噪数据中恢复原始输入数据。

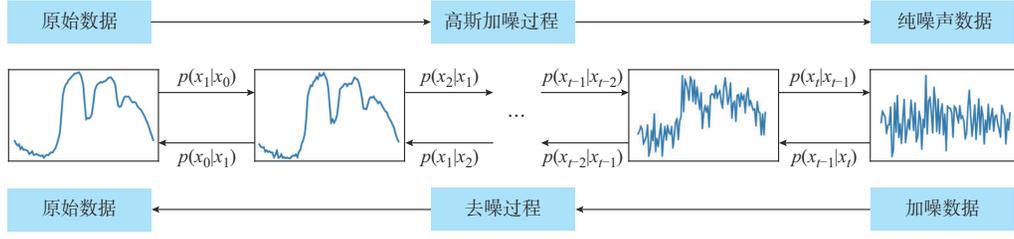


图2 前向扩散及反向扩散过程
Fig. 2 Processes of forward diffusion and reverse diffusion

2.1 前向扩散过程

前向扩散过程按照设计好的噪声强度列表 $\{\beta_s\}_{s=1}^t$ 不断向原始数据中加入噪声,其中, β_s 为第 s 个加噪强度, t 为扩散步长。噪声强度介于 0~1 之间,在多步扩散中,通常采用更大的噪声强度以达到更明显的扩散效果。为了简化推导过程,如式(4)所示,采用换元的方法将 $\{\beta_s\}_{s=1}^t$ 替换为 $\{\alpha_s\}_{s=1}^t$ 。

$$\bar{\alpha}_t = \prod_{s=1}^t \alpha_s = \prod_{s=1}^t (1 - \beta_s) \quad (4)$$

式中: α_s 为 β_s 换元之后的变量; $\bar{\alpha}_t$ 为 α_s 从 1 到 t 的累乘。

通过式(5)的化简,前向扩散可从逐步添加噪声转变为一步计算,大幅节省了时间。由结果可知,由 x_0 推出 x_t 的条件概率分布 $p(x_t|x_0)$ 满足高斯分布,如式(6)所示。

$$\begin{aligned} x_t &= \sqrt{\alpha_t} x_{t-1} + \sqrt{1 - \alpha_t} \epsilon_{t-1} = \sqrt{\alpha_t} (\sqrt{\alpha_{t-1}} x_{t-2} + \\ &\quad \sqrt{1 - \alpha_{t-1}} \epsilon_{t-2}) + \sqrt{1 - \alpha_t} \epsilon_{t-1} = \sqrt{\alpha_t \alpha_{t-1}} x_{t-2} + \\ &\quad \sqrt{(\alpha_t - \alpha_t \alpha_{t-1})^2 + (1 - \alpha_t)^2} \bar{\epsilon}_{t-2} = \dots = \\ &\quad \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon \end{aligned} \quad (5)$$

$$p(x_t|x_0) = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon \sim N(\sqrt{\bar{\alpha}_t} x_0, 1 - \bar{\alpha}_t) \quad (6)$$

式中: x_t 为加噪 t 次之后的序列; x_0 为无噪声的原始序列,即加噪次数为 0; α_t 为第 t 步的加噪强度; ϵ_{t-1} 和 ϵ_{t-2} 分别为加噪过程第 $t-1$ 步和 $t-2$ 步的高斯噪声,两者同为满足高斯分布的噪声,可进一步融合为 $\bar{\epsilon}_{t-2}$; ϵ 为整个加噪过程中的高斯噪声。

2.2 反向扩散过程

反向扩散过程是从噪声还原出数据的关键步骤,模型将加噪 t 次后的数据 x_t 逐步去噪,依次得到 x_{t-1} 、 x_{t-2} 等,直到得到 x_0 。由贝叶斯公式可得到式(7),其等号右边的部分无法求解。为顺利求解,在条件部分引入辅助量 x_0 ,可得式(8)。

$$p(x_{t-1}|x_t) = p(x_t|x_{t-1}) \frac{p(x_{t-1})}{p(x_t)} \quad (7)$$

式中: $p(x_t)$ 为 x_t 的概率密度函数。

$$p(x_{t-1}|x_t, x_0) = p(x_t|x_{t-1}, x_0) \frac{p(x_{t-1}|x_0)}{p(x_t|x_0)} \quad (8)$$

利用前向扩散过程中的式(5),可依次表示出 $p(x_t|x_{t-1}, x_0)$ 、 $p(x_{t-1}|x_0)$ 和 $p(x_t|x_0)$,即式(9)、式(10)和式(6),三者均满足高斯分布。多个高斯分布可进行合并,即式(8)可通过运算得到式(11)。

$$p(x_t|x_{t-1}, x_0) = \sqrt{\alpha_t} x_{t-1} + \sqrt{1 - \alpha_t} \epsilon_{t-1} \sim N(\sqrt{\alpha_t} x_{t-1}, 1 - \alpha_t) \quad (9)$$

$$p(x_{t-1}|x_0) = \sqrt{\bar{\alpha}_{t-1}} x_0 + \sqrt{1 - \bar{\alpha}_{t-1}} \epsilon \sim N(\sqrt{\bar{\alpha}_{t-1}} x_0, 1 - \bar{\alpha}_{t-1}) \quad (10)$$

$$\left\{ \begin{aligned} p(x_{t-1}|x_t, x_0) &= p\left(x_{t-1}|x_t, \frac{x_t - \sqrt{1 - \bar{\alpha}_t} \epsilon}{\sqrt{\bar{\alpha}_t}}\right) = \\ &\quad p(x_{t-1}|x_t, t) \sim N(\mu, \sigma^2) \\ \mu &= \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon\right) \\ \sigma^2 &= \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} (1 - \alpha_t) \end{aligned} \right. \quad (11)$$

式(11)中的噪声 ϵ 在反向扩散过程中是未知变量,故扩散模型的目标为构建神经网络,使其能够从加噪序列中估计得到 ϵ ,以便完成反向扩散过程。

2.3 扩散模型的构建及训练

为了深度挖掘历史数据集中存在的相似日数据并进行曲线生成,本文将条件确定为包含形态信息的日负荷数据及相应的负荷影响因素(包括最高温、最低温、平均温度、湿度、是否降雨、月份、日期、星期),以更加精确地生成负荷曲线。

目前,已有的扩散模型基于循环神经网络等进行构建,为进一步提高序列特征挖掘能力,本文构建基于多头自注意力机制和 U 型网络的扩散模型网络,如附录 A 图 A1 所示。其中, U 型网络结构在处理数据时能够深度挖掘多尺度特征,有利于精确生成数据。多头自注意力机制在 Transformer 等算法中表现出优越的序列处理能力,能够深度挖掘负荷

中存在的时序特征。

网络由输入模块、U型网络和输出模块三部分组成。输入模块由输入加噪序列 x_t 、加噪步长 t 和条件 C 三部分组成。输入层均为全连接网络,将输入映射为相同的维度以进行叠加。

U型网络的深度为两层,依次由4个结构相同的注意力模块构成。该模块设置全连接层以丰富输入特征信息,并加入多头自注意力层以深度挖掘序列特征。注意力模块中的全连接层参数如表1所示,多头自注意力机制模块的输入和输出维度保持一致。

表1 U型网络全连接层参数
Table 1 Fully connected layer parameters of U-net

| 名称 | 输入维度 | 输出维度 |
|-------|-------|-------|
| 全连接层1 | 512 | 1 024 |
| 全连接层2 | 1 024 | 2 048 |
| 全连接层3 | 2 048 | 1 024 |
| 全连接层4 | 1 024 | 512 |

输出模块由两部分组成,即输出无条件噪声和条件噪声。输出模块的输出维度均为96,表示模型估计的加噪序列中的噪声。

当有条件 C 输入时,构成条件模型,输出条件噪声;当无条件输入时,构成无条件模型,输出无条件噪声。该设计允许模型在不同输入条件下灵活调整其输出,以提高预测的准确性和适应性。

由2.2节可知,神经网络的核心任务是估计加噪序列中的噪声。在前向扩散过程中,噪声是人为添加的,可让模型学习估计该噪声。因此,可以构造损失函数 l 如下:

$$l = F_{\text{MAE}}(\epsilon, \epsilon_0(x_t, t)) = F_{\text{MAE}}(\epsilon, \epsilon_0(\sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon, t)) \quad (12)$$

式中: ϵ_0 为模型预测的噪声; F_{MAE} 为平均绝对误差损失函数。

本文提出无条件噪声估计和条件噪声估计相结合的损失函数,如式(13)所示。

$$L = l_1 + l_2 \quad (13)$$

式中: l_1 为无条件噪声损失函数; l_2 为条件噪声损失函数; L 为训练过程中总的损失函数。

为了使模型学习从加噪序列中解析噪声的能力,构造无条件噪声损失函数;为了使模型学习使用条件的能力,构造条件噪声损失。尽管两者最终拟合的目标均为式(12)中的噪声,但由于输入不一致,模型能够学习不同输入下的非线性生成能力。

2.4 条件增强引导的反向扩散过程

训练完毕的模型具有从加噪序列中估计噪声的能力,根据式(11)可得反向扩散过程中的表达式为:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_0(x_t, t) \right) + \sqrt{\frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t}} (1 - \alpha_t) \epsilon \quad (14)$$

根据式(14),由设定的步长 t 逐次进行条件引导,直到 t 为0,即可得到去噪后的序列。

为了利用无条件模型有效的噪声解析能力,并进一步增强模型对输入条件的关注,本文构建条件增强引导的反向扩散过程,如附录A图A2所示。首先,将加噪序列经过扩散模型,对噪声进行估计,得到无条件噪声;然后,将加噪序列和条件输入扩散模型,对噪声进行估计,得到条件噪声;最后,将条件噪声和无条件噪声进行求差,得到有条件部分,将条件部分进行 h 倍增强,并叠加无条件噪声,即为条件增强噪声。在条件增强噪声的引导下进行反向扩散过程,可强化条件对生成负荷数据的影响,提高预测精度。条件增强过程可描述如下:

$$\epsilon_e = \epsilon_1 + h(\epsilon_2 - \epsilon_1) \quad (15)$$

式中: ϵ_e 为条件增强噪声; ϵ_1 为无条件噪声; ϵ_2 为条件噪声; h 为增强的倍数。

3 基于CEDM的负荷概率预测

为了深度挖掘历史数据集存在的相似日特征,本文构造基于多头自注意力机制和U型网络的条件扩散模型,使用涵盖负荷形态信息、特征信息的日负荷数据和涵盖温度、湿度、日期等负荷影响因素的条件作为输入,更加精确地生成待预测日负荷曲线,在使用多组噪声生成多条曲线的基础上,实现电力负荷概率预测。

本文所提负荷概率预测流程如图3所示。首先,构造改进的iTransformer日负荷预测模型,对待预测日的全天负荷进行预测。其次,通过前向扩散过程,训练基于多头自注意力机制和U型网络的条件扩散模型,使其能够从加噪序列中估计噪声,还原出负荷序列。然后,将改进的iTransformer日负荷预测结果和温度、湿度、日期等特征作为条件,设定反向扩散总步长 T ,随机生成多组噪声作为扩散模型输入,通过条件增强引导的反向扩散过程生成多组数据。在得到多组噪声数据生成的多组负荷序列后,经过核密度估计得到负荷概率预测结果。

核密度估计需要先构建概率密度函数。给定一组数据 $Z = \{z_1, z_2, \dots, z_{n_i}\}$,可构建以 z 为自变量的

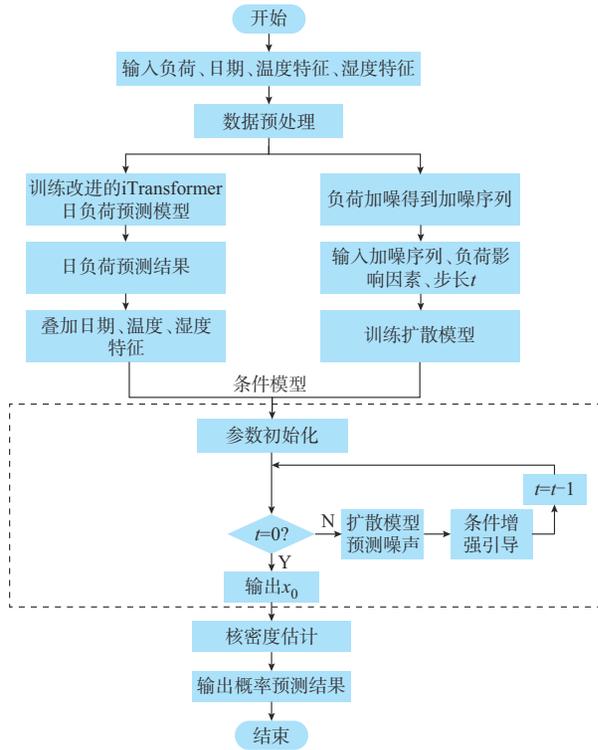


图3 负荷概率预测流程图

Fig. 3 Flow chart of load probabilistic forecasting

概率密度函数 $f(z)$,表达式如下:

$$f(z) = \frac{1}{n_1} \sum_{j=1}^{n_1} K(z - z_j) \quad (16)$$

式中: $K(\cdot)$ 为核函数,本文设置为高斯核函数; n_1 为核密度估计的样本数量。

基于 $f(z)$ 运算得到累积分布函数 $F(z)$,可得分位数 α_i 处的分位数值 q_{α_i} ,以组成置信度为 γ 时的预测区间 $P_{1,\gamma}$,表达式如下:

$$q_{\alpha_i} = F(\alpha_i)^{-1} \quad (17)$$

$$P_{1,\gamma} = [q_{\alpha_i}, q_{\bar{\alpha}_i}] \quad (18)$$

$$\gamma = \bar{\alpha}_i - \alpha_i \quad (19)$$

式中: $\bar{\alpha}_i$ 和 α_i 分别为 $P_{1,\gamma}$ 上、下边界的标称分位水平。

4 算例分析

本文基于 Windows 11 下的 Pytorch 环境(版本为 1.13.1)进行实验,具体配置如下:CPU 为 Intel 13700 KF, GPU 为 NVIDIA 4080。

4.1 数据预处理

选取中国公开的某地负荷数据集^[24]开展算例验证,数据集包含 2012 年至 2014 年共 3 年的数据,数据每隔 15 min 采集一点,即 1 d 负荷由 96 点组成。针对电力负荷历史数据中的缺失数据,使用两侧的

数据均值补充,除此之外不存在其他损坏类型。

将 2012 年、2013 年作为训练集;在 2014 年中,从每个季节随机选取 30 d,共 120 d 作为测试集,其余数据作为验证集。

数据集中的环境影响因素包括最高温度、最低温度、平均温度、湿度和降雨量,均作为负荷影响因素的组成部分。针对月份,依次将 1 至 12 月编码为 1~12;针对季节,依次将春、夏、秋、冬,编码为 0~3;针对星期,依次将周一至周日编码为 0~6。针对年份,由于 2014 年未纳入训练集,若将年份依次编码为 0、1 和 2 并纳入负荷影响因素,则代表 2014 年的 2 未被网络学习过,将作为干扰变量影响网络性能,故在此不考虑年份特征。上述所有特征均按照式(20)进行归一化。

$$P_{\text{norm}} = \frac{P - P_{\min}}{P_{\max} - P_{\min}} \quad (20)$$

式中: P 为待归一化的数据; P_{\min} 为 P 的最小值; P_{\max} 为 P 的最大值; P_{norm} 为归一化后的数据。

4.2 日负荷预测

本文基于改进的 iTTransformer 预测相邻日的日负荷数据,输入特征为电力负荷、最高温度、最低温度、平均温度、湿度、降雨量、月份特征、季节特征、周特征,输出为下一天 96 点的电力负荷数据。将模型的输出维度调整为 96,加入的 1D-CNN 设置为两层,通道数分别为 20 和 1。采用 Adam 优化器,学习率设置为 0.001,批大小设置为 256。训练过程中保存在验证集上表现最优的权重模型,测试时使用该权重计算评价指标。

本文选取的评价指标包括平均绝对百分比误差 E_{MAPE} 和均方根误差 E_{RMSE} ,表达式如下:

$$E_{\text{MAPE}} = \frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} \frac{|y_k - \hat{y}_k|}{y_k} \quad (21)$$

$$E_{\text{RMSE}} = \sqrt{\frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} (y_k - \hat{y}_k)^2} \quad (22)$$

式中: y_k 和 \hat{y}_k 分别为第 k 个元素的真实值和预测值; N_{test} 为测试集中元素的数量。

为说明改进 iTTransformer 模型在日前负荷预测中的有效性,选取 iTTransformer、LSTM 网络,结合 Attention 和 LSTM 的 AttLSTM 模型、由 1D-CNN 和 LSTM 组成的 CnnLSTM 模型作为对比方法,所有模型的输入和输出与改进 iTTransformer 保持一致。各方法均经过超参数调优,并保存在验证集上表现最好的权重参数进行测试。不同季节下,每个算法的性能指标对比如表 2 所示。

表2 不同方法的日负荷预测精度对比
Table 2 Comparison of daily load forecasting accuracies of different methods

| 算法 | $E_{MAPE}/\%$ | | | | | E_{RMSE}/MW | | | | |
|-----------------|---------------|------|------|------|------|---------------|---------|---------|---------|---------|
| | 春 | 夏 | 秋 | 冬 | 均值 | 春 | 夏 | 秋 | 冬 | 均值 |
| 改进 iTransformer | 1.62 | 1.41 | 1.40 | 1.69 | 1.53 | 147.560 | 159.202 | 131.838 | 136.884 | 143.874 |
| iTransformer | 1.71 | 1.49 | 1.53 | 1.76 | 1.62 | 154.714 | 167.688 | 144.446 | 142.066 | 152.225 |
| LSTM | 1.93 | 2.23 | 1.99 | 2.26 | 2.10 | 175.440 | 239.360 | 185.912 | 177.888 | 194.616 |
| AttLSTM | 1.75 | 2.20 | 1.71 | 1.91 | 1.89 | 153.408 | 234.464 | 166.192 | 157.488 | 177.888 |
| CnnLSTM | 1.80 | 1.99 | 1.93 | 2.36 | 2.02 | 157.216 | 222.088 | 176.256 | 189.176 | 186.184 |

表2结果显示,由于iTransformer出色的神经网络架构和多头自注意力机制强大的序列特征挖掘能力,以及为适应单序列预测所增加的1D-CNN模块,本文所提模型相比未改进的iTransformer模型提升约6%的性能,相比传统方法中表现最好的AttLSTM提升约20%的性能,充分表明改进iTransformer模型具有较好的日前负荷预测能力。

4.3 负荷概率预测

本文选取国内外近几年性能优越的负荷概率预测方法进行对比实验,包括去噪扩散概率模型(denoising diffusion probabilistic model,DDPM)^[24]、基于CGAN的概率预测模型^[15]、基于Dropout的不确定性估计模型^[25]、改进分位数长短期记忆(improved quantile long short-term memory,IQLSTM)网络^[26]、贝叶斯长短期记忆(Bayes long short-term memory,BayesLSTM)网络^[27]和概率密度估计的Transformer(Transformer for probabilistic distribution estimation,PdeTransformer)模型^[28]。各方法的实现细节如下:

1)本文所提CEDM的学习率设置为0.0004,批大小设置为256。前向扩散时,加噪策略设置最高步长为1000。反向扩散时,随机噪声的组数设置为200,噪声的维度设置为96,步长设置为50,条件增强过程的 h 取值为2。

2)DDPM在实验参数设置和网络结构上与CEDM保持一致,但反向扩散时,仅使用条件模型进行数据生成。

3)CGAN使用日负荷关键值(负荷最大值和最小值)进行日负荷生成。CGAN训练完毕后,将改进iTransformer预测的日负荷关键值作为条件,使用多组噪声生成多组数据,进行核密度估计。

4)Dropout不确定性估计模型由LSTM层、全连接层和Dropout层组成。Dropout层共两层,神经元丢失概率分别设置为0.2和0.4。模型输入与改进iTransformer的输入保持一致,输出为下一天负荷数据。推理时,Dropout层参数不固定,其他神经网

络参数固定,以Dropout层神经元随机丢失代表负荷的不确定性信息,重复预测日负荷,对多条曲线进行核密度估计。

5)IQLSTM由LSTM层和负责分位数预测的全连接层组成。为减少Pinball损失函数不连续造成的影响,使用两阶段训练的方式。第1阶段,在LSTM层后增加额外的全连接层,对日负荷进行点预测,训练完毕后去掉该全连接层;第2阶段,固定LSTM层,其已具备对负荷曲线进行有效特征提取的能力,训练负责分位数预测的全连接层。

6)BayesLSTM将神经网络的参数视为随机变量,并假设其服从高斯分布。由于BayesLSTM的神经元参数带有不确定性,其输出也带有不确定性,重复预测日负荷,对多条曲线进行核密度估计。

7)PdeTransformer假设每一时刻负荷的概率分布服从高斯分布,并基于极大似然估计和深度学习梯度优化相结合的方法,训练及预测概率分布的均值和标准差。推理时,基于预测所得的均值和标准差构建高斯分布函数,以获取概率预测区间。

针对概率预测,本文采用的性能评价指标包括预测区间覆盖偏差(prediction interval coverage error,PICE)、预测区间平均宽度(prediction interval average width,PIAW)和Winker分数 S_w 。

PICE定义为预测区间包含真实曲线占比 l_k 与置信度 γ 的差距,其绝对值越接近0,表示预测区间可靠性越高,效果越好,表达式为:

$$P_{PICE,\gamma} = \frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} l_k - \gamma \quad (23)$$

$$l_k = \begin{cases} 1 & y_k \in [q_{\alpha}, q_{\bar{\alpha}}] \\ 0 & y_k \notin [q_{\alpha}, q_{\bar{\alpha}}] \end{cases} \quad (24)$$

PIAW表示预测区间的宽度大小,表达式为:

$$P_{PIAW,\gamma} = \frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} (q_{\bar{\alpha}_i} - q_{\alpha_i}) \quad (25)$$

S_w 能够综合考虑区间可靠性和区间宽度,其值越小表示综合性能越好,表达式为:

$$S_{w,y} = \sum_{k=1}^{N_{\text{test}}} S_k \quad (26)$$

$$S_k = q_{\bar{\alpha}_i} - q_{\alpha_i} + \begin{cases} 0 & y_k \in [q_{\alpha_i}, q_{\bar{\alpha}_i}] \\ \frac{1}{\alpha_i} (q_{\alpha_i} - y_k) & y_k < q_{\alpha_i} \\ \frac{1}{1 - \bar{\alpha}_i} (y_k - q_{\bar{\alpha}_i}) & y_k > q_{\bar{\alpha}_i} \end{cases} \quad (27)$$

附录 A 图 A3 展示了不同季节、不同方法在 90% 置信度的预测区间内的对比图。与其他方法相比,本文方法给出了最窄的预测区间,并较好地覆盖了真实值曲线。因受春节假期的影响,负荷不确定性进一步提高,本文方法给出的预测区间相较于其他季节的预测区间更宽,但仍优于对比方法。

图 4 展示了不同季节下本文所提方法在不同置信度(取 10%、30%、50%、70%、90%)的日前负荷概率预测区间。由图可知,CEDM 能够灵活地根据不同时刻负荷的不确定性变化,给出宽窄相宜的预测区间。在 32、48 等负荷迅速变化时间点,均给出了较紧凑的预测区间;而在负荷峰谷等其他位置,由于不确定性较高,所提供预测区间相对宽松,但仍保持相对较窄的预测区间。整体上,本文方法精确地捕捉了负荷的形态信息及不同时刻的不确定性变化,显示了较为出色的不确定性捕捉能力。

图 5 展示了不同季节下不同方法的 PICE 绝对值、PIAW 和 S_w 指标对比图。由图 5(a)–(d) 可知,CEDM 给出的 PICE 指标能够在 0 附近波动。由图 5(e)–(h) 可知,CEDM 在各置信度下均给出了最窄的预测区间,能够有效聚集负荷的不确定性变化。由图 5(i)–(l) 可知,CEDM 在各置信度下均给出了最好的综合评分,体现了更优的负荷概率预测能力。

表 3 展示了不同季节不同方法的平均概率预测性能指标对比。由表 3 可知,CEDM 给出的平均绝对 PICE 指标 $\overline{|P_{\text{PICE},y}|}$ 均低于 3.6%, 平均 PIAW 指标 $\overline{|P_{\text{PIAW},y}|}$ 均低于 0.018, 平均 S_w 指标 $\overline{S_{w,y}}$ 均低于 3.4。其中,IQLSTM 在 PICE 指标上表现出了与 CEDM 相近的性能,体现了较为优秀的分位数预测能力,但其给出的预测区间较宽,不能有效聚集负荷的不确定性变化,导致综合评分不高。作为生成模型的代表,CEDM、DDPM 和 CGAN 均给出了比其他算法更窄的预测区间,体现了良好的不确定性聚集能力。由于考虑了负荷曲线的形态信息,CEDM 比 CGAN 具有更高的预测精度。另外,在无条件模型有效解析噪声的基础上,通过引入目的性更强的条

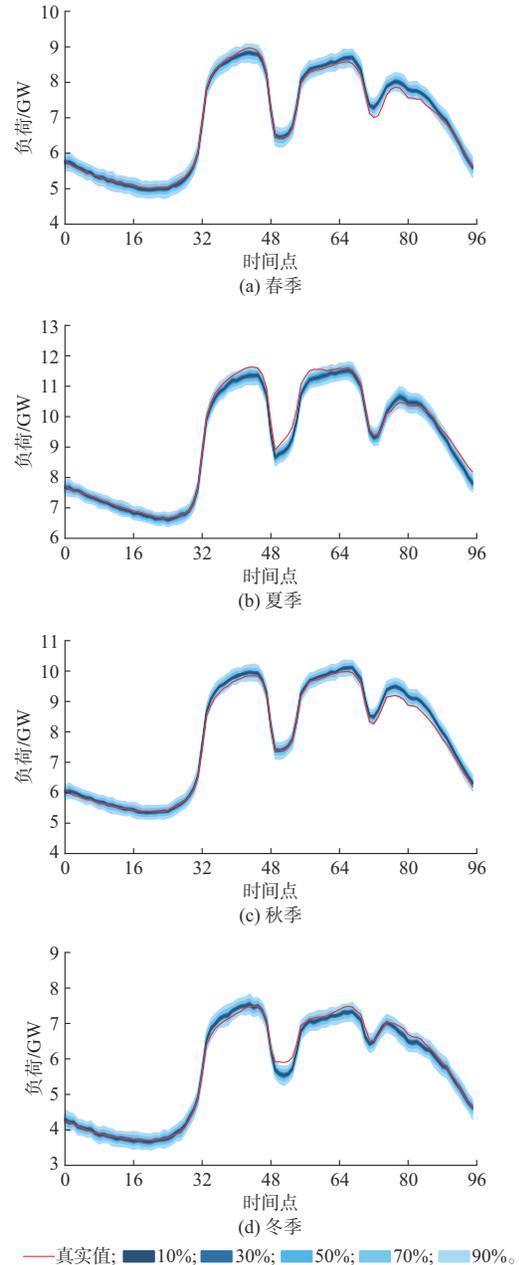


图 4 不同季节下 CEDM 的区间预测结果
Fig. 4 Interval forecasting results of CEDM in different seasons

件增强引导过程,相较于 DDPM,CEDM 生成的负荷数据波动性更小、可靠性更高、综合表现更好。综上,相比其他方法,CEDM 展示了更出色的负荷概率预测性能。

表 4 展示了不同方法的推理耗时对比,由于条件扩散模型需要逐步进行反向扩散以更加精确地生成负荷数据,本文所提方法相比其他方法耗时更长,但少于 700 ms 的推理耗时对于日前负荷概率预测及大部分负荷预测场景是可接受的。因此,本文方法仍具有良好的适用性。

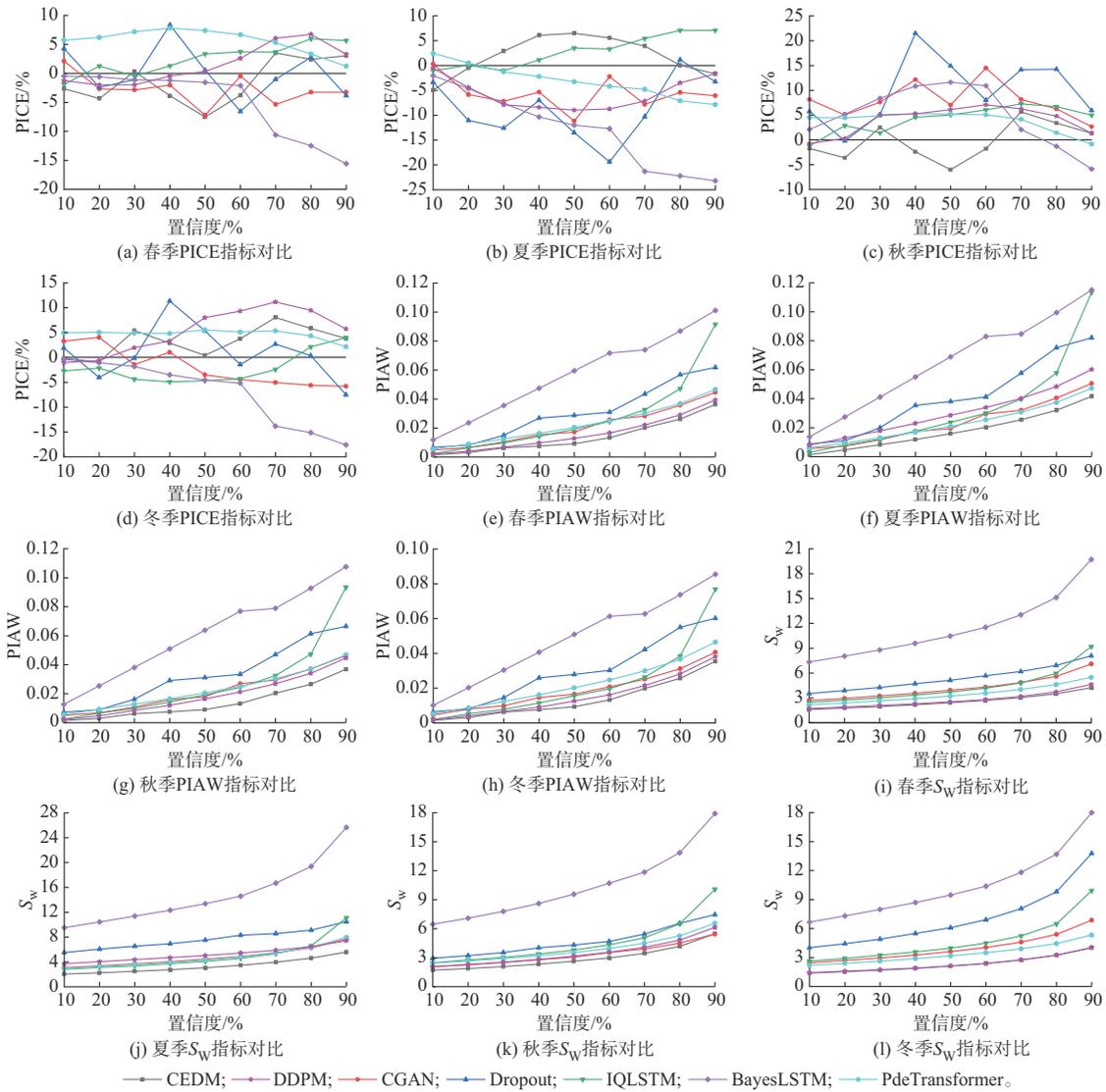


图5 不同季节不同方法的PICE、PIAW和Sw对比
 Fig. 5 Comparison of PICE, PIAW and Sw of different methods in different seasons

表3 不同季节不同方法的平均概率预测性能对比
 Table 3 Comparison of performance of average probability forecasting of different methods in different seasons

| 算法 | $\overline{ P_{PICE,\gamma} }/\%$ | | | | | $\overline{ P_{PIAW,\gamma} }$ | | | | | $\overline{S_{w,\gamma}}$ | | | | |
|----------------|-----------------------------------|------|------|------|------|--------------------------------|-------|-------|-------|-------|---------------------------|--------|--------|--------|--------|
| | 春 | 夏 | 秋 | 冬 | 均值 | 春 | 夏 | 秋 | 冬 | 均值 | 春 | 夏 | 秋 | 冬 | 均值 |
| CEDM | 3.50 | 3.57 | 3.16 | 3.46 | 3.42 | 0.014 | 0.018 | 0.014 | 0.014 | 0.015 | 2.600 | 3.369 | 2.966 | 2.332 | 2.817 |
| DDPM | 2.75 | 5.68 | 4.10 | 5.62 | 4.54 | 0.016 | 0.030 | 0.019 | 0.015 | 0.020 | 2.763 | 5.238 | 3.495 | 2.358 | 3.464 |
| CGAN | 3.24 | 5.71 | 7.96 | 3.79 | 5.17 | 0.021 | 0.024 | 0.022 | 0.019 | 0.021 | 4.242 | 4.787 | 3.340 | 3.988 | 4.089 |
| Dropout | 3.43 | 9.05 | 9.99 | 3.86 | 6.58 | 0.031 | 0.041 | 0.033 | 0.030 | 0.034 | 5.377 | 7.672 | 4.685 | 7.054 | 6.197 |
| IQLSTM | 3.03 | 3.30 | 4.46 | 3.52 | 3.58 | 0.028 | 0.034 | 0.028 | 0.023 | 0.028 | 4.367 | 4.972 | 4.599 | 4.721 | 4.665 |
| BayesLSTM | 5.08 | 12.9 | 6.49 | 7.02 | 7.87 | 0.057 | 0.065 | 0.061 | 0.048 | 0.058 | 11.513 | 14.819 | 10.425 | 10.441 | 11.800 |
| PdeTransformer | 5.65 | 3.73 | 3.99 | 4.68 | 4.51 | 0.023 | 0.023 | 0.023 | 0.022 | 0.023 | 3.445 | 4.758 | 3.887 | 3.386 | 3.869 |

4.4 不同输入条件下的模型性能

为了验证将全天负荷数据作为条件输入更具有优越性以及本文所提方法相比CGAN的优点,本文分别在CEDM和CGAN上使用两种不同的条件进

行实验:1)使用全天负荷数据作为条件;2)使用日负荷关键值作为条件。其中,日负荷数据和日负荷关键值均由本文所提出的改进iTransformer模型预测得到,最终统计结果如表5所示。

表4 不同方法的推理耗时对比
Table 4 Comparison of reasoning time-consuming of different methods

| 算法 | 耗时/ms | 算法 | 耗时/ms |
|---------|-------|----------------|-------|
| CEDM | 696 | IQLSTM | 6 |
| DDPM | 441 | BayesLSTM | 42 |
| CGAN | 1 | PdeTransformer | 2 |
| Dropout | 23 | | |

表5 不同条件下的模型性能
Table 5 Model performances under different conditions

| 条件类型 | iTransformer的 $E_{MAPE}/\%$ | 算法 | $\overline{ P_{PICE,y} }/\%$ | $\overline{ P_{PIAW,y} }$ | $\overline{S_{w,y}}$ |
|--------|--------------------------------|------|------------------------------|---------------------------|----------------------|
| 日负荷数据 | 1.53 | CEDM | 3.42 | 0.015 | 2.817 |
| | | CGAN | 38.61 | 0.004 | 7.697 |
| 日负荷关键值 | 1.22 | CEDM | 3.84 | 0.015 | 3.260 |
| | | CGAN | 5.17 | 0.021 | 4.089 |

由表5可得到如下3个结论:

1)由于日负荷关键值需要预测的量少,仅为最大值和最小值,模型拟合难度低,具有更高的预测精度。

2)对比CEDM的两个实验可以得出,即使日负荷关键值的预测更为准确,但是由于缺失了负荷的形态变化信息,将其作为条件后CEDM给出的概率预测综合性能较低;而使用日负荷数据作为条件,模型能从中挖掘包括负荷形态特征在内的更多特征信息,最终给出更高的综合性能。

3)对比CGAN的两个实验可以看到,使用日负荷数据作为输入条件时,模型给出了很窄的区间宽度和很差的可靠性。原因在于:日负荷数据提供了更多关于曲线的特征信息,然而该信息会被同时输入生成器和判别器。当两者均获得足够多的特征信息时,生成器只能不断减小生成数据的波动程度,避免被判别器判别为假数据,最终导致生成的数据随机性小、预测区间窄。在预测时,CGAN无法获得完全正确的日负荷数据作为条件进行曲线生成,即使iTTransformer预测的日负荷数据误差仅为1.53%,也将因为CGAN的预测区间太窄而无法有效覆盖真实值曲线,最终导致概率预测的区间可靠性很低。

为了更进一步对比以全天负荷数据作为条件的情况下,CEDM和CGAN的性能差距,附录A图A4展示了2014年9月20日真实值、iTTransformer预测结果、CEDM和CGAN的90%概率区间。可以看出,CGAN的输出结果过于依赖条件,且输出区间很小,只能在iTTransformer预测曲线的附近小范围

波动。而CEDM的预测结果能够综合考虑所学历史数据集信息,并结合条件输入,在不同时刻给出适宜的区间宽度,能够更好地覆盖真实值曲线。

综上,CEDM能够基于所学历史数据集和输入条件,生成更加真实可靠的负荷数据。而CGAN受限于其生成器和判别器对抗训练的方式,难以将详细的日负荷数据条件发挥作用,导致生成数据的质量得不到提升。

5 结语

本文提出基于条件增强扩散模型的短期负荷概率预测方法。首先,构建了改进的iTTransformer日负荷预测模型,进行更为精确的日前负荷数据预测;然后,构建了基于多头注意力机制和U型网络的扩散模型,使模型深度学习噪声解析能力和条件引导能力;最后,构建了条件增强引导的反向扩散过程,有效利用了无条件模型的噪声解析能力和条件模型的引导能力,在减小生成数据不确定性的同时,精确捕捉了负荷的形态特征。实验表明,本文方法具有更优的负荷概率预测性能。

扩散模型在负荷预测领域的应用较少,后续工作可至少从以下2个方向进行研究:可尝试降低反向扩散步数以缩短生成数据所需时间,但同时又不影响曲线生成精度^[20];可进一步探讨对于其他数据集的迁移能力^[29],解决部分负荷数据集样本少等问题。

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>),扫英文摘要后二维码可以阅读网络全文。

参考文献

- [1] WAN C, ZHAO J, SONG Y H, et al. Photovoltaic and solar power forecasting for smart grid energy management[J]. CSEE Journal of Power and Energy Systems, 2015, 1(4): 38-46.
- [2] FAN S, HYNDMAN R J. Short-term load forecasting based on a semi-parametric additive model[J]. IEEE Transactions on Power Systems, 2012, 27(1): 134-141.
- [3] 李丹,张远航,杨保华,等.基于约束并行LSTM分位数回归的短期电力负荷概率预测方法[J].电网技术,2021,45(4):1356-1364.
LI Dan, ZHANG Yuanhang, YANG Baohua, et al. Short time power load probabilistic forecasting based on constrained parallel-LSTM neural network quantile regression mode[J]. Power System Technology, 2021, 45(4): 1356-1364.
- [4] 颜伟,李丹,朱继忠,等.月前日负荷曲线的概率预测和随机场景模拟[J].电力系统自动化,2017,41(17):155-162.
YAN Wei, LI Dan, ZHU Jizhong, et al. Probabilistic forecasting and stochastic scenario simulation of month-ahead

- daily load curve [J]. *Automation of Electric Power Systems*, 2017, 41(17): 155-162.
- [5] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]// 31st International Conference on Neural Information Processing Systems, December 4-9, 2017, Long Beach, USA.
- [6] LIU Y, HU T, ZHANG H, et al. iTransformer: inverted transformers are effective for time series forecasting [C]// 12th International Conference on Learning Representations, May 7-11, 2024, Vienna, Austria.
- [7] 赵洪山, 吴雨晨, 温开云, 等. 基于时空注意力机制的台区多用户短期负荷预测[J]. *电工技术学报*, 2024, 39(7): 2104-2115.
ZHAO Hongshan, WU Yuchen, WEN Kaiyun, et al. Short-term load forecasting for multiple customers in a station area based on spatial-temporal attention mechanism [J]. *Transactions of China Electrotechnical Society*, 2024, 39(7): 2104-2115.
- [8] 周思思, 李勇, 郭钺秀, 等. 考虑时序特征提取与双重注意力融合的TCN超短期负荷预测[J]. *电力系统自动化*, 2023, 47(18): 193-205.
ZHOU Sisi, LI Yong, GUO Yixiu, et al. Ultra-short-term load forecasting based on temporal convolutional network considering temporal feature extraction and dual attention fusion [J]. *Automation of Electric Power Systems*, 2023, 47(18): 193-205.
- [9] 王凌云, 周翔, 田恬, 等. 基于多维气象信息时空融合和MPA-VMD的短期电力负荷组合预测模型[J]. *电力自动化设备*, 2024, 44(2): 190-197.
WANG Lingyun, ZHOU Xiang, TIAN Tian, et al. Combination forecasting model of short-term power load based on multi-dimensional meteorological information spatio-temporal fusion and MPA-VMD [J]. *Electric Power Automation Equipment*, 2024, 44(2): 190-197.
- [10] 张建文, 杨晨, 冉懿, 等. 基于PCA-GPQR的电网负荷短期概率预测[J]. *电力系统及其自动化学报*, 2020, 32(5): 24-29.
ZHANG Jianwen, YANG Chen, RAN Yi, et al. Short-term load probability forecasting based on PCA-GPQR [J]. *Proceedings of the CSU-EPSCA*, 2020, 32(5): 24-29.
- [11] 吴浩天, 孙荣富, 廖思阳, 等. 基于改进气象聚类分型的短期风电功率概率预测方法[J]. *电力系统自动化*, 2022, 46(15): 56-65.
WU Haotian, SUN Rongfu, LIAO Siyang, et al. Short-term wind power probability forecasting method based on improved meteorological clustering and classification [J]. *Automation of Electric Power Systems*, 2022, 46(15): 56-65.
- [12] 赵耀, 高少炜, 李东东, 等. 基于天气相似聚类与QRNN的短期光伏功率区间概率预测[J]. *电力系统自动化*, 2023, 47(23): 152-161.
ZHAO Yao, GAO Shaowei, LI Dongdong, et al. Short-term interval probability prediction of photovoltaic power based on weather similarity clustering and quantile regression neural network [J]. *Automation of Electric Power Systems*, 2023, 47(23): 152-161.
- [13] 陈弘川, 蔡旭, 孙国歧, 等. 基于智能优化方法的相似日短期负荷预测[J]. *电力系统保护与控制*, 2021, 49(13): 121-127.
CHEN Hongchuan, CAI Xu, SUN Guoqi, et al. Similar day short-term load forecasting based on intelligent optimization method [J]. *Power System Protection and Control*, 2021, 49(13): 121-127.
- [14] 林珊, 王红, 齐林海, 等. 基于条件生成对抗网络的短期负荷预测[J]. *电力系统自动化*, 2021, 45(11): 52-60.
LIN Shan, WANG Hong, QI Linhai, et al. Short-term load forecasting based on conditional generative adversarial network [J]. *Automation of Electric Power Systems*, 2021, 45(11): 52-60.
- [15] 孙浩, 万灿, 曹照静, 等. 基于条件生成对抗网络曲线生成的短期负荷概率预测[J]. *电力系统自动化*, 2023, 47(23): 189-199.
SUN Hao, WAN Can, CAO Zhaojing, et al. Short-term load probabilistic forecasting based on conditional generative adversarial network curve generation [J]. *Automation of Electric Power Systems*, 2023, 47(23): 189-199.
- [16] YANG L, ZHANG Z L, SONG Y, et al. Diffusion models: a comprehensive survey of methods and applications [J]. *ACM Computing Surveys*, 56(4): 105.
- [17] SONG J, MENG C, ERMON S. Denoising diffusion implicit models [C]// 9th International Conference on Learning Representations, May 3-7, 2021.
- [18] HO J, SALIMANS T. Classifier-free diffusion guidance [EB/OL]. [2023-10-05]. <https://arxiv.org/pdf/2207.12598>.
- [19] DHARIWAL P, NICHOL A. Diffusion models beat GANs on image synthesis [J]. *Advances in Neural Information Processing Systems*, 2021, 34: 8780-8794.
- [20] NICHOL A Q, DHARIWAL P. Improved denoising diffusion probabilistic models [C]// International Conference on Machine Learning, July 18-24, 2021.
- [21] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution image synthesis with latent diffusion models [C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-24, 2022, New Orleans, USA.
- [22] DONG X, MAO Z, SUN Y, et al. Short-term wind power scenario generation based on conditional latent diffusion models [J]. *IEEE Transactions on Sustainable Energy*, 2024, 15(2): 1074-1085.
- [23] CAPEL E H, DUMAS J. Denoising diffusion probabilistic models for probabilistic energy forecasting [C]// 2023 IEEE Belgrade PowerTech, June 25-29, 2023, Belgrade, Serbia.
- [24] 中国电机工程学会电工数学专委会. 第九届“中国电机工程学会杯”全国大学生电工数学建模竞赛题目 [EB/OL]. [2023-04-25]. <http://shumo.nedu.edu.cn>.
Electrical Mathematics Committee of China Electrical Engineering Society. Title of the 9th “China Electrical Engineering Society Cup” national undergraduate electrical mathematics modeling competition [EB/OL]. [2023-04-25]. <http://shumo.nedu.edu.cn>.
- [25] GAL Y, GHAHRAMANI Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning [C]// International Conference on Machine Learning, June 19-24, 2016, New York, USA.
- [26] HUANG Q, WEI S. Improved quantile convolutional neural network with two-stage training for daily-ahead probabilistic forecasting of photovoltaic power [J]. *Energy Conversion and Management*, 2020, 220: 113085.

- [27] TRAN D, DUSENBERRY M, VAN D W M, et al. Bayesian layers: a module for neural network uncertainty [EB/OL]. [2023-09-05]. <https://arxiv.org/abs/1812.03973>.
- [28] 韩宇超,同向前,邓亚平.基于概率密度估计与时序Transformer网络的风功率日前区间预测[J/OL].中国电机工程学报[2024-03-11].<http://kns.cnki.net/kcms/detail/11.2107.TM.20230928.1732.004.html>.
HAN Yuchao, TONG Xiangqian, DENG Yaping. Probabilistic distribution estimation and temporal transformer-based interval prediction in day-ahead wind power prediction [J/OL]. Proceedings of the CSEE [2024-03-11]. <http://kns.cnki.net/kcms/detail/11.2107.TM.20230928.1732.004.html>.
- [29] HU E J, SHEN Y, WALLIS P, et al. Lora: low-rank adaptation of large language models [EB/OL]. [2023-09-05]. <https://arxiv.org/abs/2106.09685>.
- 刘锦翔(1998—),男,硕士研究生,主要研究方向:电力系统不确定性预测。E-mail:22210090@zju.edu.cn
- 张江丰(1988—),男,硕士,高级工程师,主要研究方向:电力系统网源协调控制技术与应用。E-mail:zhangjiangfeng0725@163.com
- 董山玲(1990—),女,通信作者,博士,研究员,博士生导师,主要研究方向:智能电网控制技术及应用。E-mail:shanlingdong28@zju.edu.cn
- (编辑 王梦岩)

Short-term Load Probabilistic Forecasting Based on Conditional Enhanced Diffusion Model

LIU Jinxiang¹, ZHANG Jiangfeng², DONG Shanling¹, LIU Meiqin^{1,3}, ZHANG Senlin^{1,4}

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China;

2. Electric Power Research Institute of State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310027, China;

3. Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China;

4. Jinhua Institute of Zhejiang University, Jinhua 321036, China)

Abstract: Load probabilistic forecasting can provide guidance for power grid planning, and the conditional generation model can effectively improve the forecasting performance by mining historical similar-day information. However, previous studies did not pay attention to the curve shape information and the noise analysis function of unconditional models, which increased the uncertainty of the generation curve. Therefore, a short-term load probabilistic forecasting method based on conditional enhanced diffusion model is proposed. Firstly, an improved iTransformer daily load forecasting model is constructed to forecast the adjacent daily load data. Secondly, a diffusion model combining multi-head self-attention mechanism and U-net is constructed using a loss function that combines unconditional noise estimation and conditional noise estimation. Then, the daily load forecasting results and characteristics such as temperature are used as conditional inputs. Through the reverse diffusion process of conditional enhanced guidance, multiple sets of random noise are denoised to generate multiple load curves for probability density analysis. Finally, based on a publicly available dataset from a region in China and comparative tests with various models, the case study analysis demonstrates that the proposed method has higher forecasting accuracy.

This work is supported by Zhejiang Provincial Natural Science Foundation of China (No. LZ22F030006).

Key words: probabilistic forecasting; load forecasting; diffusion model; attention mechanism

