

基于自适应核密度估计算法的风电功率超短期区间预测

郭恒宽 田建艳 刘竖威 蒋莹莹 杨立志

Wind Power Ultra-short-term Interval Prediction Based on an Adaptive Kernel Density Estimation Algorithm

GUO Hengkuan, TIAN Jianyan, LIU Shuwei, JIANG Yingying, YANG Lizhi

引用本文:

郭恒宽, 田建艳, 刘竖威, 等. 基于自适应核密度估计算法的风电功率超短期区间预测[J]. 现代电力, 2026, 43(3): 499–508. DOI: 10.19725/j.cnki.1007–2322.2024.0048

GUO Hengkuan, TIAN Jianyan, LIU Shuwei, et al. Wind Power Ultra-short-term Interval Prediction Based on an Adaptive Kernel Density Estimation Algorithm[J]. Modern Electric Power, 2026, 43(3): 499–508. DOI: 10.19725/j.cnki.1007–2322.2024.0048

在线阅读 View online: <https://doi.org/10.19725/j.cnki.1007–2322.2024.0048>

您可能感兴趣的其他文章

Articles you may be interested in

基于自适应噪声完备经验模态分解样本熵长短期记忆神经网络和核密度估计的短期电力负荷区间预测

Short-term Power Load Interval Prediction Based on CEEMDAN–SE–LSTM and KDE

现代电力. 2021, 38(2): 138–146 <https://doi.org/10.19725/j.cnki.1007–2322.2020.0329>

基于注意力机制的卷积神经网络–长短期记忆网络的短期风电功率预测

Short-Term Wind Power Forecasting Based on Attention Mechanism of CNN–LSTM

现代电力. 2022, 39(2): 212–218 <https://doi.org/10.19725/j.cnki.1007–2322.2021.0108>

基于时序注意力机制的超短期风电功率概率预测

Probability Prediction of Ultra-short-term Wind Power Based on Temporal Pattern Attention

现代电力. 2023, 40(6): 906–913 <https://doi.org/10.19725/j.cnki.1007–2322.2022.0151>

基于随差遗忘长短期记忆的风电功率实时预测

Real-time Prediction of Wind Power Based on Error Following Forget Gate-based Long Short-term Memory

现代电力. 2021, 38(1): 110–118 <https://doi.org/10.19725/j.cnki.1007–2322.2020.0200>

基于模型无关优化策略的风电短时回归预测框架设计

A Regression Framework Design for Short Term Forecasting of Wind Power Based on Model-Agnostic Meta-Learning Strategy

现代电力. 2022, 39(3): 253–261 <https://doi.org/10.19725/j.cnki.1007–2322.2021.0097>

基于分时长短期记忆神经网络的光伏发电超短期功率预测

Ultra Short-term Forecast of Photovoltaic Generation Based on Time-division Long Short-term Memory Neural Networks

现代电力. 2020, 37(6): 629–637 <https://doi.org/10.19725/j.cnki.1007–2322.2019.1026>

基于自适应核密度估计算法的风电功率超短期区间预测

郭恒宽, 田建艳, 刘竖威, 蒋莹莹, 杨立志

(太原理工大学 电气与动力工程学院, 山西省 太原市 030024)

Wind Power Ultra-short-term Interval Prediction Based on an Adaptive Kernel Density Estimation Algorithm

GUO Hengkuan, TIAN Jianyan, LIU Shuwei, JIANG Yingying, YANG Lizhi

(College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan 030024, Shanxi Province, China)

摘要: 针对风电功率区间预测中核密度估计算法局部适应性不强的问题, 提出一种自适应核密度估计算法, 实现风电功率超短期区间预测。首先, 构建基于双重注意力机制改进序列到序列 (sequence to sequence, Seq2Seq) 架构的长短期记忆神经网络风电功率超短期预测模型, 对风电功率进行确定性预测。然后, 利用提出的一种自适应核密度估计算法, 对风电功率预测误差数据集进行核密度估计, 得到误差数据集的概率密度函数。最后, 通过对概率密度函数进行积分, 得到误差数据集的分布函数, 并根据分布函数求得不同置信度下的上侧和下侧分位数, 将分位数和风电功率确定性预测值求和, 实现相应置信度下的风电功率超短期区间预测。基于中国西北某风电场的数据进行仿真预测分析, 结果表明该自适应核密度估计算法对风电功率预测误差数据具有更强的适应能力, 验证了该风电功率区间预测模型的有效性。

关键词: 风电功率区间预测; 自适应核密度估计; Seq2Seq 架构; 注意力机制; 长短期记忆网络

Abstract: Aiming at the issue of insufficient local adaptability in kernel density estimation algorithm for wind power interval prediction, we propose an adaptive kernel density estimation algorithm for wind power ultra-short-term interval prediction. Firstly, a long short term-memory wind power ultra-short-term prediction model is constructed based on the improved sequence to sequence (Seq2Seq) architecture with a dual-stage attention mechanism for wind power deterministic prediction. Subsequently, an proposed adaptive kernel density estimation algorithm is utilized to estimate the kernel density of the historical wind power prediction error. The proposed algorithm is capable of adaptively adjusting the bandwidth according to

the kernel density estimation of the prediction error. Finally, the cumulative distribution function of the prediction error is obtained by integrating the kernel density estimation results. In addition, the upper and lower quantiles under different confidence coefficients are calculated, and the ultra-short-term prediction intervals of wind power under the corresponding confidence coefficients are obtained by adding the quantiles and the deterministic prediction values. Simulation analysis based on the data from a wind farm in Northwest China demonstrate that the proposed adaptive kernel density estimation algorithm exhibits superior adaptive ability to wind power prediction error data. Moreover, the effectiveness of this wind power interval prediction model is validated through these experiments.

Keywords: wind power interval prediction; adaptive kernel density estimation; Seq2Seq architecture; attention mechanism; long short term-memory (LSTM)

DOI: 10.19725/j.cnki.1007-2322.2024.0048

0 引言

风电功率预测能提供功率波动信息, 对电网安全稳定运行至关重要。近年来, 随着电网调度日益精细化, 超短期风电功率预测受到了广泛关注。基于深度学习的确定性预测方法是当前风电功率预测的研究热点, 文献 [1] 构建了长短期记忆 (long short term-memory, LSTM) 神经网络风电功率预测模型, 实验结果表明, 该模型相较于传统机器学习模型具有更高的预测精度。文献 [2] 构建了基于序列到序列 (sequence to sequence, Seq2Seq) 架构的 LSTM 电力负荷预测模型, 该模型相较于 LSTM 模型对输入数据的时序依赖关系具有更强的处理能力, 预测精度更高。文献 [3] 建立特征

基金项目: 山西省自然科学基金项目 (202303021221026)。
Project Supported by Natural Science Foundation of Shanxi Province (202303021221026).

注意力 (feature attention, FA) 机制为 LSTM 风电功率预测模型的输入特征赋权, 提升了 LSTM 模型的预测精度。文献 [4] 构建了基于卷积神经网络 (convolutional neural network, CNN) 和 LSTM 神经网络组合的风电功率预测模型 (CNN-LSTM), 并建立时序注意力 (temporal attention, TA) 机制为输入模型不同时刻的特征序列赋权, 提高了 CNN-LSTM 模型的预测精度。上述预测模型无法同时辨别输入模型的不同特征及不同时刻特征序列的重要性差异。文献 [5-6] 建立了基于双重注意力 (dual-stage attention, DA) 机制的风电功率预测模型, 然而其 DA 机制有待进一步改进。上述确定性预测模型的预测误差无法避免, 其预测结果无法对风电功率的不确定性进行定量描述 [7], 区间预测可以解决该问题。

常见的区间预测方法有分位数回归 (quantile regression, QR) 法 [8]、核密度估计 (kernel density estimate, KDE) 法 [9] 等。文献 [10] 建立了基于核极限学习机的分位数回归模型, 实现风电功率的区间预测。文献 [11] 建立了基于支持向量机 (support vector machine, SVM) 的分位数回归模型, 对风电功率进行区间预测。然而, 上述分位数回归模型无法计算风电功率预测误差的概率密度分布或风电功率预测值的概率密度分布。文献 [12] 采用 KDE 算法对风电功率预测误差进行核密度估计, 证明了相较于参数化建模方法, KDE 算法不需要对预测误差的分布进行预先假设, 对误差数据具有更强的鲁棒性。文献 [13] 构建了双向长短期记忆神经网络 (bidirectional long short-term memory, BiLSTM) 风电功率预测模型, 并将其与 KDE 算法结合, 对风电功率进行区间预测。文献 [14] 构建了 LSTM 负荷预测模型, 并将 LSTM 预测模型与 QR 算法、KDE 算法结合, 对负荷进行概率密度预测。文献 [15] 为了提升 KDE 算法的精度, 利用拇指法则计算 KDE 算法的全局最优带宽, 并将该算法与 BiLSTM 结合, 对风电功率进行区间预测。文献 [16] 采用无偏交叉验证算法优化 KDE 算法的带宽, 提升了 KDE 算法的精度。文献 [17] 基于最小化均方积分误差 (mean integrated squared error, MISE) 优化 KDE 算法的带宽, 相较于随机带宽的 KDE 算法, 该方法具有更高的风电功率区间预测精度。然而, 上述 KDE 算法的带宽无法随数据点的核密度估计值自适应调整, 因此 KDE 算法的局部适应性不强, 导致预测结果的精度有

待提升。

针对上述问题, 本文提出一种基于自适应 KDE 算法的风电功率区间预测模型。首先, 对风电功率进行确定性预测; 其次, 构建风电功率预测误差数据集; 然后, 利用本文提出的自适应 KDE 算法对误差数据集进行核密度估计; 最后, 对核密度估计结果进行处理, 得到相应置信度下的风电功率超短期预测区间。

1 风电功率区间预测模型

1.1 风电功率预测误差数据集

风电功率区间预测需要对风电功率确定性预测模型的误差数据集进行概率密度估计, 本文采用比例划分的方法生成风电功率预测误差数据集, 误差数据集的构造如图 1 所示。

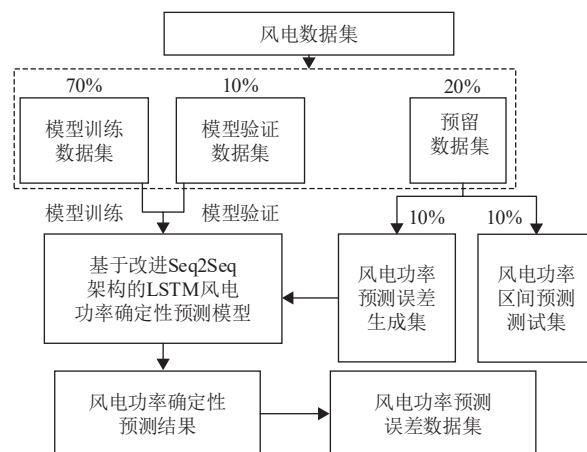


图 1 风电功率预测误差数据集的构建方法

Fig. 1 Construction method for wind power prediction errors datasets

由图 1 可知, 本文将风电数据集按照 7:1:2 的比例划分为三部分:

1) 模型训练数据集。将按时间顺序升序的风电数据集前 70% 的数据划分为模型训练集, 用于对风电功率确定性预测模型进行训练。

2) 模型验证数据集。将按时间顺序升序的风电数据集中间 10% 的风电数据划分为模型验证集, 用于增强风电功率确定性预测模型的泛化性能, 防止模型出现过拟合或欠拟合的问题。

3) 预留数据集。将按时间顺序升序的风电数据集后 20% 的数据作为风电功率预测的预留数据集。预留数据集的前一部分数据作为风电功率确定性预测误差生成集, 将该部分数据输入风电

功率确定性预测模型进行预测，并求得风电功率预测值和真实值之间的差值，从而生成风电功率预测误差数据集；预留数据集的后一部分数据划分为风电功率区间预测的测试集，将测试集输入风电功率确定性预测模型，得到风电功率预测值，并将预测值与预测误差数据集的上下分位数求和，得到风电功率的区间预测。

1.2 风电功率超短期确定性预测模型

考虑到风电功率具有较强的波动性，且模型的输入特征间存在耦合，本文搭建基于 Seq2Seq 架构的 LSTM(Seq2Seq-LSTM) 风电功率确定性预测模型，从而提升 LSTM 的性能。针对模型无法识别不同输入特征重要性差异的问题，建立特征注意力 (feature attention, FA) 机制，对模型的编码侧进行改进；因为模型无法识别不同时刻下特征序列的重要性差异，所以建立时序注意力 (temporal attention, TA) 机制对模型的解码侧进行改进。该风电功率确定性预测模型的结构如图 2 所示。

图 2 的输入层中： n 为模型输入数据的特征数量； T 为模型输入数据的时间窗口大小。通过 FA 机制对不同的输入特征赋权，使得模型编码侧能够识别输入特征的重要性差异。特征权重的计算如式 (1) 所示。

$$j_{\theta}^{(u)} = \mathbf{V}_z^T \tanh(\mathbf{W}_z[\mathbf{h}_{\theta-1}; \mathbf{s}_{\theta-1}] + \mathbf{U}_z \mathbf{x}^{(u)} + \mathbf{B}_z) \quad (1)$$

式中： $\mathbf{V}_z$ 、 $\mathbf{W}_z$ 、 $\mathbf{U}_z$ 、 $\mathbf{B}_z$ 为 FA 机制的参数矩阵； $\mathbf{h}_{\theta}$ 、 $\mathbf{s}_{\theta}$ ($\theta=1,2,\dots,T$) 分别为编码侧 LSTM 的隐含层状态和单元状态； $\mathbf{x}^{(u)}$ ($u=1,2,\dots,n$) 为第 u 个特征构成的特征数据； $j_{\theta}^{(u)}$ 为第 u 个特征在时刻 θ 的特征权重。

对特征权重归一化，如式 (2) 所示。

$$\alpha_{\theta}^{(u)} = \exp(j_{\theta}^{(u)}) / (\sum_{i=1}^n \exp(j_{\theta}^{(i)})) \quad (2)$$

利用特征权重对输入特征赋权，得到赋权的特征数据 $\tilde{\mathbf{x}}_{\theta}$ ，如式 (3) 所示。

$$\tilde{\mathbf{x}}_{\theta} = [\mathbf{x}_{\theta}^{(1)} \alpha_{\theta}^{(1)}, \mathbf{x}_{\theta}^{(2)} \alpha_{\theta}^{(2)}, \dots, \mathbf{x}_{\theta}^{(n)} \alpha_{\theta}^{(n)}] \quad (3)$$

将 $\tilde{\mathbf{x}}_{\theta}$ 输入编码侧 LSTM 网络，得到编码侧 LSTM 网络的隐含层状态 $\mathbf{h}_{\theta}$ ($\theta=1,2,\dots,T$) 为

$$\mathbf{h}_{\theta} = \mathbf{f}(\mathbf{h}_{\theta-1}, \tilde{\mathbf{x}}_{\theta}) \quad (4)$$

式中 $\mathbf{f}$ 为编码侧 LSTM 网络。 $\mathbf{h}_{\theta}$ ($\theta=1,2,\dots,T$) 作为编码侧的输出，又作为解码侧的输入。

通过 TA 机制对编码侧的输出 $\mathbf{h}_{\theta}$ ($\theta=1,2,\dots,T$) 赋权，使得解码侧能识别编码侧不同时刻输出的

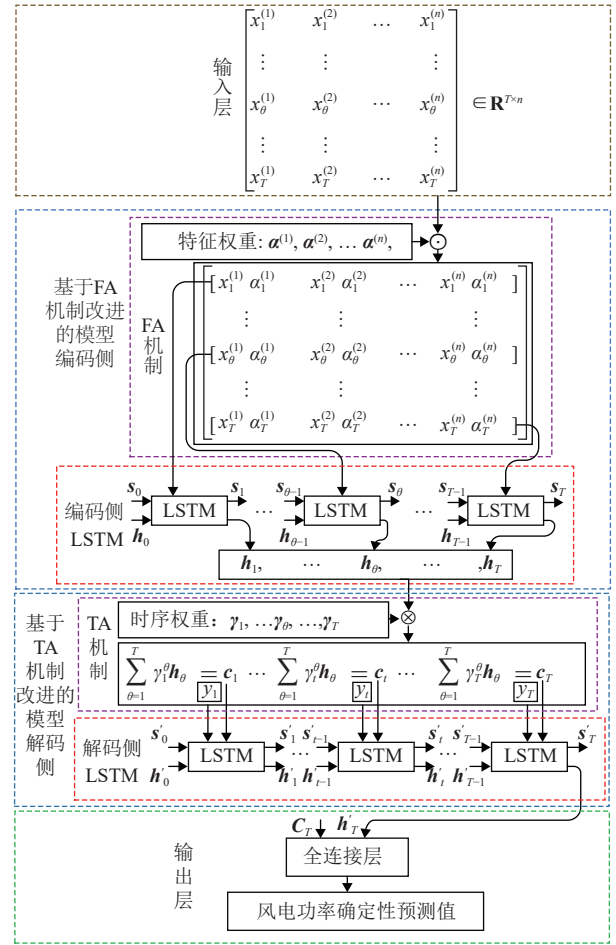


图 2 风电功率确定性预测模型结构

Fig. 2 Structure of wind power deterministic prediction model

重要性差异。由于 $\mathbf{h}_{\theta}$ 为输入模型的特征序列 $\mathbf{x}_{\theta}$ ($\theta=1,2,\dots,T$) 的映射，故 TA 机制辨别了不同时刻特征序列 $\mathbf{x}_{\theta}$ 的重要性差异，时序权重的计算式为

$$\varphi_t^{\theta} = \mathbf{V}_d^T \tanh(\mathbf{U}_d[\mathbf{h}_{t-1}'; \mathbf{s}_{t-1}'] + \mathbf{P}_d \mathbf{h}_{\theta} + \mathbf{B}_d) \quad (5)$$

式中： $\mathbf{V}_d^T$ 、 $\mathbf{U}_d$ 、 $\mathbf{P}_d$ 和 $\mathbf{B}_d$ 为 TA 机制的参数矩阵； $\mathbf{h}_{\theta}$ 为编码侧 LSTM 的隐含层状态； $\mathbf{h}_{t-1}'$ 、 $\mathbf{s}_{t-1}'$ 分别为解码侧 LSTM 的隐含层状态和单元状态； φ_t^{θ} 为 $\mathbf{h}_{\theta}$ 的时序权重。

对时序权重归一化处理，如式 (6) 所示。

$$\gamma_t^{\theta} = \exp(\varphi_t^{\theta}) / (\sum_{j=1}^T \exp(\varphi_t^j)) \quad (6)$$

将编码侧的输出与对应的时序权重加权求和，得到当前时刻 t 的风电融合数据 $\mathbf{c}_t$ ，如式 (7) 所示。

$$\mathbf{c}_t = \sum_{\theta=1}^T \mathbf{h}_{\theta} \gamma_t^{\theta} \quad (7)$$

将 c_t 和历史风电功率 y_t 融合,得到解码侧 LSTM 的输入 $\tilde{y}_t$,如式(8)所示。

$$\tilde{y}_t = W_o([c_t; y_t] + b_o) \quad (8)$$

式中 W_o 、 b_o 为数据融合的参数矩阵。

将 $\tilde{y}_t$ 输入解码侧 LSTM,得到解码侧 t 时刻的隐含层输出 h'_t ,如式(9)所示。

$$h'_t = f'(h'_{t-1}, \tilde{y}_t) \quad (9)$$

式中 f' 为解码侧 LSTM 网络。

将解码侧最后时刻的 LSTM 隐含层状态 h'_T 和风电融合数据 c_T 输入至输出层,得到风电功率预测值 $\hat{y}_{T+\Delta t}$,如式(10)所示。

$$\hat{y}_{T+\Delta t} = V_c(W_T[h'_T; c_T] + b_T) + b_c \quad (10)$$

式中: V_c 、 W_T 、 b_T 和 b_c 为输出层的参数矩阵; Δt 为预测的时间尺度。

1.3 风电功率预测误差的固定带宽核密度估计

风电功率概率预测方法主要包括参数建模算法和非参数建模算法。参数建模算法需要对输入数据的分布进行先验假设,而输入数据具有随机性和不确定性,其分布形式通常难以假设。KDE 算法属于非参数建模算法,该算法以误差数据集中的任意风电功率预测误差点为中心,并结合带宽参数建立核函数,将所有误差点的核函数求和并取平均值,最终得到误差数据集的概率密度函数,误差数据集中的任一误差点 e 处的概率密度函数如式(11)所示。

$$\hat{f}(e) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{e-e_i}{h}\right) \quad (11)$$

式中: n 为风电功率预测误差点的数量; h 为 KDE 算法的带宽; e_i 为误差数据集中第 i 个误差点; K 为核函数。

由式(11)可知,该方法的精度完全取决于带宽 h 和核函数 K 。文献[18]指出核函数的选取对核密度估计精度的影响较小,因此本文选取常用的高斯核函数,其表达式如式(12)所示。

$$K(e) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{e^2}{2}\right) \quad (12)$$

基于高斯核函数的风电功率误差数据集的概率密度函数如式(13)所示。

$$\hat{f}(e) = \frac{1}{\sqrt{2\pi}nh} \sum_{i=1}^n \exp\left(-\frac{(e-e_i)^2}{2h^2}\right) \quad (13)$$

然而,对于固定带宽的风电功率预测误差 KDE 算法,其带宽无法根据误差点的核密度估计值自适应调整,对误差点的局部适应性较差,主要体现为对误差点分布较为集中的区域容易造成欠拟合的现象,而对误差点分布较为分散的区域容易造成过拟合的现象。

1.4 风电功率预测误差的自适应核密度估计方法

针对上述风电功率预测误差固定带宽 KDE 算法存在的问题,本文提出一种自适应 KDE 算法,模型在风电功率预测误差点较为集中的区域可以自适应减小 KDE 算法的带宽,在误差点较为分散的区域可以自适应增大 KDE 算法的带宽,从而更加准确地估计误差点的概率密度分布,具体方法如式(14)~(28)所示。

对于误差数据集中任一误差点 e ,其核密度估计函数值 $\hat{f}(e)$ 的 MISE 为

$$\hat{f}_{\text{MISE}}(e) = E\left\{\int [(\hat{f}(e) - f(e))^2]de\right\} = \int \text{var}(\hat{f}(e))de + \int [E(\hat{f}(e)) - f(e)]^2de \quad (14)$$

式中: var 为方差; $\hat{f}(e)$ 为误差点的核密度估计值,即误差点的概率密度估计值; $f(e)$ 为误差点的真实概率密度值。

对式(14)分解后的第一项进行求解,可得

$$\begin{aligned} \int \text{var}(\hat{f}(e))de &= \frac{1}{n^2h^2} \int \text{var}\left(\sum_{i=1}^n K\left(\frac{e-e_i}{h}\right)\right)de \approx \\ &= \frac{1}{n^2h^2} \int E\left(\sum_{i=1}^n K^2\left(\frac{e-e_i}{h}\right)\right)de = \\ &= \frac{1}{nh^2} \int \int K^2\left(\frac{e-e_i}{h}\right)f(e_i)de_ide \end{aligned} \quad (15)$$

$f(e_i)$ 在其积分区间积分值为1,令 $\frac{e-e_i}{h}=u$,式(15)可以进行如下转换

$$\int \text{var}(\hat{f}(e))de = \frac{1}{nh} \int K^2(u)du \quad (16)$$

对式(14)分解后的第二项进行求解,计算式为

$$\begin{aligned} \int [E(\hat{f}(e)) - f(e)]^2de &= \\ \int \left[\frac{1}{nh} E\left(\sum_{i=1}^n K\left(\frac{e-e_i}{h}\right)\right) - f(e)\right]^2de &= \\ \int \left[\frac{1}{h} \int K\left(\frac{e-e_i}{h}\right)f(e_i)de_i - f(e)\right]^2de &= \end{aligned} \quad (17)$$

令 $\frac{e-e_i}{h}=u$,式(17)可以进行如下转换

$$\int [E(\hat{f}(e)) - f(e)]^2 de = \int \left[\int K(u)f(e-hu)du - f(e) \right]^2 de \quad (18)$$

令 $f(e-hu)$ 在 e 点泰勒展开, 式 (18) 可表示为

$$\int [E(\hat{f}(e)) - f(e)]^2 de = \int \left[\int K(u)(f(e) - huf^{(1)}(e) + h^2u^2 \frac{f^{(2)}(e)}{2} du - f(e)) \right]^2 de \quad (19)$$

因为 $\int uK(u)du=0$, 式 (19) 可转换为

$$\int [E(\hat{f}(e)) - f(e)]^2 de = \frac{h^4}{4} \left(\int K(u)u^2 du \right)^2 \int f^{(2)}(e) de \quad (20)$$

综上所述, $\hat{f}_{MISE}(e)$ 的求解结果为

$$\hat{f}_{MISE}(e) = \frac{1}{nh} \int K^2(u)du + \frac{h^4}{4} \left(\int K(u)u^2 du \right)^2 \int f^{(2)}(e) de \quad (21)$$

为了求得式 (21) 的最小值, 对式 (21) 关于 h 求一阶导数, 并令其为 0, 得到全局最优带宽 h_g 的表达式为

$$h_g = \left[\frac{\int K^2(u)du}{n \int (f^{(2)}(e))^2 de \left(\int K(u)u^2 du \right)^2} \right]^{0.2} \quad (22)$$

然而, h_g 无法根据误差数据集中误差点的概率密度值 $f(e)$ 自适应调整, 部分数据点的核密度估计效果仍不能令人满意, 故需要对带宽做进一步优化。

误差点核密度估计的均方误差 (mean square error, MSE) 计算公式为

$$\hat{f}_{MSE}(e) = [E(\hat{f}(e)) - f(e)]^2 + \text{var}(\hat{f}(e)) = \frac{h^4}{4} [f^{(2)}(e)]^2 \left[\int K(u)u^2 du \right]^2 + \frac{f(e)}{nh} \int K^2(u)du \quad (23)$$

对式 (23) 关于 h 求导, 并令其为 0, 得到自适应带宽 h_* 的表达式为

$$h_* = \left[\frac{f(e) \int K^2(u)du}{n(f^{(2)}(e))^2 \left(\int K(u)u^2 du \right)^2} \right]^{0.2} \quad (24)$$

将式 (24) 转换为基于全局最优带宽 h_g 的形式, 如式 (25) 所示。

$$h_* = h_g \frac{f^{0.2}(e) \left(\int (f^{(2)}(e))^2 de \right)^{0.2}}{[f^{(2)}(e)]^{0.4}} \quad (25)$$

由式 (25) 可知, h_* 能够根据每一个误差点 e 进行优化, 能够随误差点的概率密度值 $f(e)$ 自适应

调整。然而, $f(e)$ 未知, 故式 (25) 无法求解。概率密度函数 $\hat{f}(e)$ 相较于假设的分布更加接近误差点的真实概率密度分布, 故本文利用概率密度函数 $\hat{f}(e)$ 替换 $f(e)$, 对式 (25) 求解, 求解过程如式 (26)~(27)。

将 h_g 代入式 (13), 对式 (13) 求二阶导数, 得到基于带宽 h_g 的概率密度函数 $\hat{f}_g(e)$ 的二阶导数为

$$\hat{f}_g^{(2)}(e) = \frac{1}{nh_g^3} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}} (v^2 \exp(-\frac{v^2}{2}) - \exp(-\frac{v^2}{2})) \quad (26)$$

式中 $v = \frac{e - e_i}{h_g}$ 。

将式 (26) 代入式 (25) 中, 可得 h_* 为

$$h_* = \frac{0.822\pi^{-0.1} h_g^{0.8} \hat{f}_g^{0.2}(e)}{\left(\frac{1}{nh_g} \sum_{i=1}^n v^2 K(v) - \hat{f}_g(e) \right)^{0.4}} \quad (27)$$

基于自适应带宽的概率密度函数可表示为

$$\hat{f}_*(e) = \frac{1}{\sqrt{2\pi}nh_*} \sum_{i=1}^n \exp\left(-\frac{(e - e_i)^2}{2h_*^2}\right) \quad (28)$$

式中 $\hat{f}_*(e)$ 为基于自适应带宽的概率密度函数。

对概率密度函数进行积分, 得到误差数据集的分布函数 $\hat{F}_*(e)$, 如式 (29) 所示。

$$\hat{F}_*(e) = \int \frac{1}{\sqrt{2\pi}nh_*} \sum_{i=1}^n \exp\left(-\frac{(e - e_i)^2}{2h_*^2}\right) de \quad (29)$$

令分布函数 $\hat{F}_*(e)$ 分别等于 $1 - \frac{\alpha}{2}$ 和 $\frac{\alpha}{2}$, 求得置信度 $1 - \alpha$ 下的上侧分位数 $e_{1-\alpha/2}$ 和下侧分位数 $e_{\alpha/2}$, 并将上下侧分位数与风电功率确定性预测结果求和, 得到相应置信度下的风电功率预测区间, 其表达式如式 (30) 所示。

$$\begin{cases} U_i = \hat{p}_i + e_{1-\alpha/2} \\ D_i = \hat{p}_i + e_{\alpha/2} \end{cases} \quad (30)$$

式中: U_i 为风电功率预测区间的上限; D_i 为风电功率预测区间的下限; $\hat{p}_i$ 为风电功率的确定性预测值; α 为预测区间的显著性水平。

2 风电功率区间预测流程及精度评价指标

2.1 风电功率区间预测流程

本文构建的风电功率区间预测模型的预测流程如下:

1) 风电数据预处理。模型的输入数据存在

一些异常值和缺失值,利用箱线图法剔除异常值,并用前后相邻时刻数据的均值填充缺失值。然后,将输入模型的各个特征归一化,从而消除不同特征的量纲对预测精度的影响,数据的归一化如式(31)所示。

$$x_{\text{scaler}} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (31)$$

式中: x 为某一输入特征的原始值, $\min(x)$ 为该特征的最小值; $\max(x)$ 为该特征的最大值; x_{scaler} 为该特征归一化后的数据。

2) 风电功率确定性预测。本文建立了基于 DA 机制改进 Seq2Seq 架构的 LSTM 风电功率预测模型,首先构建 Seq2Seq 架构的 LSTM 模型,然后构建 FA 机制和 TA 机制对模型进行改进,模型能够同时对输入模型的不同特征及不同时刻的特征序列动态赋权,最终得到风电功率的确定性预测结果。

3) 风电功率区间预测。利用本文提出的自适应 KDE 算法对风电功率预测误差数据集进行核密度估计,求得不同置信度下的上侧和下侧分位数,并与风电功率确定性预测结果求和,得到风电功率的预测区间。

2.2 风电功率预测区间精度评价指标

本文构建的风电功率区间预测模型利用 Adam 算法训练模型,为了对模型的预测精度进行评估,选取平均绝对误差 (mean absolute error, MAE) 指标 e_{MAE} 、均方根误差 (root mean square error, RMSE) 指标 e_{RMSE} 、决定系数 R^2 作为确定性预测模型的精度评估指标;选取区间覆盖率 (prediction interval coverage probability, PICP) X_{PICP} 、区间平均宽度 (prediction interval average width, PINAW) X_{PINAW} 和温克勒得分 (winker score, WS) X_{WS} 作为区间预测的精度评估指标。各预测区间精度评价指标的计算公式为

$$X_{\text{PICP}} = \frac{1}{n} \sum_{i=1}^n \zeta_i \times 100\%, \zeta_i = \begin{cases} 0, & p_i \notin [U_i, D_i] \\ 1, & p_i \in [U_i, D_i] \end{cases} \quad (32)$$

$$X_{\text{PINAW}} = \frac{1}{n} \sum_{i=1}^n (U_i - D_i) \quad (33)$$

$$W_i = \begin{cases} U_i - D_i + 2(D_i - p_i)/\alpha, & p_i < D_i \\ U_i - D_i, & D_i \leq p_i \leq U_i \\ U_i - D_i + 2(p_i - U_i)/\alpha, & p_i > U_i \end{cases} \quad (34)$$

$$X_{\text{WS}} = \frac{1}{n} \sum_{i=1}^n W_i \quad (35)$$

式中: n 为误差点的数量; U_i 为预测区间的上限; D_i 为预测区间的下限; p_i 为风电功率真实值; α 为预测区间的显著性水平; X_{PICP} 为风电功率真实值被包含在预测区间内的比例,用于衡量预测区间的可靠程度; X_{PINAW} 为风电功率预测值波动区间的大小,衡量了预测区间预测信息的细致程度。

由上述公式可知, X_{WS} 结合了区间覆盖率和区间平均宽度两方面综合因素,对区间预测的质量进行综合评价^[19]。高质量的风电功率区间预测同时具备了高 X_{PICP} 和窄 X_{PINAW} ,但是这两个评估指标难以平衡,故本文以综合评价指标 X_{WS} 作为区间预测效果的主要评价指标,相同置信度下, X_{WS} 越小,代表模型的预测效果越好。

3 算例分析

3.1 仿真平台

为了验证本文构建的风电功率超短期区间预测模型的有效性,本文基于中国西北某风电场 2015 年 1 月 1 日至 2017 年 12 月 31 日的实测数据进行仿真分析,该风电场的装机容量为 129 MW,数据的采样间隔为 15 min,选取风向、轮毂高度风速、温度、气压、湿度以及历史风电功率作为模型的输入特征。预测模型基于 Python 3.7 搭建,实验平台配置为 Intel(R)Core(TM) i7-8700 CPU@3.2 GHz,操作系统为 Windows 10。

3.2 风电功率超短期确定性预测算例分析

为了验证本文构建的确定性预测模型的可靠性,利用该模型预测未来 15 min 的风电功率,并将其与时间卷积网络 (temporal convolutional network, TCN)-LSTM 模型、LSTM 模型、GRU 模型、SVR 模型进行对比。选取风电功率波动较为剧烈的特殊天气时段绘制预测曲线,如图 3 所示。

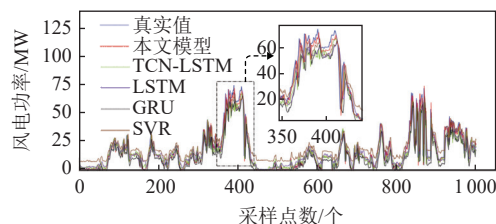


图 3 不同风电功率预测模型预测结果

Fig. 3 Prediction results of different wind power prediction models

由于图 3 中风电功率实测值的测量时间为 12 月，该月份风速波动较大，因此风电功率容易出现爬坡现象。由图 3 可知，本文构建的风电功率预测模型的预测结果相较于其他对比模型，在风电功率较为平稳和风电功率剧烈爬坡时都更加接近真实值。而 TCN-LSTM 模型、GRU 模型、LSTM 模型在风电功率波动较为剧烈时预测精度低于本文构建模型。与上述深度学习模型相比，SVR 模型的风电功率预测精度较低。

上述风电功率预测模型的全年预测误差见表 1。

表 1 不同确定性风电功率预测模型预测误差
Table 1 Prediction errors of different deterministic wind power prediction models

预测模型	$R^2/\%$	e_{RMSE}/MW	e_{MAE}/MW
本文确定性预测模型	95.2	4.461	2.217
TCN-LSTM	93.1	5.498	3.205
LSTM	92.3	6.105	3.944
GRU	91.6	6.345	4.319
SVR	88.2	7.576	6.503

由表 1 可知：本文构建的风电功率预测模型的预测精度最高，这可以归因于该模型构建了 Seq2Seq 架构，并用 DA 机制进行改进；TCN-LSTM 模型通过 TCN 网络对模型输入的风电数据进行特征提取，相较于 LSTM 模型和 GRU 模型具有更高的预测精度；SVR 模型在处理大规模非线性的风电数据时拟合能力不足，预测精度较低。

综上所述，本文构建的风电功率确定性预测模型的鲁棒性较强，在常规及特殊天气条件下均具有较高的预测精度，验证了本文所构建的确定性预测模型的有效性。

3.3 风电功率超短期区间预测算例分析

本节用到的风电功率确定性预测值与预测误差数据集，均来自 1.1 节中确定性预测模型预测未来 15 min 的风电功率及其预测误差。针对上述固定带宽 KDE 算法存在的问题，本文在 1.4 节中提出了一种自适应 KDE 算法。该算法的核心在于带宽能随着误差点的概率密度估计值自适应调整，具体机制如图 4 所示。

由图 4 可知，本文提出的自适应 KDE 算法的带宽 h 随着误差点概率密度估计值的增大而自适应减小，随着误差点概率密度估计值的减小而自适应增大。当 h 值减小时，核函数对误差点的概

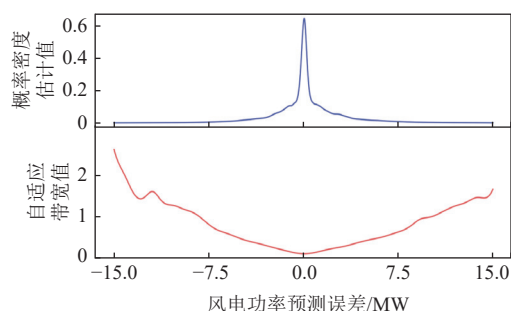


图 4 带宽随概率密度估计值的自适应调整

Fig. 4 Adaptive adjustments of bandwidth varying with probability density estimation

率密度分布具有更高的分辨率，能够分辨相近误差点概率密度的差异，适合处理分布较为集中的误差点；当 h 值增大时，核函数对误差点概率密度分布的分辨率降低，适合揭示分布较为分散误差点的概率密度分布的宏观特征，故该自适应带宽的调整策略符合误差点核密度估计的实际需求。

为了验证本文提出的预测误差自适应 KDE 算法的有效性，利用不同带宽下的核密度估计算法对误差点进行核密度估计，求得误差点的概率密度估计值，如图 5 所示。

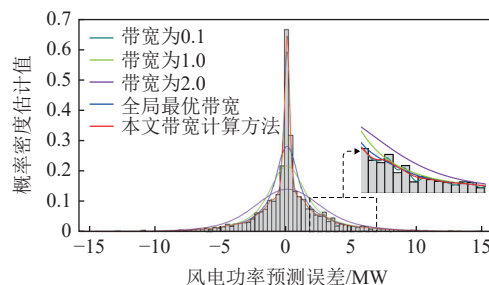


图 5 不同带宽下预测误差的概率密度估计

Fig. 5 Probability density estimation of prediction errors under different bandwidths

图 5 中频率分布直方图为风电功率预测误差的真实概率密度分布。由图 5 可知：当误差点的概率密度估计值较大时，自适应 KDE 算法概率密度函数的概率密度估计值相较于其他对比算法更加贴近真实的概率密度值；当误差点概率密度较小时，自适应 KDE 算法的概率密度函数更加光滑，能够揭示误差点概率密度分布的宏观特征。其他对比算法对于误差数据的局部适应性较差：在误差点概率密度较大的区域，其概率密度函数的误差较大，算法出现了欠拟合的问题，带宽值较大的核密度估计算法欠拟合问题尤为明显；在误差点概率密度较小的区域，带宽值为 0.1 的 KDE

算法的概率密度函数不够光滑,算法出现了过拟合的问题。通过上述不同算法的对比,验证了本文提出的预测误差自适应 KDE 算法的有效性。

对基于自适应 KDE 算法的误差数据集概率密度函数进行积分,求得不同置信度下误差数据的上下分位数,并将上下分位数与风电功率确定性预测结果求和,得到相应置信度下风电功率的预测区间。80%~95% 置信度下的风电功率区间预测常被用于电网的调度与调控^[17],故本文展示了 0.95、0.9、0.85、0.8 置信度下的风电功率预测区间,如图 6 所示。

由图 6 可知,当置信度较高时,除了个别的风电功率波峰和波谷点,风电功率真实值基本被包含在本文构建的风电功率区间预测模型的预测区间内。并且,预测区间的宽度随着置信度的减小而逐渐变小,因为预测区间越小,风电功率真实值被包含在预测区间的概率越小,置信度也就越低。本文构建的模型通过预测风电功率的波动区间,避免了确定性预测模型无法对风电功率的不确定性做出定量描述的问题^[7]。

为了量化本文所构建的风电功率区间预测模型的预测精度,利用 2.2 节提到的区间预测精度评价指标分别对 0.95、0.9、0.85、0.8 置信度下的预测结果进行评估,各置信度下的预测精度见表 2。

由表 2 可知:风电功率预测区间的区间覆盖率满足预设置信度的要求,且随着置信度的降低,预测区间的宽度逐渐减小,预测模型能够提供更加细致的风电功率波动信息;随着预测区间宽度的增大,预测区间的覆盖率逐渐升高,预测结果的可信度增大。由式 (34)—(35) 可知, X_{WS} 的值受预测区间宽度的影响,随预测区间宽度的减小而逐渐减小。

为了进一步验证本文构建的风电功率区间预测模型的有效性,将其与全局最优带宽、固定带宽为 0.1、固定带宽为 1 和固定带宽为 2 的区间预测模型进行对比。0.95、0.9、0.85、0.8 置信度下的风电功率区间预测精度如表 3 所示。

由表 3 可知:在预设置信度下,本文构建的风电功率区间预测模型相较于其他对比模型, X_{PICP} 指标更好,虽然在一定程度上增加了预测区间的宽度,但区间覆盖率满足预设置信度的要求,预测区间的可靠性更强;而对比模型的预测区间虽然相较本文模型更窄,但其 X_{PICP} 指标偏离预设

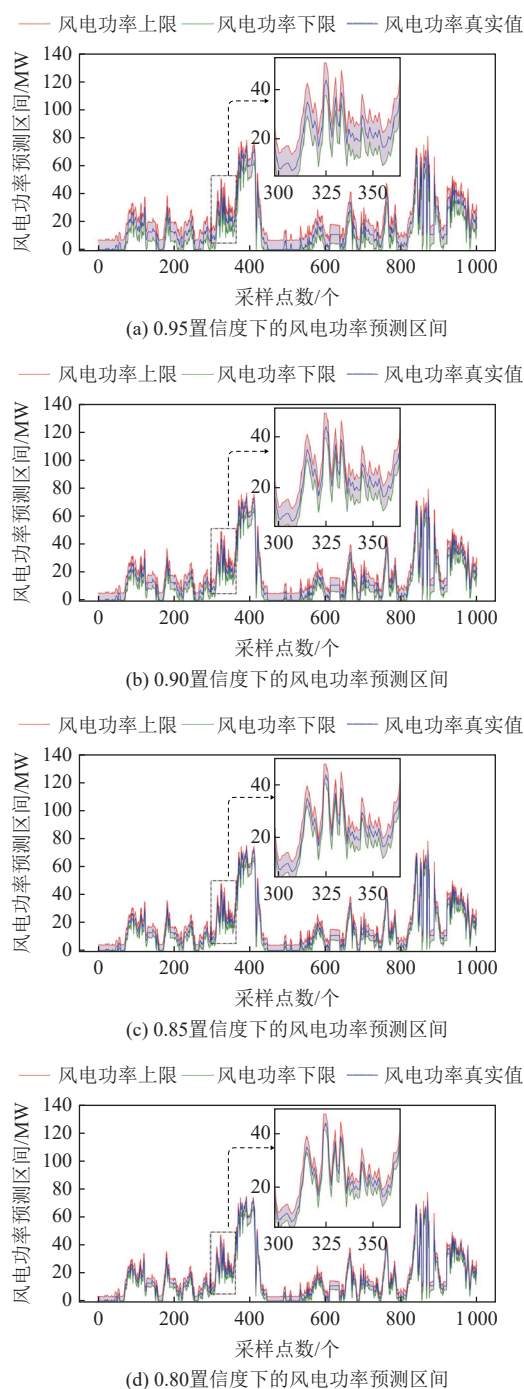


图 6 不同置信度下的风电功率预测区间

Fig. 6 Wind power prediction intervals under different confidence levels

置信度的程度较大,预测区间的可信度不足,这将给电力系统的调度带来安全隐患;在相同置信度下,本文模型的 X_{WS} 指标相较于对比模型更小,尤其当预设置信度较高时,该优势更加明显。基于该算法的区间预测模型相较于其他对比模型,在 0.95 置信度下的 X_{WS} 指标分别降低了 13.78%、

表 2 本文模型的风电功率区间预测精度
Table 2 Accuracy of wind power interval prediction by the proposed model

置信度	评估指标		
	$X_{PICP}/\%$	X_{PINAW}/MW	X_{WS}/MW
0.95	95.06	10.99	27.84
0.90	89.24	8.22	19.91
0.85	85.19	6.44	16.31
0.80	80.48	5.19	14.03

表 3 不同方法的风电功率区间预测精度
Table 3 Accuracy of wind power interval prediction by different methods

置信度	预测方法	评估指标		
		$X_{PICP}/\%$	X_{PINAW}/MW	X_{WS}/MW
0.95	本文方法	95.06	10.99	27.84
	全局最优带宽	87.93	8.09	32.29
	固定带宽为0.1	87.78	7.95	32.66
	固定带宽为1	88.27	8.28	31.86
	固定带宽为2	90.58	9.29	30.02
0.90	本文方法	89.24	8.22	19.91
	全局最优带宽	83.61	6.42	21.23
	固定带宽为0.1	83.19	6.27	21.40
	固定带宽为1	84.27	6.63	21.00
	固定带宽为2	87.47	7.76	20.12
0.85	本文方法	85.19	6.44	16.31
	全局最优带宽	78.96	5.30	16.83
	固定带宽为0.1	78.36	5.21	16.90
	固定带宽为1	80.05	5.52	16.68
	固定带宽为2	84.96	6.75	16.14
0.80	本文方法	80.48	5.19	14.03
	全局最优带宽	74.74	4.44	14.29
	固定带宽为0.1	73.79	4.28	14.37
	固定带宽为1	76.05	4.70	14.16
	固定带宽为2	82.02	5.98	13.80

14.76%、12.62%、7.26%；在 0.9 置信度下的 X_{WS} 指标分别降低了 6.22%、6.96%、5.19%、1.04%。 X_{WS} 指标作为本文的主要评价指标，说明本文模型的预测区间相较于其他对比模型具有更高的综合质量和预测精度。综上所述，本文构建的风电功率区间预测模型相较于对比模型，预测结果更加可靠，验证了该模型的有效性。

4 结论

1) 本文构建了基于 DA 机制改进 Seq2Seq 架

构的 LSTM 风电功率超短期确定性预测模型，对输入数据具有更强的特征提取能力和时序依赖关系处理能力，该模型相较于对比模型具有更高的预测精度，在常规及特殊天气条件下具有较高的稳定性。

2) 本文提出了一种针对风电功率区间预测的自适应 KDE 算法，该算法求得的预测误差概率密度估计值相较于其他对比算法更加接近真实的概率密度值。基于该算法的区间预测模型相较于对比模型，在 0.95 置信度下的温克勒得分指标分别降低了 13.78%、14.76%、12.62%、7.26%；在 0.9 置信度下的温克勒得分指标分别降低了 6.22%、6.96%、5.19%、1.04%。

3) 相较于其他对比模型，本文构建的风电功率区间预测模型的区间覆盖率能够满足预设置信度的要求，其预测区间具有更高的综合质量和可信度。

4) 由于环境气象等因素的不断变化，风电功率预测模型会出现模型漂移问题。因此，在后续研究中，将通过数字孪生技术解决风电功率预测模型的漂移问题，进一步提升模型的预测精度。

参考文献

[1] 朱乔木, 李弘毅, 王子琪, 等. 基于长短期记忆网络的风电场发电功率超短期预测 [J]. 电网技术, 2017, 41(12): 3797–3802.
ZHU Qiaomu, LI Hongyi, WANG Ziqi, *et al.* Short-term wind power forecasting based on LSTM[J]. Power System Technology, 2017, 41(12): 3797–3802(in Chinese).
[2] MU Y, WANG M, ZHENG X, *et al.* An improved LSTM-Seq2Seq-based forecasting method for electricity load[J]. Frontiers in Energy Research, 2023, 10(1): 1–13.
[3] XIONG B, LOU L, MENG X, *et al.* Short-term wind power forecasting based on attention mechanism and deep learning[J]. Electric Power Systems Research, 2022, 206(5): 1–9.
[4] 姚越, 刘达. 基于注意力机制的卷积神经网络-长短期记忆网络的短期风电功率预测 [J]. 现代电力, 2022, 39(2): 212–218.
YAO Yue, LIU Da. Short-term wind power forecasting based on attention mechanism of CNN-LSTM[J]. Modern Electric Power, 2022, 39(2): 212–218(in Chinese).
[5] MENG A, CHEN S, OU Z. *et al.* A novel few-shot learning approach for wind power prediction applying second-

- ary evolutionary generative adversarial network[J]. *Energy*, 2022, 261(10): 1–15.
- [6] 苏向敬, 周汶鑫, 李超杰, 等. 基于双重注意力 LSTM 神经网络的可解释海上风电出力预测 [J]. *电力系统自动化*, 2022, 46(7): 141–151.
- SU Xiangjing, ZHOU Wenxin, LI Chaojie, *et al.* Interpretable offshore wind power output forecasting based on dual-stage attention long short-term memory[J]. *Automation of Electric Power Systems*, 2022, 46(7): 141–151(in Chinese).
- [7] 吴问足, 乔颖, 鲁宗相, 等. 风电功率概率预测方法及展望 [J]. *电力系统自动化*, 2017, 41(18): 167–175.
- WU Wenzu, QIAO Ying, LU Zongxiang, *et al.* Methods and prospects for probability forecasting of wind power[J]. *Automation of Electric Power Systems*, 2017, 41(18): 167–175(in Chinese).
- [8] CUI W, WAN C, SONG Y. Ensemble deep learning-based non-crossing quantile regression for nonparametric probabilistic forecasting of wind power generation[J]. *IEEE Transactions on Power Systems*, 2023, 38(4): 3163–3178.
- [9] CHOI J, EOM H, BAEK S M. A wind power probabilistic model using the reflection method and multi-kernel function kernel density estimation[J]. *Energies*, 2022, 15(24): 1–17.
- [10] 杨锡运, 邢国通, 马雪, 等. 一种核极限学习机分位数回归模型及风电功率区间预测 [J]. *太阳能学报*, 2020, 41(11): 300–306.
- YANG Xiyun, XING Guotong, MA Xue, *et al.* A model of quantile regression with kernel extreme learning machine and wind power interval prediction[J]. *Acta Energetica Sinica*, 2020, 41(11): 300–306(in Chinese).
- [11] HAQUE A U, NEHRIR M H, MANDAL P. A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting[J]. *IEEE Transactions on Power Systems*, 2014, 29(4): 1663–1672.
- [12] JUBAN J, FUGON L, KARINIOTAKIS G. Probabilistic short-term wind power forecasting based on kernel density estimators[C]// *European Wind Power Conference and Exhibition*. Milan, Italy: EWEA, 2007: 7–10.
- [13] 朱满庭. 基于深度学习的风电功率预测研究 [D]. 北京: 华北电力大学(北京), 2021.
- ZHU Manting. Research on wind power prediction based on deep learning[D]. Beijing: North China Electric Power University (Beijing), 2021(in Chinese).
- [14] 杨秀, 陈斌超, 朱兰, 等. 基于相关性分析和长短期记忆网络分位数回归的短期公共楼宇负荷概率密度预测 [J]. *电网技术*, 2019, 43(9): 3061–3071.
- YANG Xiu, CHEN Binchao, ZHU Lan, *et al.* Short-term probability density prediction of public building load based on correlation analysis and quantile regression of long short-term memory networks[J]. *Power System Technology*, 2019, 43(9): 3061–3071(in Chinese).
- [15] 王森, 孙永辉, 侯栋宸, 等. 基于多窗宽核密度估计的风电功率超短期自适应概率预测 [J]. *高电压技术*, 2024, 50(7): 3070–3079.
- WANG Sen, SUN Yonghui, HOU Dongchen, *et al.* Ultra-short-term adaptive probabilistic prediction of wind power based on multi-window wide kernel density estimation[J]. *High Voltage Engineering*, 2024, 50(7): 3070–3079(in Chinese).
- [16] WAHBAH M, MOHANDS B, EL-FOULY T H M, *et al.* Unbiased cross-validation kernel density estimation for wind and PV probabilistic modelling[J]. *Energy Conversion and Management*, 2022, 266(6): 1–10.
- [17] 麻向津. 基于 LSTM 网络和非参数核密度估计的风电功率预测方法研究 [D]. 沈阳: 东北大学, 2020.
- MA Xiangjin. Research on wind power forecasting method based on long short-term memory networks and nonparametric kernel density estimation[D]. Shenyang: Northeastern University, 2020(in Chinese).
- [18] 邓齐林, 邱天宇, 申富饶, 等. 一种自适应在线核密度估计方法 [J]. *软件学报*, 2020, 31(4): 1173–1188.
- DENG Qilin, QIU Tianyu, SHEN Furao, *et al.* Adaptive online kernel density estimation method[D]. *Journal of Software*, 2020, 31(4): 1173–1188(in Chinese).
- [19] 石天. 考虑新能源接入的母线短期净负荷概率预测 [D]. 南京: 东南大学, 2021.
- SHI Tian. Probability forecasting of bus short-term net load considering renewable energy access[D]. Nanjing: Southeast University, 2021(in Chinese).

收稿日期: 2024-03-12

作者简介:

郭恒宽 (1996), 男, 硕士研究生, 研究方向为风电功率预测, E-mail: 1525558537@qq.com;

田建艳 (1966), 女, 博士, 教授, 通信作者, 研究方向为复杂系统建模、综合能源系统管控等, E-mail: tut_tianjy@163.com;

刘坚威 (1994), 男, 硕士研究生, 研究方向为数字孪生系统设计与综合能源管控, E-mail: 2549120912@qq.com;

蒋莹莹 (1998), 女, 硕士研究生, 研究方向为光伏出力预测, E-mail: 630451226@qq.com;

杨立志 (2000), 男, 硕士研究生, 研究方向为数字孪生系统设计与综合能源管控, E-mail: 2042784576@qq.com。