

Copula 分位数回归方法在风电超短期出力预测上的应用

郭军红^{1,2)}, 王小萱^{1,2)}, 汪月新^{1,2)}, 李 薇^{1,2)}, 丁 一³⁾, 贾宏涛^{1,2)}✉

1) 华北电力大学环境科学与工程学院, 北京 102206 2) 华北电力大学资源环境系统优化教育部重点实验室, 北京 102206 3) 国投电力控股股份有限公司, 北京 100034

✉通信作者, E-mail: 2543649929@qq.com

摘 要 风电出力具有较强的随机性和波动性, 相比于传统预测, 分位数预测方法能够提供全面的风电功率概率分布信息, 可实现更可靠的风电出力预报, 对电网系统的安全和稳定运行具有重要意义. 以甘肃某风电场为案例, 将数据按 6 : 2 : 2 划分为训练集、验证集和测试集, 采用基于 Copula 的分位数回归方法 (QCopula) 进行功率区间预测, 并与三个传统的分位数回归方法进行比较. 结果显示, 在不同置信区间下 QCopula 的修正预测区间精度范围在 0.701 ~ 0.773 之间, 预测精度平均值比传统分位数回归 (QR)、随机森林分位数回归 (QRF) 和长短期记忆神经网络分位数回归 (QLSTM) 分别高出 15%、9% 和 13%, 优于其他三种分位数预测方法. 分位数交叉验证中, QCopula 未出现分位数交叉, 每个样本点的功率预测值均随概率值单调递增, 而 QR、QRF、QLSTM 均出现不同程度的分位数交叉现象. 综上所述, QCopula 可以表征更小的区间宽度和更高的区间覆盖率, 且分位数曲线不存在交叉, 可信度较高.

关键词 Copula 函数; 分位数回归; 风电; 超短期; 出力预测

分类号 TK89

Enhancing ultra-short-term wind power forecasting using the Copula quantile regression method

GUO Junhong^{1,2)}, WANG Xiaoxuan^{1,2)}, WANG Yuexin^{1,2)}, LI Wei^{1,2)}, DING Yi³⁾, JIA Hongtao^{1,2)}✉

1) College of Environment Science and Engineering, North China Electric Power University, Beijing 102206, China

2) MOE Key Laboratory of Resources and Environmental Systems Optimization, North China Electric Power University, Beijing 102206, China

3) SDIC Power Holdings Co., Ltd., Beijing 100034, China

✉Corresponding author, E-mail: 2543649929@qq.com

ABSTRACT In recent years, the shift toward renewable energy in China's power industry has been remarkable, with the installed capacity of renewables surpassing that of coal-fired power. Among these, wind power output plays a pivotal role, although it is characterized by its strong randomness and volatility. Traditional prediction methods fall short as they cannot provide comprehensive probability distribution information on the wind power output. To bridge this gap, quantile prediction methods have emerged as superior options for achieving reliable wind power output predictions, which are crucial for the safe and stable operation of power grid systems. To address the inherent unpredictability of wind power, this study introduces a quantile regression method based on Copula (QCopula). The Copula function captures the correlation between the marginal distribution functions of the random variables and their joint distribution function. The process begins with selecting an optimal Copula function using the Akaike Information Criterion (AIC). This function elucidates the relationship between wind power and wind speed, enabling the expression of the conditional probability distribution function of power. By considering different conditional probability values, we obtained wind power prediction results at different quantiles, leading to interval prediction results across different confidence intervals. These results were compared with three

收稿日期: 2023–12–15

基金项目: 国家电网公司总部科技项目 (1400-202456282A-1-1-ZN)

traditional quantile regression methods (Quantile Regression (QR), Quantile regression Random Forests (QRF), and Quantile regression Long Short-Term Memory (QLSTM)) using three elevation metrics: Predictive Interval Coverage Probability (PICP), Predictive Interval Normalized Average Width (PINAW), and Corrected Predictive Interval Accuracy (CPIA). This comparison was aimed at evaluating the interval prediction accuracy of the four quantile regression methods. Finally, the crossover of the quantile curves for each method was analyzed. A case study was conducted at a wind power plant in Gansu Province, utilizing wind speed and power data (measured in MW at 15-minute intervals) from September 2022 to June 2023. With 29088 sample points in total, the data were divided into training, validating and testing sets in an 6 : 2 : 2 ratio. The training set facilitated model development through various quantile regression methods, the validating set was used for model parameterization, whereas the testing was used to evaluate the accuracy of each model. The results showed that the QCopula consistently outperformed the other methods across different confidence intervals, with its modified prediction interval accuracy ranging between 0.701 and 0.773. On average, it exceeded QR, QRF, and QLSTM by 15%, 9%, and 13%, respectively. Notably, the QCopula maintained a consistent increase in the predicted power values for each sample point with probability, without any instances of quantile crossing, a common issue observed in QR, QRF, and QLSTM. In summary, the QCopula offers narrower interval widths and higher interval coverage without the drawback of quantile curve crossing, thereby ensuring higher reliability.

KEY WORDS Copula function; quantile regression; wind power; ultra-short-term; output forecasts

随着“十四五”规划、碳达峰和碳中和政策的推出, 可再生清洁能源发电成为了中国未来发展的重点领域和主要布局点, 风力发电、清洁能源等行业的大力发展步入快车道, 未来行业的发展环境将持续向好。风电不同于传统火力发电, 稳定性较差, 原因是其出力受气象因素(风速)影响巨大, 具有较强的随机性与波动性^[1-2]; 另外由于负荷中心与集中型风电基地的距离问题, 需要进行长距离电力输送, 不稳定的风电大规模直接接入输电网, 会极大威胁电力系统的安全稳定运行^[3-4]。实现精准的风电功率预测, 可以保证系统之间的供用电平衡, 实现电力系统的安全稳定运行。

目前, 风电出力预测模型的构建方法主要分为确定性和不确定性方法。其中确定性方法较多, 很多研究基于人工智能算法来搭建风电出力预测模型。如侯倩等^[5]提出了基于混合深度学习模型的风电功率预测及其一次调频应用方法, 构建了卷积神经网络、双向长短期记忆和注意力机制的混合深度学习模型对风电功率进行预测; Xiong 等^[6]提出了一种基于互补集成经验模式分解、样本熵、随机森林、改进的爬行搜索算法等的混合模型用于风电功率预测。也有学者引入一些方法对预测模型进行改进和优化, 如李进友等^[7]提出一种基于极端梯度提升与广义回归神经网络相耦合(XGBoost-GRNN)的风功率预测算法, 建立考虑分段式风电数据的风电机组功率预测模型。王绍敏等^[8]基于精度加权集成和频域分解方法开展了分布式小风电功率高精度预测研究, 对增强小风机自身稳定性和配电网支撑力具有重要意义。

然而, 传统的确定性方法往往忽略了风速等气象要素的随机性对风电出力的影响。基于分位数的不确定性预测能够提供全面的风电功率概率分布信息, 这类方法逐渐受到了关注^[9]。一些学者开始使用基于分位数回归的概率区间方法进行风电功率预测研究, 常用的分位数回归方法有两类, 一是传统分位数回归, 如王勃等^[10]利用分位数回归方法结合 logistic 函数构建了风电机组功率曲线建模算法, 得到了不同置信区间的功率曲线模型。Wang 等^[11]提出了一种基于共形非对称多分位数生成转换器的预测方法, 实现了高质量的日前风电功率区间预测。二是人工智能算法分位数回归, 如 Liu 等^[12]提出一种基于改进预处理方法和熵加权门控循环单元分位数回归网络(QR-EGRU)相结合的混合短时风电点-区间预测模型。Yu 等^[13]结合混合神经网络和分位数回归的优点, 提出一种新型分位数回归算法, 对区域风电功率进行短期非参数概率预测。Zhu 等^[14]提出了一种采用混沌粒子群优化的非交叉多输出分位数回归深度神经网络, 用于短时风电概率预测。邓宇文^[15]提出一种基于时域神经网络结合分位数回归与动态误差修正的风电功率概率预测方法, 该方法可以在保证高可靠性的同时获得较窄的预测区间。丰胜成等^[16]提出了基于变分模态分解结合改进的松鼠算法优化门控循环单元(GRU)分位数回归方法, 对短期风电功率概率密度进行预测。苏向敬等^[17]基于多任务联合分位数损失理念, 构建了海上风电双重注意力概率预测模型, 验证了该模型提高预测精度的可行性。

但是,超短期尺度上的风电出力对风速等气象条件更加敏感,这些方法仍不能有效解决风速随机性带来的风电出力波动性问题,它们在刻画单个变量的概率分布和多个变量的联合分布之间的关系上仍存在不足.同时不同分位数区间预测方法的区域差异性和适用性问题也亟待分析与研究.

因此,本文采用甘肃某风电站的风速及功率观测数据,使用一种基于 Copula 的分位数回归方法(QCopula)进行功率区间预测,针对区间覆盖率、区间平均宽度、修正预测区间精度、平均分位点交叉数等评价指标对区间预测的效果进行评估,并与传统分位数回归(QR)、随机森林分位数回归(QRF)及长短期记忆神经网络分位数回归(QLSTM)三种对比方法的区间预测结果进行比较.

1 模型与方法

1.1 Copula 函数

Copula 函数的理论最早由 Sklar^[18]提出,后经 Joe^[19]和 Nelsen^[20-21]的完善愈发成熟, Copula 的优点是能够描述随机变量各自的边缘分布函数与变量间联合分布函数的相关关系,充分体现出变量间的相关性^[22]. Copula 方法的核心是 Sklar 定理:

假设有 n 个随机变量 $x_1, x_2, x_3, \dots, x_n$, 其对应的边缘分布函数分别为 $F(x_1), F(x_2), F(x_3), \dots, F(x_n)$, 则必有一个 n 维 Copula 函数 $C[F(x_1), F(x_2), F(x_3), \dots, F(x_n)]$ 满足式(1):

$$G(x_1, x_2, x_3, \dots, x_n) = C[F(x_1), F(x_2), F(x_3), \dots, F(x_n)] \quad (1)$$

式(1)中 $G(x_1, x_2, x_3, \dots, x_n)$ 代表 n 个随机变量的联合分布函数,此 n 维 Copula 函数满足以下三个性质:

(1) 由于边缘分布函数 $F(x_1), F(x_2), F(x_3), \dots, F(x_n)$ 均为 $[0, 1]$ 区间内定义的函数,故联合分布函数的定义域 I^n 为 $[0, 1]^n$;

(2) $C[F(x_1), F(x_2), F(x_3), \dots, F(x_n)]$ 具有零基面同时满足 n 维递增;

(3) $C[F(x_1), F(x_2), F(x_3), \dots, F(x_n)]$ 不会因随机变量 $x_1, x_2, x_3, \dots, x_n$ 的严格单调变换而改变其相关性测度.

1.2 Copula 函数选择原则

本文采用赤池信息准则^[23](Akaike information criterion, AIC)和均方根误差(Root mean square error, RMSE)两个指标对 Copula 函数进行选择.

对于变量对 $\{x_j, x_k\}$, 其采样集合分别为 $X_j =$

$(X_{j,1} \cdots X_{j,N}), X_k = (X_{k,1} \cdots X_{k,N})$, 利用最大似然估计法对各种 Copula 函数参数进行估计:

$$\hat{\theta} = \max_{\theta} \sum_{n=1}^N \ln c(F_x(X_{j,n}), F_y(X_{k,n})|\theta) \quad (2)$$

其中, θ 为 Copula 函数的参数集合, N 为样本数, $c(F_x(X_{j,n}), F_y(X_{k,n})|\theta)$ 为变量对 $\{x_j, x_k\}$ 的 Copula 函数.

AIC 准则的作用是评估统计模型拟合效果.在计算公式中,除了考虑拟合优良性的似然函数之外,还考察了模型复杂性,这样就可以协调模型复杂性和模型拟合优度之间的关系,从而避免出现过拟合现象^[24].文献[25]表明 AIC 准则更加适用于 Copula 函数的选择.

$$AIC_{\min} = \min \left(-2 \sum_{n=1}^N \ln c(F_x(X_{j,n}), F_y(X_{k,n})|\theta) + 2K \right) \quad (3)$$

其中, K 表示 Copula 函数中所包含的参数数目. AIC 数值越小,表示拟合效果越好.

均方根误差计算如式(4)所示:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (4)$$

其中, x_i, y_i 分别为模拟值与实际值. RMSE 数值越小,表示拟合效果越好.

1.3 基于 Copula 的分位数回归模型

基于 Copula 函数所描述的风电功率与风速之间的相关关系,以方程形式表达出风电功率的条件概率分布函数.风电功率 P 关于风速 W 的二维条件概率分布函数 $F(x_P|x_W)$ 表示为

$$F(x_P|x_W) = h_{WP}(F(x_P), F(x_W)) \quad (5)$$

式中,函数 h_{WP} 与 Copula 函数的关系为 $h_{WP} = \partial C(P, W)/\partial W$.

基于式(5),可推导出风电功率条件分位数的回归方程.对风速给定取值: $x_W = R_W$, 利用该回归方程即可计算出风电功率的 α 分位点.

$$\begin{cases} F(\hat{q}_\alpha) = h_{WP}(F(x_P), F(R_W)) \\ \hat{q}_\alpha = F^{-1}(F(\hat{q}_\alpha)) \end{cases} \quad (6)$$

其中, F^{-1} 表示 F 的反函数; $\hat{q}_\alpha$ 表示风电功率的 α 分位点估计值.

在以上分位数回归方程中,条件概率 α 为输入变量,因此一旦模型训练结束, Copula 参数确定后,只需改变条件概率的取值,即可获得任意分位点的风电功率预测.因此回归方程的 Copula 参数一旦通过模型参数训练确定后,通过条件概率 α 的不同取值便可实现任意分位点的风电功率预测.

1.4 模型参数

本文基于 QCopula 方法对风电超短期出力进行预测, 为了验证预测效果, 引入了传统分位数回归(QR)、随机森林分位数回归(QRF)、长短期记忆神经网络分位数回归(QLSTM)三种模型进行对比. 四种模型的分位数区间参数为分位数间隔(39)和区间范围(0.025 ~ 0.975), 其他参数如表 1 所示.

表 1 模型参数表

Table 1 Model Parameter Table

Model	Parameters	Value
QCopula	Copula function	Frank
QR	—	—
QRF	n_estimators	300
	max_depth	6
	max_features	sqrt
QLSTM	Batch_size	32
	Epochs	600
	Optimizer	Adam

1.5 模型评价方法

1.5.1 模型验证评价指标

采用均方根误差(RMSE)、平均绝对误差(MAE)、平均相对误差(MRE)对模型预测结果进行评价.

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - y_i| \quad (7)$$

$$MRE = \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - y_i}{y_i} \right| \quad (8)$$

其中, x_i , y_i 分别为模拟值与实际值.

1.5.2 区间覆盖率

区间覆盖率(Predictive interval coverage probability, PICP)是评估风电功率预测效果最重要的指标之一. 其定义是指风电功率的观测值位于预测区间的上、下界内的概率. 它反映了预测模型的准确程度. 从定义来看, 预测效果最好的情况下, PICP=100%, 在实际预测中, PICP 的值越大, 预测效果越好. 通常我们希望观测值位于预测区间内的概率 p 的值能高于给定置信水平 β , 即

$$p = P(V_i \in [\underline{V}_i, \overline{V}_i]) \geq \beta \quad (9)$$

其中, p 表示预测区间包络风电功率观测值的概率; P 为概率函数; V_i 表示风电功率观测值; $\underline{V}_i$ 表示预测区间的下界; $\overline{V}_i$ 表示预测区间的上界.

依据伯努力大数定律, 本文采用风电功率观测值位于预测区间的频率近似来估计区间覆盖率 PICP, 即:

$$PICP = \frac{1}{N} \sum_{i=1}^N A_i \times 100\%, P(V_i \in [\underline{V}_i, \overline{V}_i]) \quad (10)$$

其中, N 表示待预测风电功率时刻的总数; A_i 为布尔量, 其定义为式(11)

$$A_i = \begin{cases} 1, & V_i \in [\underline{V}_i, \overline{V}_i] \\ 0, & V_i \notin [\underline{V}_i, \overline{V}_i] \end{cases} \quad (11)$$

1.5.3 区间平均宽度

若一味要求 PICP 值大于给定的置信水平, 将可能导致预测区间宽度过大, 使区间预测显得过于保守, 丧失了区间预测的实用性. 因此, 预测区间的平均宽度(Predictive interval normalized average width, PINAW)也是一个评估预测区间效果的重要指标, 如式(12)所示:

$$PINAW = \frac{1}{NR} \sum_{i=1}^N (\overline{V}_i - \underline{V}_i) \quad (12)$$

其中, R 表示观测值的极差.

区间预测平均宽度结合区间覆盖率, 可以对区间预测效果作出更加客观合理的评价. 理论上预测效果最好的情况下, PINAW=0, 但此时区间预测退化为点预测, 所以本文区间预测希望 PINAW 值在 PINAW>0 的前提下尽可能小.

1.5.4 修正预测区间精度

修正预测区间精度(CPIA)^[26]综合考虑了区间覆盖率和区间平均宽度, 是一种更加全面合理的衡量区间预测精度的新指标, 定义如下:

$$CPIA = \frac{1}{N} (1 - PINAW) \sum_{i=1}^N C_i \quad (13)$$

其中, C_i 的定义为:

$$C_i = \begin{cases} 1 - \frac{y_i - \overline{V}_i}{d_i}, & y_i \in \left[\overline{V}_i, \overline{V}_i + \frac{d_i}{2} \right] \\ 1, & y_i \in [\underline{V}_i, \overline{V}_i] \\ 1 - \frac{\underline{V}_i - y_i}{d_i}, & y_i \in \left[\underline{V}_i - \frac{d_i}{2}, \underline{V}_i \right] \\ 0, & y_i \notin \left[\underline{V}_i - \frac{d_i}{2}, \overline{V}_i + \frac{d_i}{2} \right] \end{cases} \quad (14)$$

其中, $d_i = \overline{V}_i - \underline{V}_i$. CPIA 的取值在 0 ~ 1 之间, 数值越大表示精确度越高.

1.5.5 平均分位点交叉数

平均分位点交叉数($\bar{N}_{QC}$)可以表征分位数预测方法在区间预测上的可信度, 定义如下:

$$\overline{N}_{QC} = \frac{1}{N} \sum_{i=1}^M \sum_{j=i+1}^M c_{ij} \tag{15}$$

其中, $\overline{N}_{QC}$ 表示平均分位点交叉数; N 表示待预测样本总数; M 表示研究设置的分位点个数, 本文在 2.5% ~ 97.5% 之间设置 39 个分位点, 相邻分位点间隔为 2.5%; c_{ij} 表示第 i 、 j 个两个分位点的风电功率曲线之间交叉点的个数.

2 案例分析

2.1 数据样本

本文数据来自甘肃某风电场, 经纬度为 37.0°N, 104.0°E, 总装机容量为 26800 MW, 使用时间范围为 2022 年 9 月至 2023 年 6 月的风速和风电功率

数据(单位为 MW, 间隔为 15 min), 共 29088 个样本点, 采用 6 : 2 : 2 的比例划分训练集、验证集和测试集, 2022 年 9 月至 2023 年 2 月数据用于训练模型, 2023 年 3 月至 4 月数据用于验证模型, 2023 年 5 月至 6 月数据用于测试模型的区间预测效果.

2.2 Copula 函数选择

本文采用高斯混合分布(GMM)作为风速与功率的边际分布函数, 模拟结果显示, 风速和功率的 GMM 分布 P -value 分别为 0.942 和 0.751, 均大于 0.05, GMM 分布对于风速和功率的拟合效果较好.

在 Copula 函数选择上, 采用 Clayton、Gumbel、Frank、Joe 四种函数对风速、功率之间的联合分布进行拟合, 结果如图 1 所示.

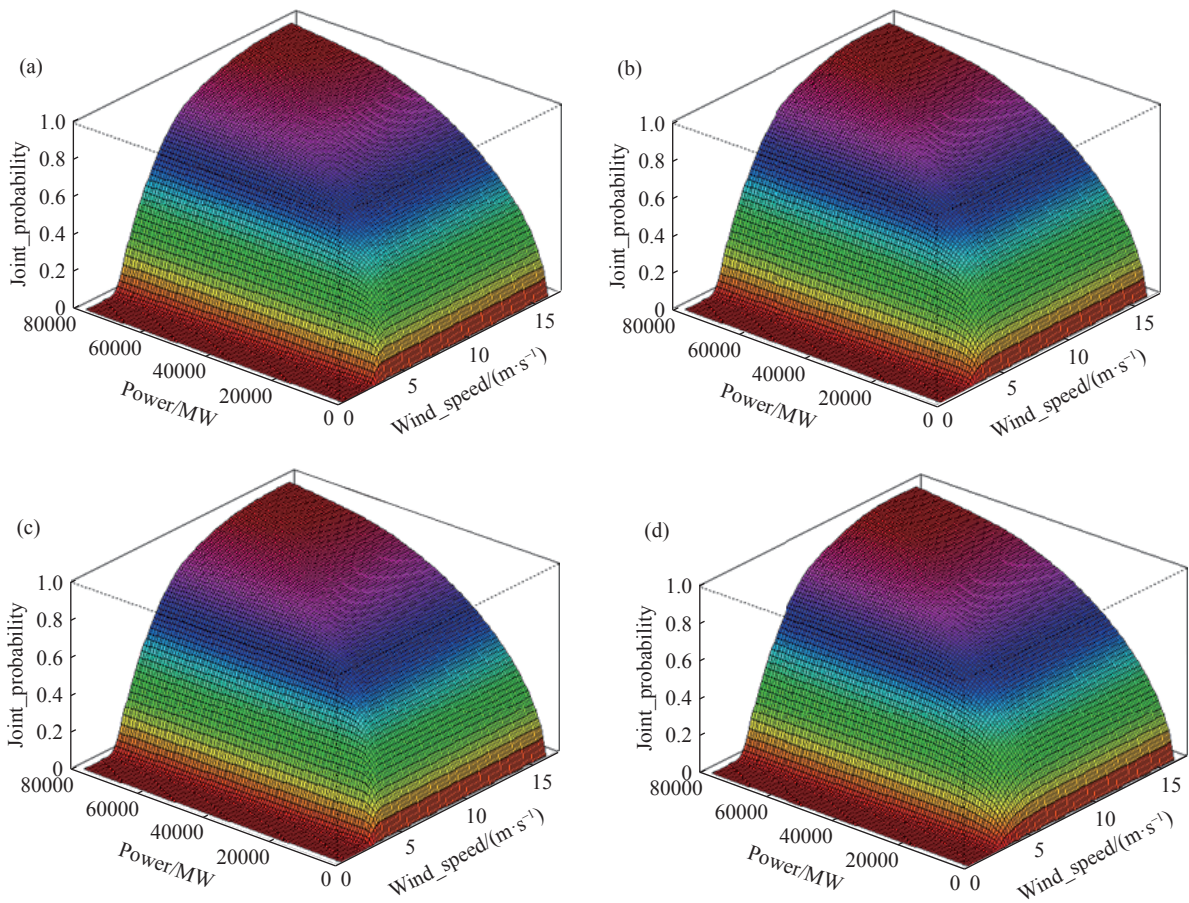


图 1 四种 Copula 函数对风速、功率联合分布拟合图. (a) Clayton; (b) Gumbel; (c) Frank; (d) Joe

Fig.1 Four Copula functions fitting the joint distribution of wind speed and power: (a) Clayton; (b) Gumbel; (c) Frank; (d) Joe

以 AIC 准则作为定量评价指标, 确定最优二元 Copula 函数类别. 表 2 结果显示, Frank Copula 函数的 AIC 最小(-146435), 其对应的 RMSE 也最小(0.010), 通过对比发现, Frank 函数的拟合效果最好, 因此本文选择 Frank Copula 函数拟合风速与功率之间的关系.

2.3 结果分析

2.3.1 模型效果评价

四种模型在验证期和测试期的评估结果如表 3 所示. 结果显示, QCopula 在验证期的 RMSE、MAE 和 MRE 分别为 7.381、301 和 0.037, 在测试期为 1.602、703 和 0.048, 这些误差都比其他模型要低,

表 2 Copula 函数的评价指标

Table 2 Evaluation Metrics of the Copula Function

Copula function	RMSE	AIC
Clayton	0.030	-110700
Gumbel	0.035	-105426
Frank	0.010	-146435
Joe	0.062	-87868

表 3 模型效果评价指标

Table 3 Evaluation Metrics of Model Performance

Model	RMSE	MAE	MRE
QCopula	7.381 (1.602)	301 (703)	0.037 (0.048)
QR	13.309 (4.288)	796 (772)	0.25 (0.16)
QRF	10.003 (3.217)	706 (809)	0.055 (0.061)
QLSTM	8.413 (15.455)	608 (724)	0.093 (0.062)

Note: Test set results are in parentheses, validation period results are outside parentheses.

说明 QCopula 方法在模拟风电出力上具有一定的优越性。

2.3.2 区间预测分析

使用 QCopula 方法和其他三种分位数方法的区间预测结果如图 2 所示, 四种分位数方法预测的区间变化趋势与实际功率曲线 (obs) 一致, QCopula 方法在功率波动较大的时段 (如样本号 2000 ~ 4000), 预测区间也能很好地覆盖实际功率曲线, 且随着置信水平增加, 预测区间变宽。尤其在样本号 0 ~ 1000 之间, QCopula 方法的模拟结果

相比 QRF 分位数方法更好, 与实际功率更为接近。QCopula 也能更好地模拟出实际功率的峰谷功率 (如样本号 2500 ~ 3500)。

为了更准确地比较 QCopula 方法与其他分位数方法在区间预测上的效果, 本文采用 PICP、PINAW、CPIA 三个评价指标, 对区间预测结果进行评价, 其中置信度选取 95%、90%、85%、80%。就 PINAW 而言, QR 和 QLSTM 方法较大, 分别为 0.330、0.310, QRF 方法最小为 0.182, QCopula 结果适中; 就 PICP 而言, QCopula 方法与 QR、QLSTM 的结果基本一致, 均大于 0.8, QRF 值最低 (0.653)。

图 3 展示了四种分位数回归方法的区间评价指标随置信度的变化。结果显示, 随着置信度的提高, 四种方法的 PINAW 均呈现增大趋势, QCopula 及 QRF 的 PINAW 较低, 增长趋势相对平缓 (图 3(a)); 四种方法的 PICP 也随着置信度增加而升高, QCopula、QR 和 QLSTM 三种方法的 PICP 较高, 且数值相近, 而 QRF 的 PICP 远低于其他三种方法 (图 3(b)); 在四个置信区间下, QCopula 方法 CPIA 分别为 0.701、0.747、0.767 和 0.773, 各置信区间的平均值比 QR、QRF 和 QLSTM 分别高出 15%、9% 和 13% (图 3(c))。综上结果, QCopula 的区间预测效果明显优于三种对比分位数预测方法。

2.3.3 分位数交叉检验

四种分位数回归方法交叉检验结果显示 (图 3(d)), 相对于其他分位数方法, QCopula 方法未出现分位点交叉, 说明 QCopula 处理分位数回

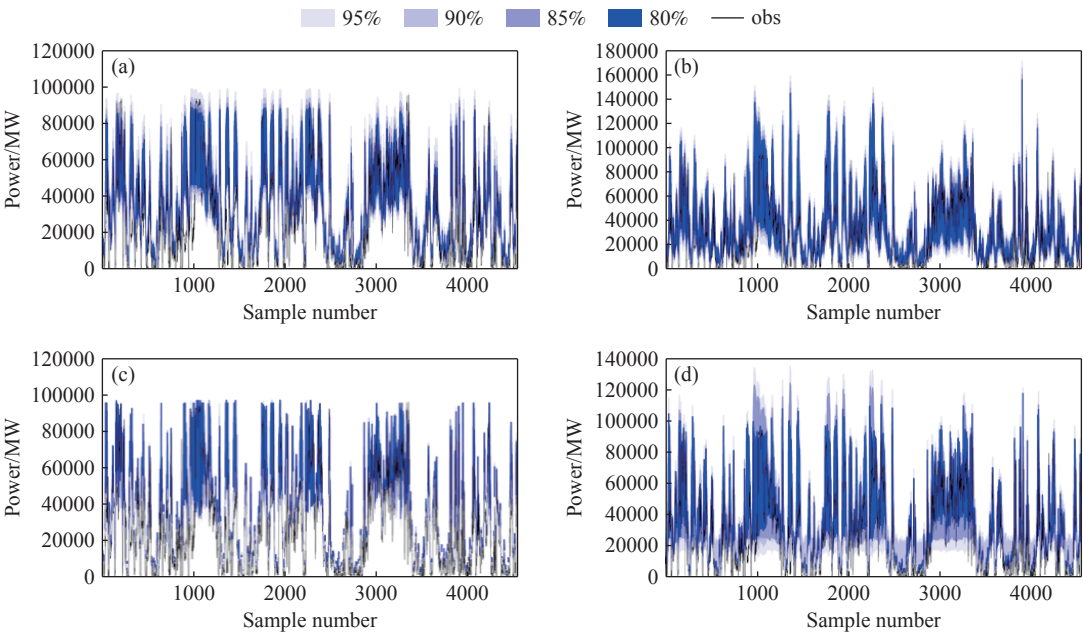


图 2 四种分位数回归模型预测结果. (a) QCopula; (b) QR; (c) QRF; (d) QLSTM

Fig.2 Prediction results of four quantile regression models: (a) QCopula; (b) QR; (c) QRF; (d) QLSTM

归问题的优越性. 图 4 显示了四种分位数回归方法在不同概率条件下的部分分位点预测曲线. 图中可以看出, 采用 QR、QRF 和 QLSTM 方法时, 同

一时刻各分位点预测曲线之间存在严重的交叉, 尤其是 QRF 方法, 平均分位点交叉数达到 134, 为四种方法最高, 表明此三种方法概率预测结果中

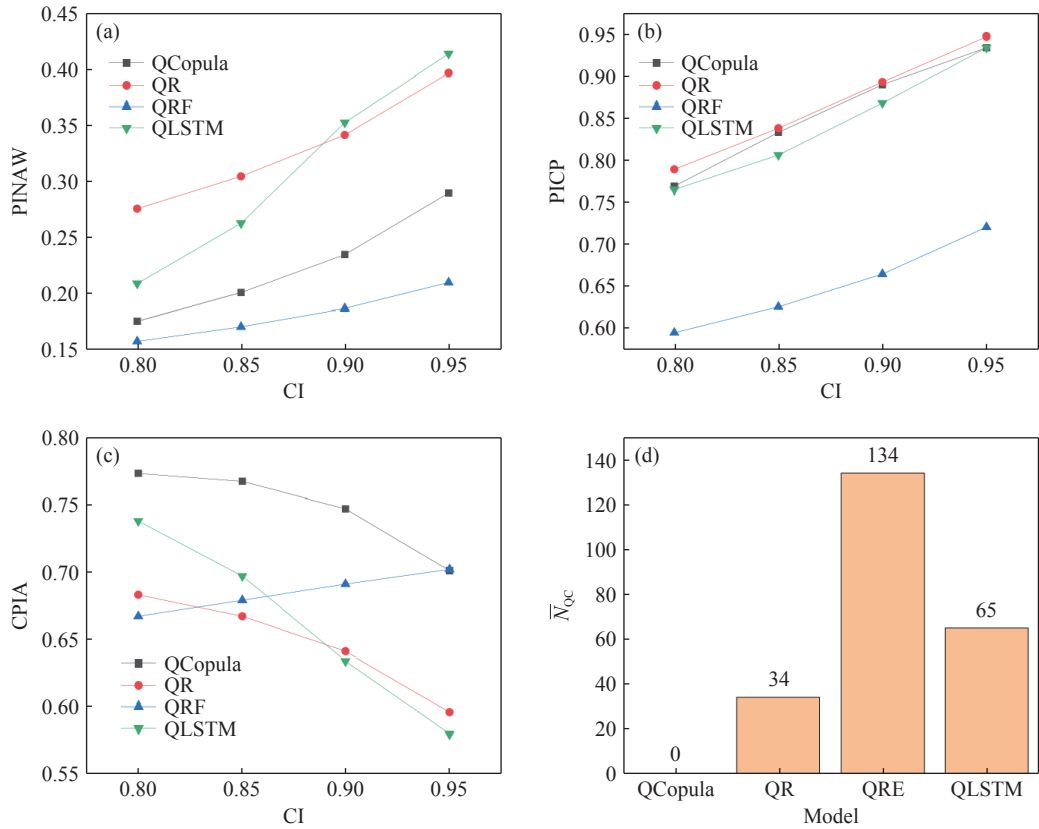


图 3 不同置信度区间评价指标对比图. (a) PINAW; (b) PICP; (c) CPIA; (d) $\bar{N}_{QC}$

Fig.3 Comparison of evaluation metrics across different confidence intervals: (a) PINAW; (b) PICP; (c) CPIA; (d) $\bar{N}_{QC}$

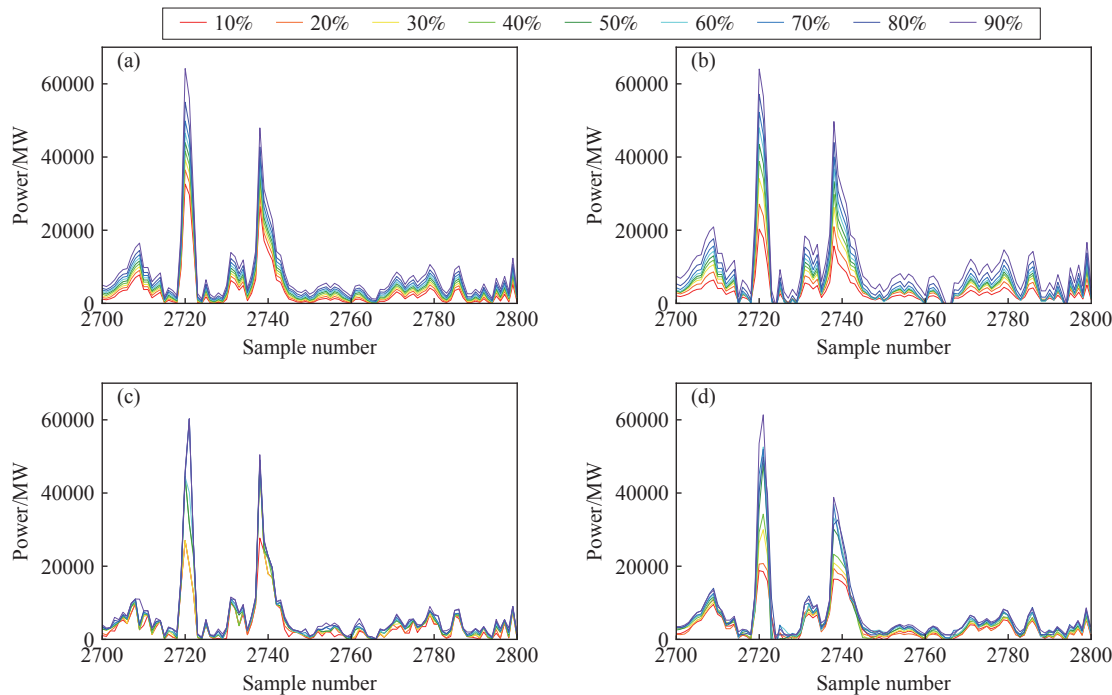


图 4 四种分位数回归模型分位数交叉图. (a) QCopula; (b) QR; (c) QRF; (d) QLSTM

Fig.4 Cross quantile plot of four quantile regression models: (a) QCopula; (b) QR; (c) QRF; (d) QLSTM

存在不合理的样本点, 该点功率预测值不随概率值单调递增. 原因可能是在 QR、QRF、QLSTM 三种方法中, 不同分位点的预测结果是通过不断训练多个分位点回归模型得到的, 但在模型训练过程中, 不同分位点的回归模型之间并没有条件对单调性进行约束, 由此造成分位点交叉; 而 QCopula 方法是基于 Copula 函数构建而成的分位数回归模型, 并不依赖多次训练模型得到分位点预测结果, 能够防止分位点交叉的发生.

3 结论

由于风电具有较强的随机性和波动性, 分位数概率预测相比于传统方法可以为电站出力预测提供更加科学的决策支持. 本文以甘肃某风电站为案例, 提出一种基于 Copula 的分位数回归模型, 并与 QR、QRF 及 QLSTM 三种分位数回归方法进行对比, 结论如下:

(1) 在四个置信区间下, QCopula 方法的修正区间预测精度范围在 0.701 ~ 0.773 之间, 各置信区间的平均值比 QR、QRF 和 QLSTM 分别高出 15%、9% 和 13%, 说明 QCopula 方法在该风电场的超短期功率预测中效果更好.

(2) 在分位数交叉检验时, 采用 QCopula 方法时, 各分位点功率预测曲线之间不存在交叉, 而采用 QR、QRF、QLSTM 方法时, 同一时刻各分位点预测曲线之间存在交叉现象, 不同分位点模型的独立训练和缺乏单调性约束可能是造成这种现象的主要原因.

(3) 相比于确定性预测, 分位数回归等不确定性预测方法可以更好地解决风电出力预测的随机性问题. 未来可以在获得更大数据样本的基础上, 优选更好的边缘分布函数和 Copula 函数进行分位数回归, 进一步提高区间预测的可靠性.

参 考 文 献

- [1] Liu Z H, Wei Z N, Gao S Y, et al. Adaptive-linearized probabilistic power flow calculation for power grid integrated with high proportion wind power in source-load interactive environment. *Power Syst Technol*, 2019, 43(11): 3926
(柳志航, 卫志农, 高昇宇, 等. 源-荷互动环境下含高比例风电并网的自适应线性化概率潮流计算. *电网技术*, 2019, 43(11): 3926)
- [2] Cao H Q. *Day-Ahead Economic Dispatch of Power System with Wind Power Considering Demand Response* [Dissertation]. Wuhan: Wuhan University, 2018
(曹慧秋. 计及需求响应的含风电电力系统日前经济调度[学位论文]. 北京: 华北电力大学, 2021)
- [3] Liu Z Y, Zhang Q P, Dong C, et al. Efficient and security transmission of wind, photovoltaic and thermal power of large-scale energy resource bases through UHVDC projects. *Proceedings of the CSEE*, 2014, 34(16): 2513
(刘振亚, 张启平, 董存, 等. 通过特高压直流实现大型能源基地风、光、火电力大规模高效率安全外送研究. *中国电机工程学报*, 2014, 34(16): 2513)
- [4] Liu D W, Huang Y H, Wang W S, et al. Analysis on provincial system available capability of accommodating wind power considering peak load dispatch and transmission constraints. *Autom Electr Power Syst*, 2011, 35(22): 77
(刘德伟, 黄越辉, 王伟胜, 等. 考虑调峰和电网输送约束的省级系统风电消纳能力分析. *电力系统自动化*, 2011, 35(22): 77)
- [5] Hou Q, Hao X G, Jin F, et al. Wind power prediction and primary frequency regulation application based on hybrid deep learning. *J Eng Thermal Energy Power*, 2023, 38(10): 167
(侯倩, 郝晓光, 金飞, 等. 基于混合深度学习的风电功率预测及一次调频应用. *热能动力工程*, 2023, 38(10): 167)
- [6] Xiong J L, Peng T, Tao Z, et al. A dual-scale deep learning model based on ELM-BiLSTM and improved reptile search algorithm for wind power prediction. *Energy*, 2023, 266: 126419
- [7] Li J Y, Li Y, Huang L Q, et al. Piecewise wind power forecast based on the XGBoost-GRNN algorithm. *Comput Integr Manuf Syst*, <http://kns.cnki.net/kcms/detail/11.5946.tp.20230927.1028.004.html>
(李进友, 李媛, 黄露秋, 等. 基于 XGBoost-GRNN 算法的分段式风功率预测. *计算机集成制造系统*, <http://kns.cnki.net/kcms/detail/11.5946.tp.20230927.1028.004.html>)
- [8] Wang S M, Wang S X, Zhao Q Y, et al. Distributed wind power forecasting method based on frequency domain decomposition and precision-weighted ensemble. *Electr Power Constr*, 2023, 44(5): 84
(王绍敏, 王守相, 赵倩宇, 等. 基于频域分解和精度加权集成的分布式风电功率预测方法. *电力建设*, 2023, 44(5): 84)
- [9] Zhang H. *Research on Time-Space Uncertainty Forecasting Method of Wind Power* [Dissertation]. Beijing: North China Electric Power University, 2021
(张浩. 风电功率时空不确定性预测方法研究[学位论文]. 北京: 华北电力大学, 2021)
- [10] Wang B, Sun Y, Li Z Y, et al. Wind turbine power curve modelling with Logistic functions based on quantile regression. *Electr Meas Instrum*, <https://kns.cnki.net/kcms/detail/23.1202.th.20230412.1608.002.html>
(王勃, 孙勇, 李振元, 等. 基于 Logistic 函数与分位数回归的风电机组功率曲线建模方法. *电测与仪表*, <https://kns.cnki.net/kcms/detail/23.1202.TH.20230412.1608.002.html>)
- [11] Wang W, Feng B, Huang G, et al. Conformal asymmetric multi-quantile generative transformer for day-ahead wind power interval prediction. *Appl Energy*, 2023, 333: 120634
- [12] Liu T H, Qi S, Qiao X, et al. A hybrid short-term wind power

- point-interval prediction model based on combination of improved preprocessing methods and entropy weighted GRU quantile regression network. *Energy*, 2024, 288: 129904
- [13] Yu Y X, Han X, Yang M, et al. Probabilistic prediction of regional wind power based on spatiotemporal quantile regression // 2019 *IEEE Industry Applications Society Annual Meeting*. Baltimore, 2019: 1
- [14] Zhu J H, He Y, Yang X, et al. Ultra-short-term wind power probabilistic forecasting based on an evolutionary non-crossing multi-output quantile regression deep neural network. *Energy Convers Manag*, 2024, 301: 118062
- [15] Deng Y W. Probabilistic power prediction of short-term wind generation based on temporal convolutional network and quantile regression. *Electr Eng*, 2023(21): 49
(邓宇文. 基于时域卷积网络和分位数回归的短期风电功率概率预测. 电工技术, 2023(21): 49)
- [16] Feng S C, Guo J C, Fu H, et al. Wind power probability density prediction based on ISSA and GRU quantile regression. *Adv Technol Electr Eng Energy*, 2023, 42(10): 55
(丰胜成, 郭继成, 付华, 等. 基于 ISSA 与 GRU 分位数回归的风电功率概率密度预测. 电工电能新技术, 2023, 42(10): 55)
- [17] Su X J, Yu H B, Fu Y, et al. Probabilistic forecasting of offshore wind power based on dual-stage attentional LSTM and joint quantile loss function. *Electr Power*, 2023, 56(11): 10
(苏向敬, 宇海波, 符杨, 等. 基于 DALSTM 和联合分位数损失的海上风电功率概率预测. 中国电力, 2023, 56(11): 10)
- [18] Sklar M. Distribution functions in N dimensions and their margins // *Annals of the ISUP*. Paris, 1959: 229 (Sklar M. Fonctions de répartition à N dimensions et leurs marges // *Annales de l'ISUP*. Paris, 1959: 229)
- [19] Joe H. *Multivariate Models and Multivariate Dependence Concepts*. Boca Raton: CRC press, 1997
- [20] Nelsen R B. Properties and applications of copulas: A brief survey // *Proceedings of the First Brazilian Conference on Statistical Modeling in Insurance and Finance*. Sao Paulo, 2003: 10
- [21] Nelsen R B. *An Introduction to Copulas*. New York: Springer, 2006
- [22] Li Z X. *Research on Wind Power Interval Prediction Based on Copula Function* [Dissertation]. Xi'an: Xi'an University of Technology, 2021
(李智轩. 基于 Copula 函数的风电功率区间预测研究[学位论文]. 西安: 西安理工大学, 2021)
- [23] Xu B, Xu Q S, Huang Y, et al. Day-ahead probabilistic forecasting of photovoltaic power based on vine copula quantile regression. *Power Syst Technol*, 2021, 45(11): 4426
(许彪, 徐青山, 黄煜, 等. 基于藤 copula 分位数回归的光伏功率日前概率预测. 电网技术, 2021, 45(11): 4426)
- [24] Xu Q S, Huang Y, Liu J K, et al. Probabilistic load flow method using Monte Carlo simulation based on Gaussian mixture model and marginal transformation. *Autom Electr Power Syst*, 2016, 40(16): 23
(徐青山, 黄煜, 刘建坤, 等. 采用混合高斯模型及边缘变换技术的蒙特卡洛随机潮流方法. 电力系统自动化, 2016, 40(16): 23)
- [25] Brechmann E C, Czado C, Aas K. Truncated regular vines in high dimensions with application to financial data. *Can J Statistics*, 2012, 40(1): 68
- [26] Leng T. *Hierarchical Power Load Forecasting Based on Quantile Regression Neural Network* [Dissertation]. Hunan: Hunan Normal University, 2021
(冷婷. 基于分位数回归神经网络的分层电力负荷预测[学位论文]. 湖南: 湖南师范大学, 2021)