

文章编号:1006-2467(2023)07-0845-14

DOI: 10.16183/j.cnki.jsjtu.2021.377

# 基于在线 LASSO VAR 和 EGARCH 模型的风场功率集成概率预测

王 鹏, 李艳婷, 张 宇

(上海交通大学 机械与动力工程学院, 上海 200240)

**摘要:** 由于风速波动性大, 风力发电往往呈现一定的不确定性。传统风能预测模型以均值为 0、方差固定的正态分布度量不确定性, 但方差可能随时间变化, 即具有异方差性。为提升预测精度, 基于在线最小绝对收缩和选择算子的向量自回归 (LASSO VAR) 和指数自回归条件异方差 (EGARCH) 模型, 提出一种考虑异方差性的风场级功率集成概率预测模型。首先使用在线 LASSO VAR 模型预测风力机的有功功率, 再利用自回归条件异方差检验验证残差的异方差性, 并利用信息冲击曲线和动态显著线评估正负残差对未来条件方差的不对称影响。然后针对异方差性和不对称性, 使用 EGARCH 模型对单风力机有功功率的残差进行预测, 得到有功功率的条件方差。最后, 考虑各风力机有功功率的相关性, 将风场中各风力机的有功功率求和, 得到整个风场总有功功率的概率预测结果。将该方法应用于中国华东某地风场, 验证了该模型能有效提高预测精度。

**关键词:** 在线 LASSO VAR; 异方差; 指数条件异方差模型; 概率预测

中图分类号: TM 614 文献标志码: A

## Probabilistic Forecasting of Wind Power Generation Using Online LASSO VAR and EGARCH Model

WANG Peng, LI Yanting, ZHANG Yu

(School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

**Abstract:** Wind power generation has uncertainty due to the high fluctuation of wind speed. In traditional wind power prediction models, the uncertainty is measured by normal distribution with zero mean and constant variance. However, the variance may vary with time, which means the variance has heteroscedasticity. To improve the prediction accuracy, this paper proposes a new integrated probabilistic wind power prediction model for wind farm considering heteroscedasticity based on online least absolute shrinkage and selection operator and vector autoregression (LASSO VAR) and the exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model. First, online LASSO VAR is used to forecast power output. Then, heteroscedasticity of residuals is validated by autoregressive conditional heteroskedasticity test. Considering heteroscedasticity, the news impact curve and dynamic significance line verify that positive and negative residuals affect future volatility asymmetrically. Thus, the EGARCH model is used to forecast the residuals to obtain the conditional variance of point prediction results. Finally, the probabilistic result of total power output is obtained by summing the power output of turbines.

收稿日期:2021-09-26 修回日期:2021-12-20 录用日期:2021-12-31

基金项目:国家自然科学基金面上项目(72072114)

作者简介:王 鹏(1998-),硕士生,研究方向为数据驱动的能源预测等。

通信作者:李艳婷,副教授,博士生导师; E-mail:ytli@sjtu.edu.cn.

in the wind farm considering the correlation of the active wind power of wind turbines. This method is applied to forecast the power output of a wind farm in East China and is proved effective in improving the prediction accuracy.

**Key words:** online least absolute shrinkage and selection operator and vector autoregression (LASSO VAR); heteroscedasticity; exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model; probabilistic forecasting

由于传统能源的不可再生性,清洁绿色能源受到越来越广泛的关注。风能作为一种清洁的可再生能源,是近年来全球增长最快的能源之一。风能是重要的可再生能源,风力发电(简称风电)累计装机容量持续增长,从 2010 年的 198 GW 增长到 2021 年的 837 GW,预计到 2030 年将达到 3 200 GW<sup>[1]</sup>。然而,风速不确定性高、波动性大,这使得风力机有功功率往往也呈现一定的不确定性,因此有必要进行有功功率的概率预测。但传统的概率预测模型未考虑方差可能随时间变化,即具有异方差性的问题,影响了功率预测的准确性。因此,亟需探索新方法以准确预测整个风电场的输出功率。

风电预测模型一般可分为确定性预测和概率预测两类。现有风电预测模型大多是仅提供特定预测值的确定性模型。确定性模型可分为基于数值天气预测的模型<sup>[2-3]</sup>、传统统计模型<sup>[4-6]</sup>、基于机器学习的模型<sup>[7-8]</sup>和这些方法的集成模型<sup>[9-10]</sup>。不同确定性预测模型的详细概述可见文献[11-12]。考虑时空相关的回归模型是同时考虑时间和空间相关性的确定性预测模型,基于最小绝对收缩和选择算子(LASSO)的向量自回归(Vector Autoregression, VAR)模型计算效率较高,是一种常用的时空模型<sup>[13-14]</sup>。确定性预测模型的一个明显缺点是不能量化预测的不确定性。相反,概率预测模型可以给出预测值的置信区间或预测值的概率分布函数来评估预测的不确定性。

概率预测模型主要分为参数、非参数和集成模型。参数模型需要预测指定数据和参数的分布。高斯分布<sup>[15]</sup>和贝塔分布<sup>[16]</sup>是描述风电分布的两种合理选择。非参数模型并不需要预先指定数据和参数的分布。分位数回归(Quantile Regression, QR)是广泛应用的非参数风电概率预测模型。文献[17-18]中利用局部 QR 得到风能的概率预测。核密度估计(Kernel Density Estimation, KDE)模型由于灵活的数据适应性,也常用作非参数概率预测模型。KDE 模型可以提供风电预测概率密度函数的所有信息,然而,由于传统的 KDE 为所有历史数据分配相等的权重,所以会产生具有高同质性的密度函数。

因此,有研究提出采用与历史数据相似性相关的加权核密度估计<sup>[19]</sup>。

近年来,随着基于数值天气预测模型的发展,基于集成的概率预测模型也被应用于一些研究中。其中,贝叶斯模型平均(Bayes Model Averaging, BMA)是最重要的一个。风力机有功功率预测的概率密度函数由几个条件概率密度函数的加权和得到,更高的权重被分配给预测更准确的成员。BMA 已成功应用于风速<sup>[20]</sup>和风电<sup>[21]</sup>的概率预测。集成方法可以基于集成成员的预测结果进行偏差校正,由于条件分布设置灵活,所以具有较高的适应性。

然而,以上概率预测模型未充分考虑异方差性,自回归条件异方差(Autoregressive Conditional Heteroskedasticity, ARCH)族模型为预测方差提供了更好的选择。文献[22]中结合自回归移动平均(Autoregressive Moving Average, ARMA)模型和广义自回归条件异方差(Generalized Autoregressive Conditional Heteroskedasticity, GARCH)模型建立了 ARMA-GARCH 模型预测风功率数据,并针对不同情况的预测误差给予对应的补偿方案,提升了风电功率的预测精度。文献[23]中首先建立风电波动性与风电预测误差的动态相依性模型,并结合差分整合 ARMA 模型和 GARCH 模型建立时变分布边缘模型,最终给出不同波动水平下的风电条件预测误差分布。

综上所述,针对风电不确定性高和现有概率预测模型未充分考虑异方差性的问题,首先使用在线 LASSO VAR 模型对风电场中多个风力机的有功功率时间序列进行建模,以描述风力机有功功率间的时空相关性,减少预测所需的运行时间并提升预测准确度。同时,考虑风能点预测模型可能存在的异方差问题,使用 ARCH 检验验证点预测的异方差性,并使用信息冲击曲线和动态显著线验证正负偏差对未来残差波动的不同影响。之后使用指数自回归条件异方差(EGARCH)模型对点预测模型的不确定性建模,追踪点预测模型不确定性的趋势。最后基于风力机功率的空间相关性,对风场中风力机的

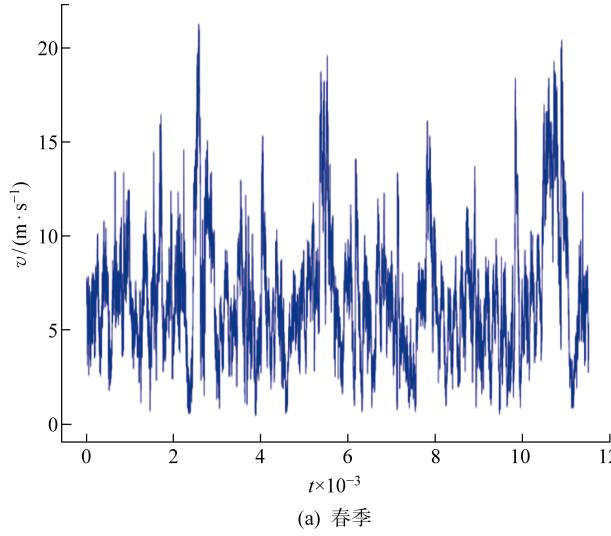
有功功率进行集成,计算风场总有功功率的均值和方差,得到风场总功率的概率预测.

## 1 数据集介绍

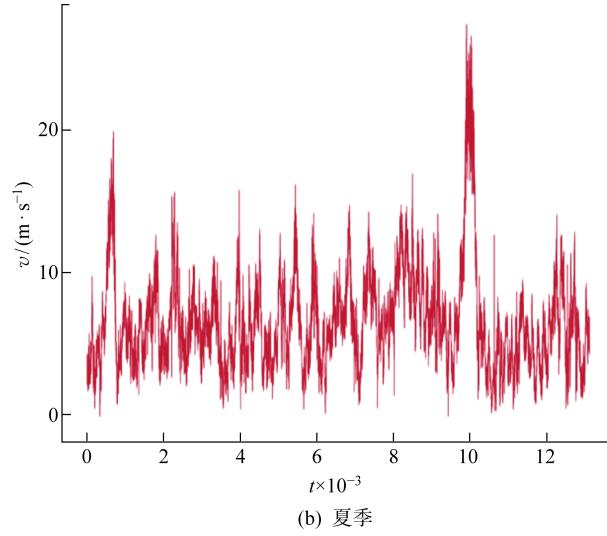
使用中国华东某地海上风场的数据.该风场共有 24 台风力机,风力机额定功率为 3.6 MW,总装机容量为  $20.42 \times 10^4$  kW.风场所处位置四季分明,年平均气温为 16 ℃,年平均降水量为 1 100 mm.风力机的传感器记录风力机所在位置的实时数据,包括风速、风向、有功功率.为了提升后续时序模型的效果,需要使时序数据更加稳定,因此对每小时的数据取平均值以平滑数据.

图 1 为风场一年四季的风速时序图对比.图中: $t$  为相应季节中的时间步索引,每两个时间步的间隔为 10 min;  $v$  为风速.

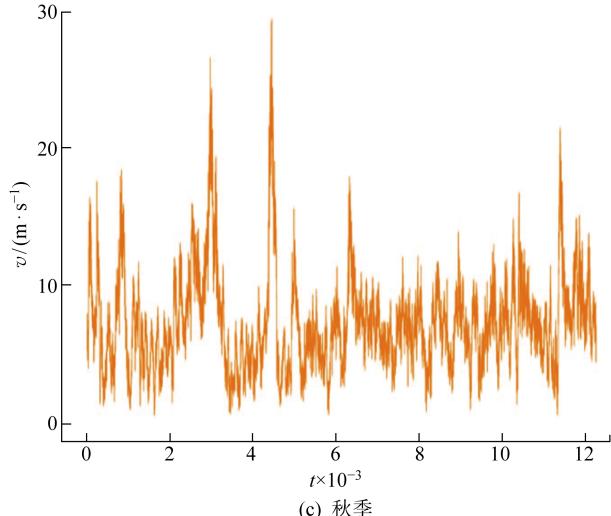
图 2 为风向玫瑰图对比.风速和风向在不同季



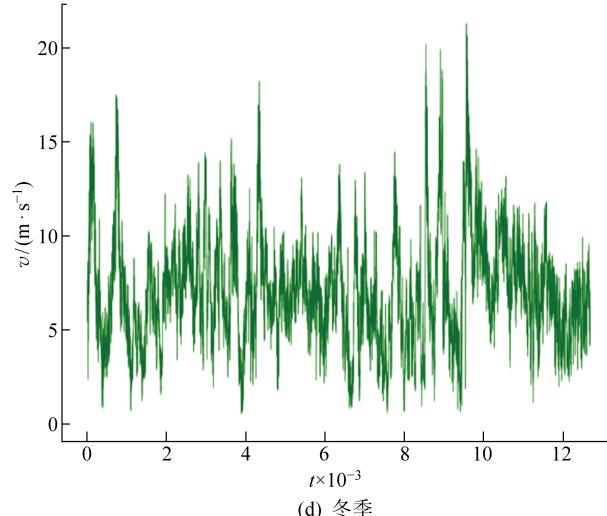
(a) 春季



(b) 夏季



(c) 秋季



(d) 冬季

图 1 不同季节风速时序图

Fig. 1 Time series of wind speed in different seasons

节呈现不同趋势和模式.风速在不同季节均具有较大的波动性,与其他季节相比,冬季的风速波动最小.不同季节的盛行风不同,春季盛行西南风,夏季盛行南风,而秋季和冬季主要盛行东风.

图 3 展示了不同时期下(全年、1月、2月、3月)前 13 台风力机( $B_1 \sim B_{13}$ )记录的有功功率相关系数( $\delta$ ).可以看出,大多数相关系数都在 0.8 以上,这意味着风力机的有功功率具有较强的空间相关性.因为有功功率可以支撑空间关系的研究,所以仅使用有功功率数据,该特征时间跨度为 2019 年 1 月 2 日至 2019 年 12 月 31 日.在不同时期内,风力机的有功功率相关性也在随时间不断变化.例如 1 月,1 号风力机与其他风力机的相关性较低;2 月,1 号与其他风力机的相关性都在 0.7 以上,而 9 号和 13 号风力机与其余风力机相关性较 1 月份明显降低;3 月,9 号和 10 号风力机与其余风力机的相关性较低.而

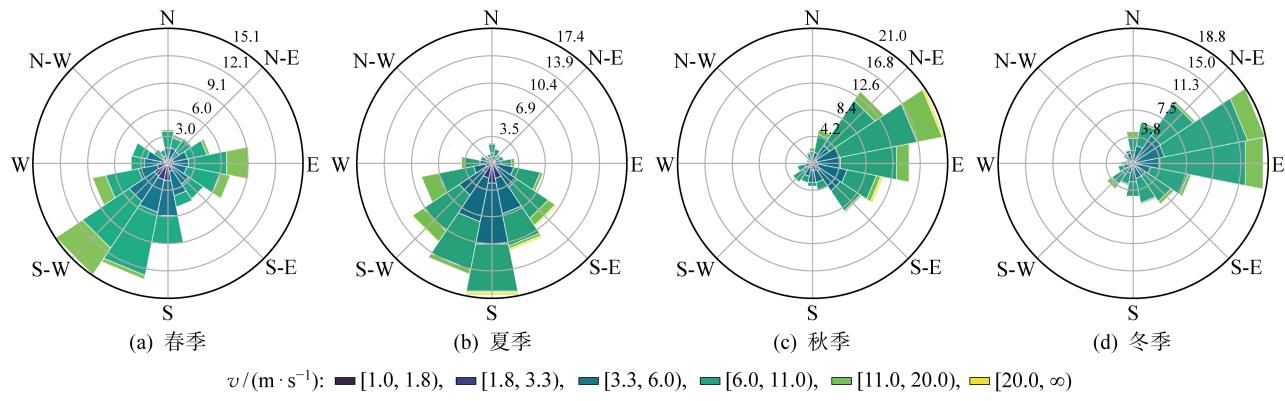


图 2 不同季节风向玫瑰图

Fig. 2 Rose diagram of wind direction in different seasons

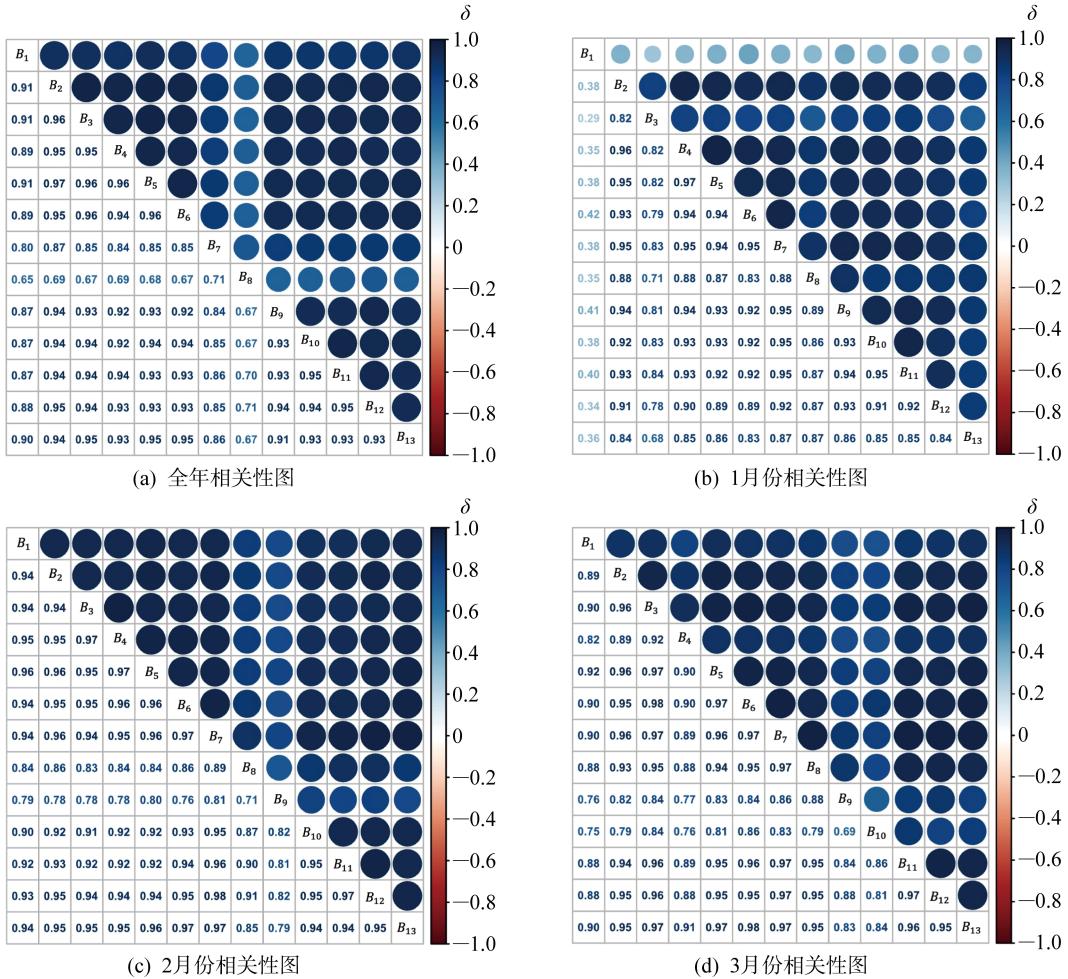


图 3 风力机有功功率相关性图

Fig. 3 Correlation of power output of turbines

从全年来看,8号风力机与其余风力机相关性较低.

## 2 研究方法

### 2.1 在线 LASSO VAR 模型

VAR 模型可以分析风力机有功功率间的空间相关性,并且计算效率高、效果好,可以用于有功功

率的点预测<sup>[24]</sup>,设  $P_{t,i}$  为风力机  $i$  在时刻  $t$  的有功功率,  $\mathbf{P}_t$  为有功功率向量,包含  $t$  时刻所有风力机的有功功率数据,引入 VAR 模型<sup>[25]</sup>则有

$$\mathbf{P}_t = \sum_{l=1}^L \mathbf{A}_l \mathbf{P}_{t-l} + \boldsymbol{\epsilon} \quad (1)$$

式中:  $\mathbf{A}_l$  为系数矩阵;  $l$  为滞后阶数;  $L$  为模型阶数;

白噪声项  $\epsilon_t$  假设服从多元正态分布,  $\epsilon_t \sim (\mathbf{0}, \Sigma)$ ,  $\Sigma$  为协方差矩阵。通常,  $\mathbf{A}_t$  的估计可以通过最小化训练集上的均方误差得到。然而, 随着维度增大, 所需拟合的参数数量增长速度是非线性的。因此, 当训练数据不足时, 大量待拟合参数会带来严重的过拟合问题, 从而导致预测效果不佳。

LASSO 正则化可以有效地解决过拟合问题。文献[26]中提出风能预测中 LASSO VAR 的应用, LASSO VAR 的损失函数为

$$\sum_{t=1}^T \|\mathbf{P}_t - \hat{\mathbf{P}}_t\|_2^2 + \lambda \sum_{l=1}^L \|\mathbf{A}_l\|_1 \quad (2)$$

式中:  $\hat{\mathbf{P}}_t$  为  $\mathbf{P}_t$  的预测值;  $\|\cdot\|_1$  为矩阵 1-范数,  $\|\cdot\|_2$  为向量 2-范数,  $T$  为求和的时间长度。式(2)中的损失函数可以使  $\mathbf{A}_t$  有解析解。正则化参数  $\lambda$  可以控制  $\mathbf{A}_t$  的稀疏程度, 过滤掉部分不重要的相关系数, 从而使  $\mathbf{A}_t$  中只包含对预测影响最明显的一部分系数。式(2)的最优化可以由循环坐标下降法求解<sup>[27]</sup>。循环坐标下降法在每次求解中固定其他系数, 只对损失函数求解一个系数的最优值, 不断循环迭代直至收敛。在每次求解中, 系数  $A_{l,i,j}$  可由下式更新:

$$A_{l,i,j} =$$

$$\frac{S(\sum_{t=1}^T P_{t-l,j}(P_{t,i} - \hat{P}_{t,i} + A_{l,i,j}P_{t-l,j}), \lambda)}{\sum_{t=1}^T P_{t-l,j}^2} \quad (3)$$

$$S(z, \lambda) = \begin{cases} z - \lambda, & z > 0 \text{ 且 } \lambda < |z| \\ z + \lambda, & z < 0 \text{ 且 } \lambda < |z| \\ 0, & \lambda > |z| \end{cases}$$

式中:  $\hat{P}_{t,i}$  为  $t$  时刻风力机  $i$  功率的预测值;  $S(\cdot)$  为软阈值算子。

为了使模型更好地捕捉到有功功率序列的变动, 每当获得新的有功功率数据时, 系数矩阵  $\mathbf{A}_t$  都需要被重新估计。然而, 重新估计一个高维的 LASSO VAR 模型非常耗时。因此, 基于文献[28]中的工作, 对有功功率 LASSO VAR 模型采用如下在线估计方法。该方法重新将  $\mathbf{A}_t$  更新公式表示为

$$A_{l,i,j} = \frac{S(\mathbf{A}, \lambda)}{\mathbf{Z}_{l,l,T,j}} \quad (4)$$

$$\left. \begin{aligned} \mathbf{A} &= \mathbf{Z}_{0,l,T,i,j} - \sum_{m=1}^L \mathbf{Z}_{m,l,T,j}^T \mathbf{A}_{m,i}^T + \\ &\quad \mathbf{Z}_{l,l,T,j} A_{l,i,j} \\ \mathbf{Z}_{l,l,T,j,j} &= \sum_{t=1}^T P_{t-l,j}^2 \end{aligned} \right\} \quad (5)$$

式中:  $\mathbf{A}$  为式(3)中软阈值算子的第一项;  $\mathbf{A}_{m,i}$  为模型的系数矩阵;  $\mathbf{Z}$  为两台风力机的功率数据相乘并

求和后的表示算子, 下标的前两项为两风力机历史时间的索引, 下标的末两项为两风力机的索引, 可以用矩阵表示为

$$\mathbf{Z}_{m,l,T} = \sum_{t=1}^T \mathbf{P}_{t-m} \mathbf{P}_{t-l}^T$$

$\mathbf{Z}_{m,l,T}$  可以变换为

$$\mathbf{Z}_{m,l,T} = \mathbf{Z}_{m,l,T-1} + \mathbf{P}_{T-m} \mathbf{P}_{T-l}^T \quad (6)$$

因为  $\mathbf{Z}_{m,l,T-1} = \sum_{t=1}^{T-1} \mathbf{P}_{t-m} \mathbf{P}_{t-l}^T (m = 0, 1, \dots, L; l = 1, 2, \dots, L)$  可以由之前的时间步直接得到, 所以, 只需要存储  $L$  步的数据与  $\mathbf{Z}_{m,l,T-1}$ , 就可以得到  $\mathbf{Z}_{m,l,T}$ , 进而可以快速更新  $\mathbf{A}_t$  的元素。

超参数  $\lambda$  的选取沿用文献[28]中的方法, 即给定一个  $\lambda$  的取值集合, 在每一次更新中, 挑选出使  $\sum_{t=1}^{T-1} \|\mathbf{P}_t - \hat{\mathbf{P}}_t\|_2^2$  最小的  $\lambda$ 。

在时刻  $t$ , 给定系数矩阵  $\mathbf{A}_l, l = 1, 2, \dots, L$ , 在线 LASSO VAR 在时刻  $t+1$  的有功功率预测向量  $\mathbf{P}_{t+1} = [P_{t+1,1} \ P_{t+1,2} \ \dots \ P_{t+1,Q}]^T$  可以估计为

$$\hat{\mathbf{P}}_{t+1} = \sum_{l=1}^L \mathbf{A}_l \mathbf{P}_{t+1-l} \quad (7)$$

## 2.2 异方差模型

**2.2.1 ARCH 检验** 在线 LASSO VAR 模型预测有功功率时, 假设误差项为白噪声, 方差为常数, 即  $\epsilon \sim N(\mathbf{0}, \Sigma)$ 。但实际上, 有功功率预测的误差项可能和残差历史值以及时间相关, 即具有异方差效应。

传统的检验异方差效应的方法主要包括图示检验法、Breusch-Pagan 检验、Goldfeld-Quandt 检验等<sup>[29]</sup>。但上述异方差检验方法主要针对多元回归问题, 对时间序列预测问题而言, ARCH 检验更合适。因此, 采用 ARCH 检验验证风力机有功功率预测的异方差性。

采用 ARCH 检验对在线 LASSO VAR 模型得到的所有风力机有功功率残差向量  $a_t$  进行检验。单台风力机对应的有功功率残差值  $a_t = P_t - E(P_t | F_{t-1})$ , 其中  $P_t$  为时刻  $t$  所有风力机有功功率之和,  $F_{t-1}$  为时刻  $t-1$  时所有信息的信息集,  $E(P_t | F_{t-1})$  为时刻  $t$  的在线 LASSO VAR 预测的有功功率值。构造如下自回归模型:

$$a_t^2 = \alpha_0 + \sum_{l=1}^q \alpha_l a_{t-l}^2 + e_t, \quad t = q+1, q+2, \dots, M \quad (8)$$

式中:  $e_t$  为误差项;  $q$  为模型阶数;  $M$  为样本量。用普通最小二乘法估计自回归模型的参数并计算预测  $t$  时刻的残差  $\hat{e}_t$ , 令回归模型对因变量的拟合程度和残差平方和分别为

$$e_{\text{SSR}_0} = \sum_{t=q+1}^M (a_t^2 - u)^2, \quad e_{\text{SSR}_1} = \sum_{t=q+1}^M \hat{e}_t^2$$

其中  $u = \frac{1}{M} \sum_{t=1}^M a_t^2$ , 得到检验统计量为

$$F = \frac{(e_{\text{SSR}_0} - e_{\text{SSR}_1})/q}{e_{\text{SSR}_1}/(M-2q-1)} \quad (9)$$

零假设为参数  $\alpha_1 = \alpha_2 = \dots = \alpha_q = 0$ , 零假设成立时  $F$  近似服从  $F(q, M-2q-1)$ ; 否则, 拒绝零假设, 认为有功功率预测值存在异方差效应.

**2. 2. 2 EGARCH 模型 GARCH 模型<sup>[30]</sup>** 是 ARCH 模型<sup>[31]</sup> 的推广, 可以修正点预测的异方差性, 计算点预测的不确定性. GARCH 模型对在线 LASSO VAR 模型的误差项  $\epsilon$  建模, GARCH( $p'$ ,  $q$ ) 过程由下式给出:

$$\left. \begin{aligned} a_t | F_{t-1} &= \epsilon_t \sqrt{h_t} \\ h_t &= \alpha_0 + \sum_{i=1}^{p'} \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} \end{aligned} \right\} \quad (10)$$

式中:  $\epsilon_t$  为零均值单位方差的独立同分布白噪声, 其分布常取标准正态分布、标准化的  $t$  分布、广义误差分布;  $a_t^2$  的模型阶数  $p' > 0$ ,  $q \geq 0$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ , 自回归部分待求解的模型参数  $\beta_j \geq 0$ ; 条件方差  $D(P_t | F_{t-1}) = D(a_t | F_{t-1}) = h_t$ , 即有功功率预测的条件方差可以由有功功率残差平方  $a_t^2$  滞后值的线性组合表示.

类似于经济学中, 回报低于预期时股市波动率往往上升, 回报高于预期时股市波动率往往下降<sup>[32]</sup>; 风能概率预测中, 有功功率点预测的正负偏差对未来波动率的影响也可能不同, 后续将从数据维度进行验证. 基于上述分析, 因为正负偏差对波动率有不对称影响, 所以使用 EGARCH 模型<sup>[33]</sup> 对有功功率的条件方差进行建模更合适.

EGARCH( $p'$ ,  $q$ ) 模型中, 有功功率残差  $a_t$  的分布与 GARCH 模型相同, 为式(10). 令  $g(\epsilon_t) = \alpha \epsilon_t + \gamma(|\epsilon_t| - E|\epsilon_t|)$ ,  $g(\epsilon_t)$  的期望为 0,  $\mathbf{z}'_t = (1, g(\epsilon_{t-1}), \dots, g(\epsilon_{t-p'}), \ln h_{t-1}, \ln h_{t-2}, \dots, \ln h_{t-q})$ ,  $\boldsymbol{\omega}^\top = (\alpha_0, \alpha_1, \dots, \alpha_p, \beta_1, \beta_2, \dots, \beta_q)$ , EGARCH 模型的  $h_t$  由下式给出:

$$\ln h_t = \mathbf{z}'_t \boldsymbol{\omega}, \quad \beta_1 \equiv 1 \quad (11)$$

式中:  $\mathbf{z}'_t$  为输入变量, 为  $\mathbf{z}_t$  的转置;  $\alpha$ 、 $\gamma$  和  $\boldsymbol{\omega}$  为待学习的参数.

EGARCH 模型的参数通常采用极大似然估计法进行估计. 假设  $\epsilon_t \sim N(0, 1)$ , EGARCH 模型的对数似然损失  $L(\boldsymbol{\omega})$  为

$$\left. \begin{aligned} L(\boldsymbol{\omega}) &= \frac{\sum_{t=1}^M l_t(\boldsymbol{\omega})}{M} \\ l_t(\boldsymbol{\omega}) &= -\frac{1}{2} \ln h_t - \frac{1}{2} a_t^2 h_t^{-1} \end{aligned} \right\} \quad (12)$$

式中:  $l_t(\boldsymbol{\omega})$  为模型参数为  $\boldsymbol{\omega}$  时  $t$  时刻的损失值, 对参数  $\boldsymbol{\omega}$  进行求导有

$$\frac{\partial l_t}{\partial \boldsymbol{\omega}} = \frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial \boldsymbol{\omega}} \left( \frac{a_t^2}{h_t} - 1 \right) \quad (13)$$

$$\frac{\partial h_t}{\partial \boldsymbol{\omega}} =$$

$$h_t \left( \mathbf{z}_t + \sum_{j=1}^{p'} \alpha_j \frac{\partial g(\epsilon_{t-j})}{\partial \boldsymbol{\omega}} + \sum_{i=1}^q \beta_i \frac{1}{h_{t-i}} \frac{\partial h_{t-i}}{\partial \boldsymbol{\omega}} \right) \quad (14)$$

具体估计细节可以参考文献[34], 另外可以采用文献[35]中提出的迭代步骤来获得极大似然估计量.  $\boldsymbol{\omega}_o$  表示第  $o$  次迭代后的参数值,  $\boldsymbol{\omega}_{o+1}$  可由下式计算:

$$\boldsymbol{\omega}_{o+1} = \boldsymbol{\omega}_o + \lambda_o \left( \sum_{t=1}^M \frac{\partial l_t}{\partial \boldsymbol{\omega}} \frac{\partial l_t}{\partial \boldsymbol{\omega}}^\top \right)^{-1} \sum_{t=1}^M \frac{\partial l_t}{\partial \boldsymbol{\omega}} \quad (15)$$

式中:  $\partial l_t / \partial \boldsymbol{\omega}$  在  $\boldsymbol{\omega}_o$  处估计,  $\lambda_o$  为一个可变的步长, 选定  $\lambda_o$  以最大化给定方向上的似然函数, 给定方向由  $\partial l_t / \partial \boldsymbol{\omega}$  的  $T \times 1$  向量的最小二乘回归得到.

## 2. 3 风场的有功功率概率预测

一般来说, 风场中的风力机作为整体, 针对特定区域供电, 就风力机风能预测问题而言, 更关注整个风场的有功功率预测. 因此, 在得到单个风力机的有功功率后, 需要将所有风力机的有功功率求和. 不同风力机间的有功功率相关系数计算公式为

$$\rho_{ij} = \frac{\sum_{t=1}^M (P_{t,i} - u_i)(P_{t,j} - u_j)}{\sqrt{\sum_{t=1}^M (P_{t,i} - u_i)^2} \sqrt{\sum_{t=1}^M (P_{t,j} - u_j)^2}} \quad (16)$$

式中:  $\rho_{ij}$  为风力机  $i$ 、 $j$  有功功率的相关系数;  $u_i$ 、 $u_j$  分别为风力机  $i$ 、 $j$  有功功率的平均值. 而存在相关性的风力机有功功率求和的均值和方差计算公式为

$$E\left(\sum_{i=1}^n P_{t,i} | F_{t-1}\right) = E(P_{t,1} | F_{t-1}) + \dots + E(P_{t,n} | F_{t-1}) \quad (17)$$

$$\left. \begin{aligned} D\left(\sum_{i=1}^n P_{t,i} | F_{t-1}\right) &= \\ \sum_{i=1}^n D(P_{t,i} | F_{t-1}) + 2 \sum_{i \neq j} O_{i,j,t} &= \\ \sum_{i=1}^n h_{t,i} + 2 \sum_{i \neq j} O_{i,j,t} &= \\ O_{i,j,t} = \rho_{ij} \sqrt{D(P_{t,i} | F_{t-1})} \times & \\ \sqrt{D(P_{t,j} | F_{t-1})} = & \\ \rho_{ij} \sqrt{h_{t,i}} \sqrt{h_{t,j}} & \end{aligned} \right\} \quad (18)$$

式中:  $E(P_{t,i} | F_{t-1})$  为在线 LASSO VAR 预测时刻  $t$  时风力机  $i$  有功功率的条件均值;  $D(P_{t,i} | F_{t-1})$  为风力机  $i$  有功功率的条件方差;  $O_{i,j,t}$  为时刻  $t$  时风力机  $i$  和  $j$  有功功率的协方差预测值。由式(17)和式(18)可以得到整个风场有功功率及其不确定性。

### 3 模型评估

从三方面评估新模型的效果:在线 LASSO VAR 方法对有功功率的点预测效果、EGARCH 模型对单风力机有功功率的概率预测效果、考虑功率相关性的整个风场的点预测与概率预测效果。实例研究的代码基于 Windows 系统的 R version 4.0.2 执行。

#### 3.1 在线 LASSO VAR 有功功率点预测

首先使用在线 LASSO VAR 对 24 台风力机 2019 年全年的有功功率进行建模,时间分辨率为 1 h,共有 8 402 个观测。为了方便在线 LASSO VAR 模型与其他时空模型进行对比,将前 7 000 个观测归为训练集,最后的 1 402 个观测作为测试集。使用均方根误差(RMSE)、平均绝对误差(MAE)与对称平均绝对百分比误差(SMAPE)作为评估指标,其数学表达式如下:

$$\left. \begin{aligned} \alpha_{\text{RMSE}} &= \sqrt{\frac{1}{M} \sum_{t=1}^M \| \mathbf{P}_t - \hat{\mathbf{P}}_t \|_2^2} \\ \alpha_{\text{MAE}} &= \frac{1}{M} \sum_{t=1}^M \| \mathbf{P}_t - \hat{\mathbf{P}}_t \|_1 \\ \alpha_{\text{SMAPE}} &= \frac{1}{M} \sum_{t=1}^M \frac{2 | \mathbf{P}_t - \hat{\mathbf{P}}_t |}{\mathbf{P}_t + \hat{\mathbf{P}}_t} \end{aligned} \right\} \quad (19)$$

从在线学习、稀疏性、气象数据 3 个维度验证在线 LASSO VAR 模型的有效性。在线学习维度中,使用在线自回归(Autoregressive, AR)模型、Batch LASSO VAR 模型<sup>[28]</sup>进行对比;在稀疏性维度中,使用弹性网(Elastic Net, ENET)、分层稀疏(Hierarchical Sparse)、平滑削边绝对偏离(Smoothly Clipped Absolute Deviation, SCAD)进行对比。在气象数据维度中,使用基于风速和风向的功率曲线进行对比。

表 1 为在线维度中各时序模型的预测结果。模型后的数字代表模型的最高阶数。Batch LASSO VAR 的数学模型与在线 LASSO VAR 相同,但模型参数只在获得所有训练集数据后更新一次。出于对模型估计效率的考虑,对于 VAR 模型,令最高时间阶数  $L$  为 4 阶。由表 1 可见,VAR 模型在各个指标上都远远优于非 VAR 模型,在线 AR 模型的预

测效果在所有模型中最差,这主要由于在线 AR 模型只考虑单风力机的时间相关性,未考虑多风力机间的空间相关性。上述结果验证了考虑空间相关性的 VAR 模型的有效性。

表 1 在线维度中各模型预测结果

Tab. 1 Forecasting results in online dimension

模型	RMSE	MAE	SMAPE
在线 LASSO VAR(4)	397.08	272.03	55.65
在线 LASSO VAR(3)	397.24	272.27	53.00
在线 LASSO VAR(2)	406.98	280.50	54.16
在线 LASSO VAR(1)	406.80	280.10	54.04
Batch LASSO VAR(4)	399.05	273.76	53.07
Batch LASSO VAR(3)	399.31	274.03	53.11
Batch LASSO VAR(2)	399.64	274.17	53.16
Batch LASSO VAR(1)	408.54	279.34	53.99
在线 AR(4)	420.52	287.19	55.05

1 阶的两种 VAR 模型效果相差无几,但更高阶的在线 VAR 模型优于 Batch VAR。同时,由于 Batch LASSO VAR 模型需要对所有的训练数据进行拟合优化,而在线 LASSO VAR 每次只需要利用到距离当前时刻最近的一部分训练数据,所以后的运行效率远远高于前者。拟合一个在线更新的 3 阶 LASSO VAR 模型只需要 20.8 s,而拟合一个同阶的 Batch LASSO VAR 模型却需要 659.6 s,这验证了在线方法的优越性。

图 4 为在线 LASSO VAR(4)模型的系数矩阵图,  $A_1 \sim A_4$  分别为系数。图 4(a)和 4(c)展示了在线 LASSO VAR(4)一阶滞后和三阶滞后的系数矩阵,可以看出大多数系数的绝对值都小于 0.2,这表明 LASSO VAR 削弱了无关风力机间的数据相关性。图 4(b)和 4(d)为四阶在线 LASSO VAR(4)的系数矩阵。空白色块表示等于 0 的系数,这意味着 LASSO 确实限制了非零系数的数量。由图 4(d)可见,在线 LASSO VAR(4)的四阶滞后系数矩阵中非零系数很少,结合预测效果分析,这代表考虑第四阶的滞后已很难从时空样本中提取到有利于改善预测的信息,从而未能明显提升 LASSO VAR 的效果,因此选择滞后阶数 3 是合理的。

综合考虑预测精度与运行效率,在线 LASSO VAR(3)的表现最优,因此,后续利用该模型的预测结果进行稀疏性、天气数值预测(Numerical Weather Prediction, NWP)数据维度的比较,以及风场总功率的概率预测。

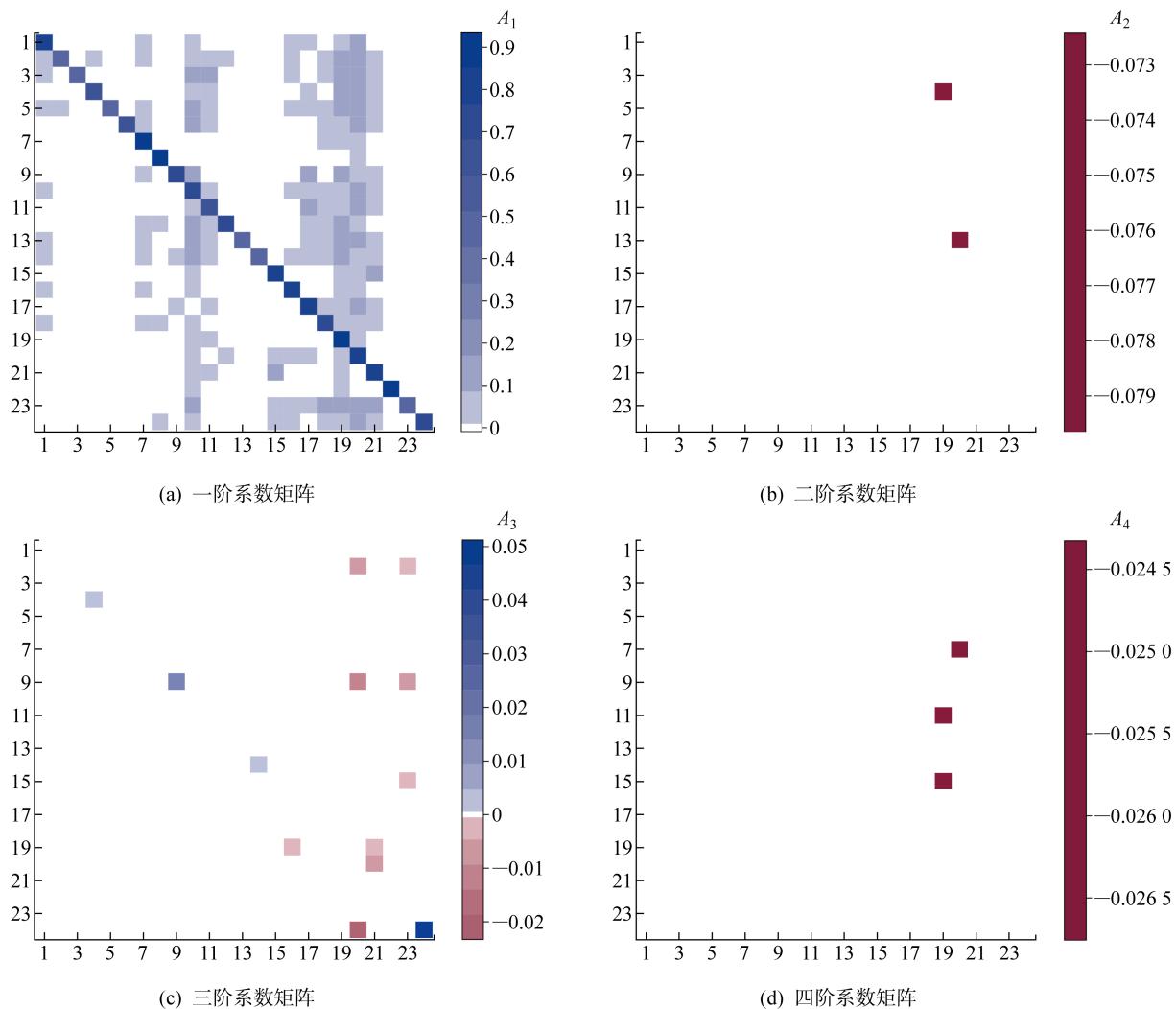


图 4 在线 LASSO VAR(4) 的系数矩阵

Fig. 4 Coefficient matrix of online LASSO VAR(4)

表 2 为稀疏性、气象数据维度中各时序模型的预测结果。稀疏性方法主要有特征降维和正则化矩阵两种。特征降维主要有主成分分析、T-SNE 等,但本文研究的时空预测问题输入为各风力机的有功功率数据,特征降维得到的新特征解释性差,不适合对输入矩阵进行稀疏处理。针对 VAR 模型的正则化矩阵主要有 LASSO、弹性网、层次稀疏、SCAD 等。使用后 3 种方法替换 LASSO 重新进行预测,发现 SCAD 的计算量很大、运行效率低,同样的数据量下,运行时间为 LASSO 的 120 倍,不适合进行大规模时空预测的稀疏处理;弹性网和层次稀疏的运行时间与 LASSO 相近,但没有提升预测结果。因此,使用 LASSO 稀疏参数矩阵是合理的。

为了研究不使用气象数据是否会影响预测结果,一方面使用在线 LASSO VAR(3) 模型预测风速和风向,并基于风速和风向构建功率曲线。在线 LASSO VAR(3) 模型已被证明能有效进行时空预

表 2 稀疏性、气象数据维度中各模型预测结果

Tab. 2 Forecasting results in sparse and meteorological data dimensions

模型	RMSE	MAE	SMAPE
LASSO VAR(3)	399.31	274.03	53.11
ENET VAR(3)	399.03	278.56	53.90
Hierarchical Sparse VAR(3)	418.82	295.31	56.55
气象预测 + 功率曲线	449.6	308.56	54.22
添加气象数据的 LASSO VAR(3)	407.1	280.4	54.12

测;基于风速和风向构建功率曲线,预测线性度  $R^2$  达到 0.96。由表 2 的预测结果可以看出,使用该方法得到的功率预测结果并不好,这主要是因为影响风力机功率输出的因素有很多,只使用风速和风向构造功率曲线,精度很难保证。另一方面,将风力机的风速和风向数据添加到只包含有功功率的输入矩阵中,运行算法发现预测结果反而变差。综上可知,

如果各风力机间的功率存在较强的时空相关性,不考虑气象数据不会影响模型预测结果。

在测试集中,以 35 号风力机为例,在线 LASSO

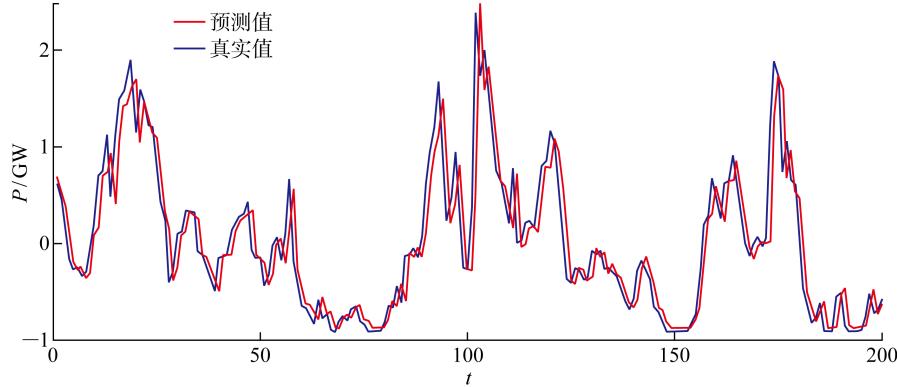


图 5 35 号风力机的在线 LASSO VAR(3)预测结果

Fig. 5 Forecasting result of online LASSO VAR (3) of Turbine 35

### 3.2 EGARCH 模型概率预测

可靠度、锐度、预报技巧分和连续概率排位分数 (Continuous Ranked Probability Score, CRPS) 是评估风场总功率概率预测结果的常用指标<sup>[36]</sup>,各指标的定义如下。

可靠度的定义为

$$R^{1-\beta} = \frac{1}{M} \sum_{t=1}^M \eta_t^{1-\beta} - (1-\beta) \quad (20)$$

式中:  $1-\beta$  为名义覆盖率,其预测区间为  $[\hat{q}_t^{\beta/2}, \hat{q}_t^{1-\beta/2}]$ ,其中  $\beta$  表示未在置信区间内的概率;  $P_t$  为  $t$  时刻的风场总功率。当  $\hat{q}_t^{\beta/2} < P_t < \hat{q}_t^{1-\beta/2}$  时,  $\eta_t^{1-\beta} = 1$ ; 其余情况下,  $\eta_t^{1-\beta} = 0$ 。

锐度的定义为

$$S^{1-\beta} = \frac{1}{M} \sum_{t=1}^M \delta_t^{1-\beta} \quad (21)$$

式中:  $\delta_t^{1-\beta} = \hat{q}_t^{1-\beta/2} - \hat{q}_t^{\beta/2}$ 。锐度是名义覆盖率预测区间的平均宽度。锐度反映了预测区间凝聚概率信息的能力,相同可靠度下,锐度越小的区间效果越好。

预报技巧分的定义为

$$K_t = \sum_{k=1}^r (\xi_t^{\alpha_k} - \alpha_k)(P_t - \hat{q}_t^{\alpha_k}) \quad (22)$$

式中:  $r$  为预测分位数的个数;  $\alpha_k$  为置信度; 当  $P_t \leq \hat{q}_t^{\alpha_k}$  时,  $\xi_t^{\alpha_k} = 1$ , 否则  $\xi_t^{\alpha_k} = 0$ .  $K_t$  反映  $t$  时刻所有预测分位数的预测效果,技巧分越大代表区间预测越好,最大值为 0.

CRPS<sup>[37]</sup>的定义为

$$C(N(u, \sigma^2), x) = \sigma \left\{ \frac{x-u}{\sigma} \left[ 2\Phi\left(\frac{x-u}{\sigma}\right) - 1 \right] + 2\phi\left(\frac{x-u}{\sigma}\right) - \frac{1}{\sqrt{\pi}} \right\} \quad (23)$$

VAR(3)的预测值与真实值对比如图 5 所示,可以看出,预测值较好地追踪了真实值的趋势,预测结果较好。

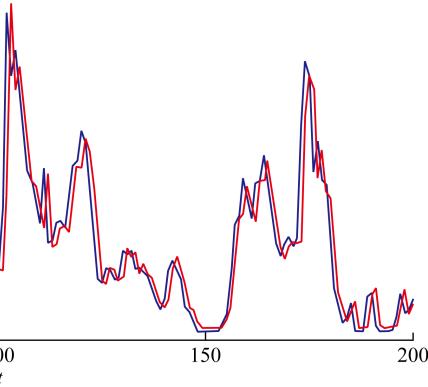


图 6 35 号风力机的有功功率与残差散点图

式中:  $u$  和  $\sigma^2$  分别为预测正态分布的均值和方差;  $x$  为真实值;  $\phi$  和  $\Phi$  分别为标准正态分布概率密度函数和累积概率密度函数。

首先对在线 LASSO VAR(3)的预测残差进行异方差检验。一是绘制有功功率和残差的散点图,以图 6 中的 35 号风力机为例,可以看出,随着有功功率值变大,残差散点整体呈现扩大的趋势,判断可能存在异方差效应。二是对预测残差作 ARCH 检验,发现 24 台风力机的检验  $p$  值均小于 0.05,说明在线 LASSO VAR(3)的预测残差存在异方差性。

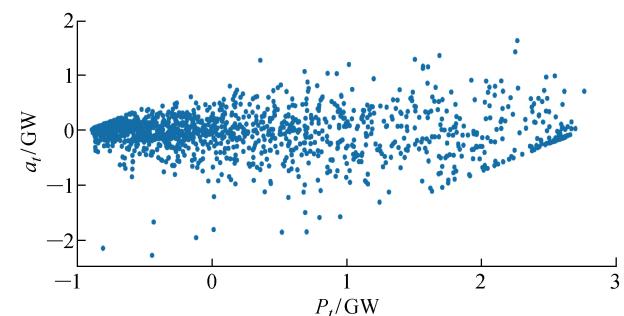


图 6 35 号风力机的有功功率与残差散点图

Fig. 6 Forecasting result of online LASSO VAR (3) of Turbine 35

接着使用 EGARCH(1, 1)模型对在线 LASSO VAR(3)验证集的残差值进行重新预测。假设 EGARCH(1, 1)模型的残差服从标准正态分布。35 号风力机拟合得到的 EGARCH(1, 1)中  $\ln h_t$  的表达式如下,其中各系数均在 0.01 水平下显著。

$$\begin{aligned} \ln h_t = & 1.255 + 0.271 \epsilon_{t-1} + \\ & 0.422 \left( |\epsilon_{t-1}| - \sqrt{\frac{\pi}{2}} \right) + 0.893 \ln h_{t-1} \end{aligned} \quad (24)$$

图 7 为 35 号风力机的 EGARCH(1, 1)模型预测结果. 可知, 虽在某些极值点中, EGARCH(1, 1)模型没能正确预测, 低估了真实值; 在某些残差很小的点中, EGARCH(1, 1)高估了真实值, 造成一定的浪费; 但总体来说, 其能较好地追踪残差的趋势, 预测结果较好.

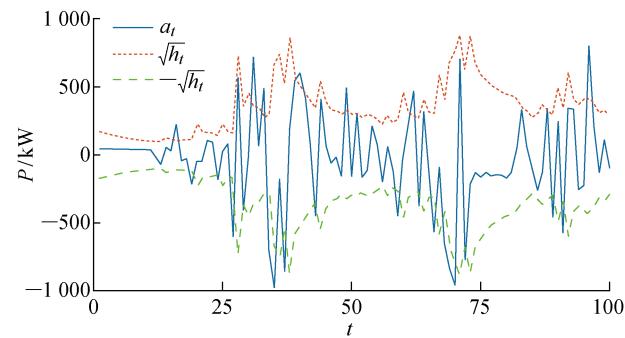
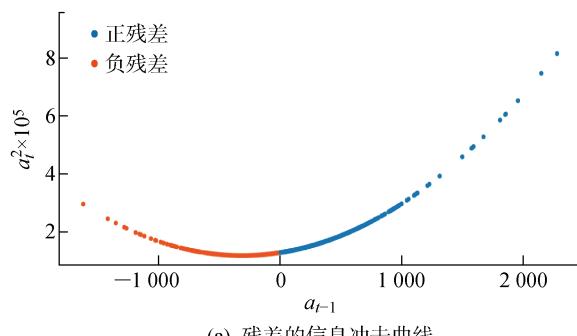


图 7 EGARCH(1, 1)模型预测 35 号风力机残差

Fig. 7 Residuals of Turbine 35 predicted by EGARCH(1, 1) model

得到 EGARCH(1, 1)模型的系数后, 使用信息冲击曲线和动态显著线验证 EGARCH 模型的合理性. 使用二次函数拟合残差和未来波动  $h_t$  的关系, 得到正负残差对未来波动  $h_t$  的不对称影响<sup>[38]</sup>, 如图 8 所示. 图 8(a)为残差的信息冲击曲线, 可以看出, 在绝对值相同的情况下, 正残差对应的下一个时间点的波动大于负残差对应的下一个时间点的波动. 图 8(b)为采用长度为 500 的滑动时间窗的方法, 使用不同的残差序列估计 EGARCH(1, 1)模型得到的对应  $\alpha$ , 其中横坐标为残差序列的索引值; 可以看出,  $\alpha$  均为正值, 并且在显著性水平为 0.05 时均显著. 综上可知, 图 8 验证了正负残差对条件方差的不对称影响.

假设单个风力机有功功率服从正态分布  $N(\mu, \sigma^2)$ ,  $\mu$  由 Online LASSO VAR(3)点预测值估



(a) 残差的信息冲击曲线

计,  $\sigma^2$  由 EGARCH(1, 1)模型预测结果估计. 考虑风场中各个风力机有功功率间的相关性, 对整个风场的总功率输出进行概率预测, 并将其与 KDE<sup>[39]</sup>、BMA 方法、QR 模型、分位数回归森林(Quantile Regression Forest, QRF)模型<sup>[40]</sup>模型进行比较. 其中, BMA 模型只选取与当前时刻最近的  $M$  个观测训练. 本文选取  $M$  为 24 和 48, 分别对应 1 d 与 2 d 的数据.

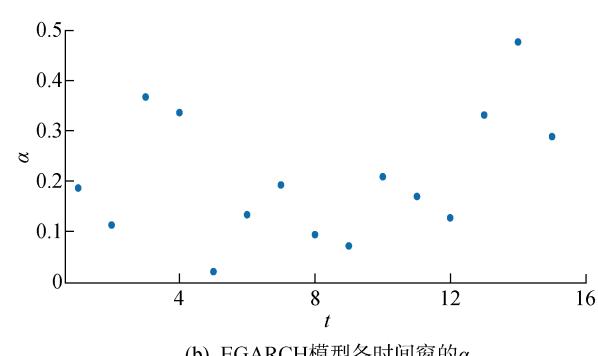
表 3 为各概率预测模型的结果. 可知, 考虑异方差的两个模型 EGARCH(1, 1)与 GARCH(1, 1)的 RMSE、预报技巧分、CRPS 相较于其他模型都有很大的提高. 就 RMSE 而言, 两个异方差模型小于 BMA 模型; 就预报技巧分而言, 异方差模型更接近 0. 同时, 就 CRPS 而言, 异方差模型优于其他模型. 两个异方差模型相比, 在中位数预测相同的情况下, EGARCH(1, 1)模型的预报技巧分和 CRPS 两个预测指标都要优于 GARCH(1, 1). 因此, 预报技巧分和 CRPS 的优势表明, EGARCH(1, 1)模型在概率预测上能取得非常好的效果.

图 9 为 EGARCH 模型与其他模型的结果对比图. 图 9(a)展示了各模型的可靠度. 可以看出, 除了

表 3 EGARCH 模型与其他模型的对比

Tab. 3 Comparison of EGARCH and other models

模型	中位数预测		CRPS
	RMSE	$K_t$	
EGARCH (1, 1)	8 019. 4	-40 401. 8	4 057. 5
GARCH (1, 1)	8 019. 4	-40 532. 8	4 070. 8
BMA_48	8 421. 7	-43 291. 0	4 389. 9
BMA_24	8 936. 9	-45 535. 7	4 640. 5
KDE	32 461. 5	-158 444	11 345. 67
QR	14 918. 7	-70 989. 3	7 201. 1
QRF	9 182. 3	-45 429. 4	4 594. 0



(b) EGARCH模型各时间窗的 $\alpha$

图 8 正负残差对波动性不同影响的验证

Fig. 8 Different effects of positive and negative residuals on volatility

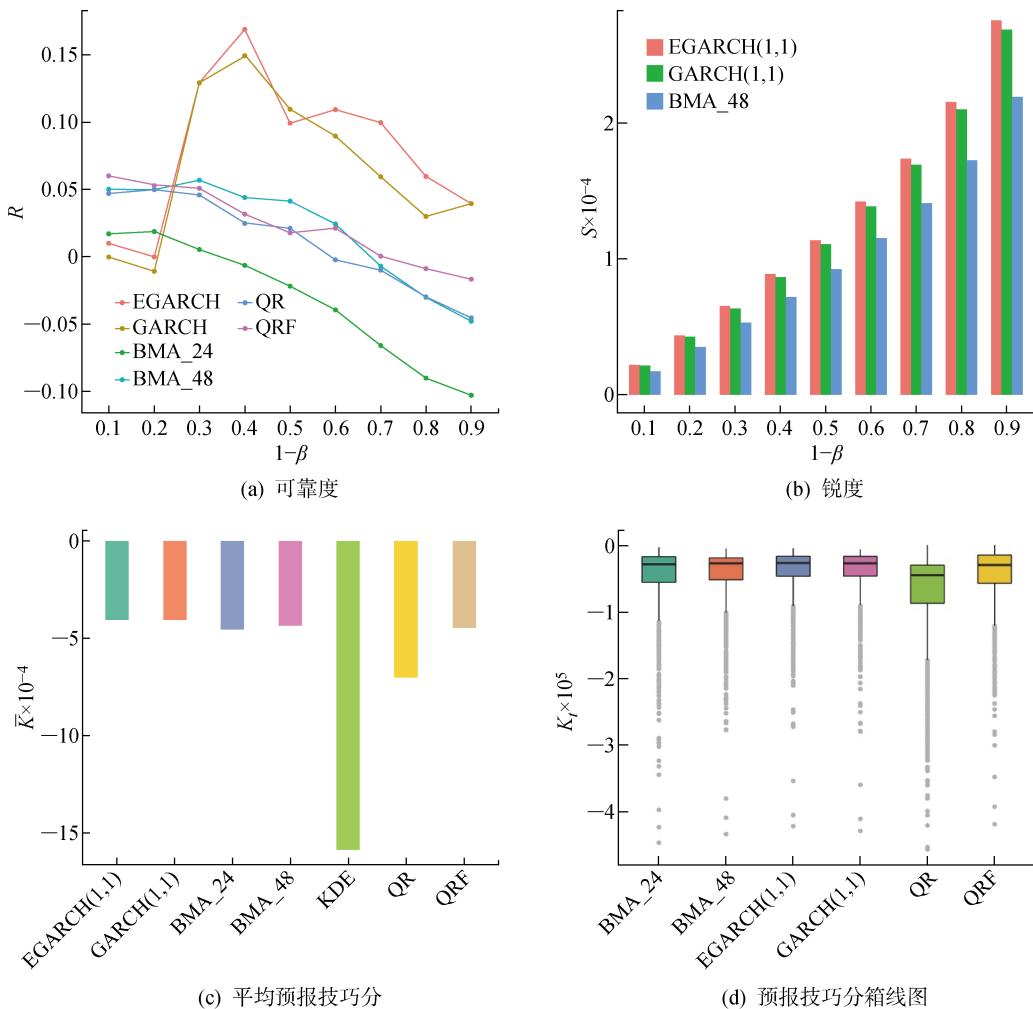


图 9 EGARCH 模型与其他模型的对比

Fig. 9 Comparison of EGARCH and other models

名义覆盖率为 10% 和 20%, EGARCH 模型和 GARCH 模型真实覆盖率与名义覆盖率的差明显高于其他模型. 图 9(b)展示了可靠性较高的 EGARCH 模型、GARCH 模型和 BMA\_48 模型的锐度. 可以看出, 异方差模型给出的区间宽度要高于 BMA\_48, 这可能导致了其预测区间对一些波动较大的点的捕获能力高, 最终得到较高的预报技巧分和 CRPS. 图 9(c)展示了各模型的平均预报技巧分 ( $\bar{K}$ ), 可以看出 EGARCH 模型最优, 并且异方差模型要明显好于非异方差模型. 图 9(d)展示了各模型的预报技巧分箱形图. 可以看出, 异方差模型的预报技巧分的箱形宽度较小, 均值更接近于 0, 预报技巧分的方差也较小, 这表明异方差模型的准确度和鲁棒性均高于其他模型.

图 10 展示了各模型在连续测试集上的预测区间, 红线表示真实的总功率  $P_t$ , 灰线表示模型的中位数预测  $\hat{P}_t$ . 名义覆盖率为 0.1~0.9 的预测区间也由不同的颜色表示, 较深的颜色表示较小的区间

覆盖率. 可以看出, EGARCH 模型的预测区间较窄, 而且能较好地覆盖真实值, 体现了 EGARCH 模型具有较好的概率预测效果. EGARCH 模型相较于 GARCH 模型, 在正偏差时的置信区间更宽、在负偏差时的置信区间更窄, 验证了 EGARCH 模型可以体现出正负残差对波动性的不同影响; 并且可以看出, 在两段功率明显上升导致正残差的时间段内, EGARCH 模型的预测区间比 GARCH 模型的预测区间宽, 覆盖了原始点, 这有助于 EGARCH 模型预测效果的提升. BMA 在功率变化较小的连续值中, 预测区间宽度较大, 而且在 25~50 h 中, 预测区间很窄, 不能很好地覆盖真实值. QR 和 QRF 模型的效果明显不如 EGARCH 模型和 BMA\_48 模型, 其预测区间较宽, 而且在部分连续时间段内未能很好地覆盖真实值. 综上可以看出, 采用 EGARCH 模型度量在线 LASSO VAR 模型的不确定性, 并将风场中各风力机的功率进行集成可以明显改进概率预测结果.

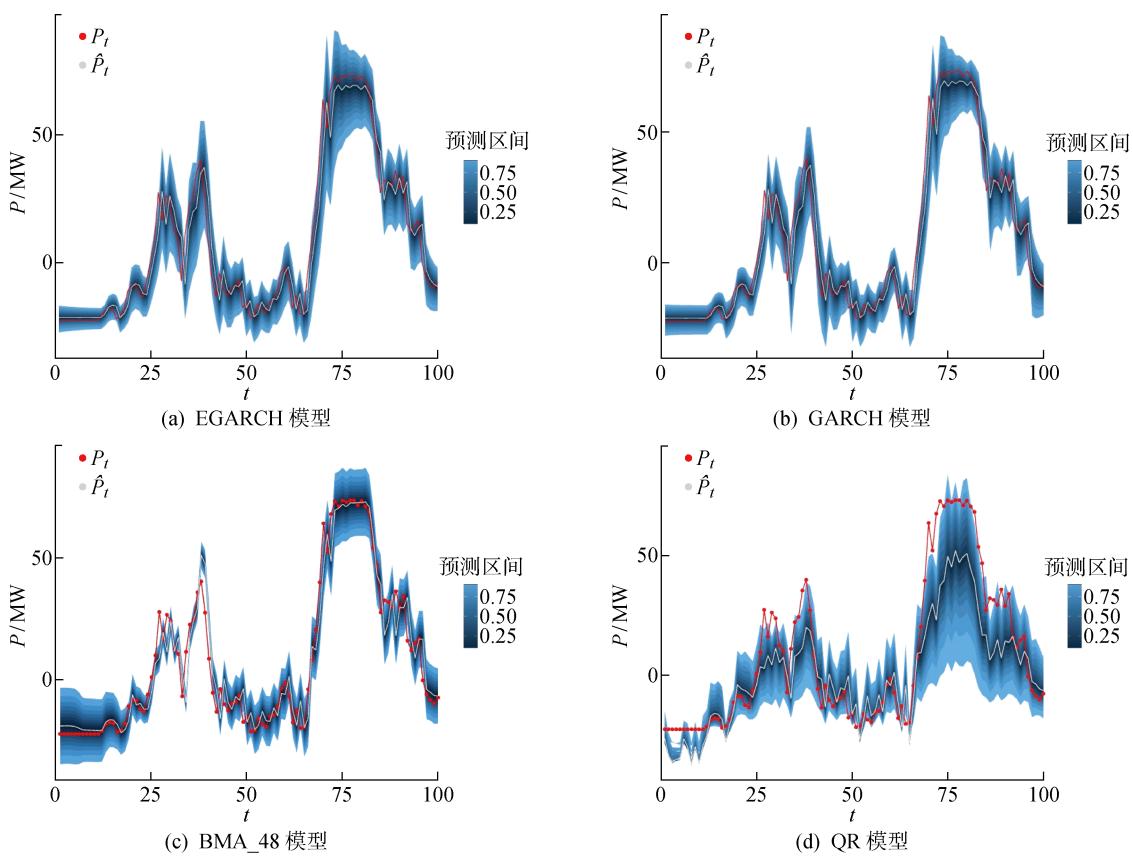


图 10 各模型在连续测试集上的预测区间

Fig. 10 Prediction interval of models on continuous test set

## 4 结论

针对风电不确定性高、现有模型未充分考虑异方差性的问题,使用在线 LASSO VAR 预测各风力机的有功功率,使用 EGARCH 预测残差,最终得到风场总有功功率的概率预测结果。将各模型在中国华东某地风场的功率预测效果进行比较,得出以下结论:

(1) 在线 LASSO VAR 模型进行时空预测,可以考虑到风力机间有功功率相关性的变化,在减小运算量的同时提升点预测的准确性。

(2) EGARCH 模型可以针对风能点预测模型的不对称性建模,预测结果相较于 GARCH 模型得到提升。

(3) 采用 EGARCH 模型预测在线 LASSO VAR 模型的条件方差,并将各风力机的有功功率求和,得到风场总有功功率,相较于 QR、QRF、KDE、BMA 直接对总有功功率建模,可以明显改进概率预测结果。

本文对风能区间预测、风能预测异方差处理、正负残差不对称性影响等问题具有一定的参考意义,

未来将对时空异方差进行进一步研究。

## 参考文献:

- [1] Global Wind Energy Council. Global wind report 2022[R/OL]. (2022-12-31)[2023-02-16]. <https://gwec.net/global-wind-report-2022/>.
- [2] LANDBERG L, WATSON S J. Short-term prediction of local wind conditions[J]. **Boundary-Layer Meteorology**, 1994, 70(1): 171-195.
- [3] FOCKEN U, LANGE M, WALDL H. Previento-A wind power prediction system with an innovative upscaling algorithm[C/OL]. (2001-01-01)[2021-07-30]. [https://www.researchgate.net/publication/250448566\\_Previento\\_-\\_A\\_Wind\\_Power\\_Prediction\\_System\\_with\\_an\\_Innovative\\_Upscaling\\_Algorithm](https://www.researchgate.net/publication/250448566_Previento_-_A_Wind_Power_Prediction_System_with_an_Innovative_Upscaling_Algorithm).
- [4] 方江晓, 周晖, 黄梅, 等. 基于统计聚类分析的短期风电功率预测[J]. **电力系统保护与控制**, 2011, 39(11): 67-73.  
FANG Jiangxiao, ZHOU Hui, HUANG Mei, et al. Short-term wind power prediction based on statistical clustering analysis[J]. **Power System Protection & Control**, 2011, 39(11): 67-73.
- [5] AASIM, SINGH S N, MOHAPATRA A. Repeated

- wavelet transform based ARIMA model for very short-term wind speed forecasting[J]. **Renewable Energy**, 2019, 136: 758-768.
- [6] YANG D Z. On post-processing day-ahead NWP forecasts using Kalman filtering[J]. **Solar Energy**, 2019, 182: 179-181.
- [7] 戚创创, 王向文. 考虑风向和大气稳定度的海上风电功率短期预测[J]. **电网技术**, 2021, 45(7): 2773-2780.
- QI Chuangchuang, WANG Xiangwen. Short-term prediction of offshore wind power considering wind direction and atmospheric stability[J]. **Power System Technology**, 2021, 45(7): 2773-2780.
- [8] 李永刚, 王月, 刘丰瑞, 等. 基于 Stacking 融合的短期风速预测组合模型[J]. **电网技术**, 2020, 44(8): 2875-2882.
- LI Yonggang, WANG Yue, LIU Fengrui, et al. Combination model of short-term wind speed prediction based on stacking fusion[J]. **Power System Technology**, 2020, 44(8): 2875-2882.
- [9] WANG Y, HU Q H, MENG D Y, et al. Deterministic and probabilistic wind power forecasting using a variational Bayesian-based adaptive robust multi-Kernel regression model[J]. **Applied Energy**, 2017, 208: 1097-1112.
- [10] 李洪涛, 马志勇, 芮晓明. 基于数值天气预报的风能预测系统[J]. **中国电力**, 2012, 45(2): 64-68.
- LI Hongtao, MA Zhiyong, RUI Xiaoming. Forecasting system based on numerical weather prediction [J]. **Electric Power**, 2012, 45(2): 64-68.
- [11] OKUMUS I, DINLER A. Current status of wind energy forecasting and a hybrid method for hourly predictions[J]. **Energy Conversion and Management**, 2016, 123: 362-371.
- [12] 钱政, 裴岩, 曹利宵, 等. 风电功率预测方法综述[J]. **高电压技术**, 2016, 42(4): 1047-1060.
- QIAN Zheng, PEI Yan, Cao Lixiao, et al. Review of wind power forecasting method[J]. **High Voltage Engineering**, 2016, 42(4): 1047-1060.
- [13] DOWELL J, PINSON P. Very-short-term probabilistic wind power forecasts by sparse vector autoregression[J]. **IEEE Transactions on Smart Grid**, 2016, 7(2): 763-770.
- [14] MESSNER J, PINSON P. Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting[J]. **International Journal of Forecasting**, 2019, 35(4): 1485-1498.
- [15] LANGE M. On the uncertainty of wind power predictions—Analysis of the forecast accuracy and statistical distribution of errors[J]. **Journal of Solar Energy Engineering**, 2005, 127(2): 177-184.
- [16] 刘兴杰, 谢春雨. 基于贝塔分布的风电功率波动区间估计[J]. **电力自动化设备**, 2014, 34(12): 26-30.
- LIU Xingjie, XIE Chunyu. Wind power fluctuation interval estimation based on beta distribution [J]. **Electric Power Automation Equipment**, 2014, 34(12): 26-30.
- [17] 李海燕. 基于数据挖掘与非线性分位数回归的风电功率概率密度预测方法[D]. 合肥: 合肥工业大学, 2018.
- LI Haiyan. Wind power probability density forecasting method based on data mining and non-linear quantile regression[D]. Hefei: Hefei University of Technology, 2018.
- [18] ZHANG Y, LIU K, QIN L, et al. Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods[J]. **Energy Conversion & Management**, 2016, 112: 208-219.
- [19] QIN X, CAO L, RUNDENSTEINER E A, et al. Scalable kernel density estimation-based local outlier detection over large data streams[J]. **Advances in Database Technology-EDBT**, 2019, 3: 421-432.
- [20] SLOUGHTER J M, GNEITING T, RAFTERY A E. Probabilistic wind speed forecasting using ensembles and Bayesian model averaging[J]. **Journal of the American Statistical Association**, 2010, 105(489): 25-35.
- [21] WANG G, JIA R, LIU J, et al. A hybrid wind power forecasting approach based on Bayesian model averaging and ensemble learning[J]. **Renewable Energy**, 2020, 145: 2426-2434.
- [22] 刘帅, 朱永利, 张科, 等. 基于误差修正 ARMA-GARCH 模型的短期风电功率预测[J]. **太阳能学报**, 2020, 41(10): 268-275.
- LIU Shuai, ZHU Yongli, ZHANG Ke, et al. Short-term wind power forecasting based on error correction arma-garch model[J]. **Acta Energiae Solaris Sinica**, 2020, 41(10): 268-275.
- [23] 李力行, 苗世洪, 涂青宇, 等. 考虑异方差效应的风电不确定性建模及其在调度中的应用[J]. **电力系统自动化**, 2020, 44(8): 36-47.
- LI Lixing, MIAO Shihong, TU Qingyu, et al. Modelling of wind power uncertainty considering heteroskedasticity effect and its application in power system dispatching [J]. **Automation of Electric Power Systems**, 2020, 44(8): 36-47.
- [24] POGGI P, MUSELLI M, NOTTON G, et al. Fore-

- casting and simulating wind speed in Corsica by using an autoregressive model[J]. **Energy Conversion & Management**, 2003, 44(20): 3177-3196.
- [25] 孙春顺, 王耀南, 李欣然. 小时风速的向量自回归模型及应用[J]. **中国电机工程学报**, 2008, 28(14): 112-117.
- SUN Chunshun, WANG Yaonan, LI Xinran. A vector autoregression model of hourly wind speed and its applications in hourly wind speed forecasting [J]. **Proceedings of the CSEE**, 2008, 28(14): 112-117.
- [26] CAVALCANTE L, BESSA R J, REIS M, et al. LASSO vector autoregression structures for very short-term wind power forecasting[J]. **Wind Energy**, 2017, 20(4): 657-675.
- [27] FRIEDMAN J, HASTIE T, HÖFLING H, et al. Pathwise coordinate optimization[J]. **The Annals of Applied Statistics**, 2007, 1(2): 302-332.
- [28] MESSNER J W, PINSON P. Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting[J]. **International Journal of Forecasting**, 2019, 35(4): 1485-1498.
- [29] 朱金蝶. 回归模型中异方差检验方法研究[D]. 太原: 山西大学, 2019.
- ZHU Jindie. Study on the methods of testing heteroscedasticity in regression model [D]. Taiyuan: Shanxi University, 2019.
- [30] ENGLE R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation[J]. **Econometrica**, 1982, 50(4): 987.
- [31] BOLLERSLEV T. Generalized autoregressive conditional heteroskedasticity[J]. **Journal of Econometrics**, 1986, 31(3): 307-327.
- [32] BLACK F. Studies of stock market volatility changes [J]. **Proceedings of the American Statistical Association Business & Economic Statistics Section**, 1976, 177-181.
- [33] NELSON D B. Conditional heteroskedasticity in asset returns: A new approach[J]. **Econometrica**, 1991, 59(2): 347-370.
- [34] HAFNER C M, LINTON O. An almost closed form estimator for the egarch model[J]. **Econometric Theory**, 2017, 33(4): 1013-1038.
- [35] BERNDT E K, HALL B H, et al. Annals of economic and social measurement [M]. USA: NBER, 1974: 653-665.
- [36] WANG Z, WANG W S, LIU C, et al. Probabilistic forecast for multiple wind farms based on regular vine copulas[J]. **IEEE Transactions on Power Systems**, 2018, 33(1): 578-589.
- [37] GRIMIT E P, GNEITING T, BERROCAL V J, et al. The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification [J]. **Quarterly Journal of the Royal Meteorological Society**, 2006, 132 (621C): 2925-2942.
- [38] CHEN H, GAO S. A study on the structure of asymmetric volatility in load series[C]// **2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies**. Nanjing, China: IEEE, 2008: 737-744.
- [39] 张晓英, 张晓敏, 廖顺, 等. 基于聚类与非参数核密度估计的风电功率预测误差分析[J]. **太阳能学报**, 2019, 40(12): 3594-3604.
- ZHANG Xiaoying, ZHANG Xiaomin, LIAO Shun, et al. Prediction error analysis of wind power based on clustering and non-parametric kernel density estimation[J]. **Acta Energiae Solaris Sinica**, 2019, 40(12): 3594-3604.
- [40] MEINSHAUSEN N, RIDGEWAY G. Quantile regression forests[J]. **Journal of Machine Learning Research**, 2006, 7(6): 983-999.

(本文编辑:王历历)