

基于滑动分块百分位数 Bootstrap 法的风电功率概率区间预测

杨锡运¹, 张 璜¹, 关文渊², 董德华¹

(1. 华北电力大学控制与计算机工程学院, 北京 102206; 2. 中船电子科技有限公司, 北京 100070)

摘 要: 提出一种基于非参数滑动分块百分位数 Bootstrap 法(MBPB)的风电功率概率区间预测方法。由于风功率数据存在显著的时间相依结构,该方法首先对预测功率进行等间隔划分,再以某区间内的预测误差序列为样本,借助滑动分块 Bootstrap 法(MBB)抽样产生多个伪样本,然后对伪样本数据通过滑动分块百分位 Bootstrap 法和四分位法相结合的统计推断生成一定置信水平下的误差上下限,进而得到该预测功率段内的概率预测区间。同时建立包含区间覆盖率和区间平均带宽的评价指标,通过将其与百分位法、百分位数 Bootstrap (PB)法的预测结果对比,表明基于 MBPB 的概率性预测区间的覆盖率更高,平均带宽更窄,精度更好且效果也更优。

关键词: 风电功率; 预测; 概率区间; 滑动分块百分位数 Bootstrap 法; 四分位数

中图分类号: TM71

文献标识码: A

0 引 言

风力发电的间歇性、波动性特性给电网安全稳定运行带来极大挑战^[1],对风电功率进行有效预测至关重要^[2,3]。目前风功率短期预测一般为确定性点预测,包括物理法和统计法^[4,5]。由于各种不确定因素影响,点预测在实际工程中通常存在误差,且无法给出预测数值出现的概率及波动范围。为满足决策者在电网规划、可靠性评估等方面的需要,采用风电功率的概率性区间预测法量化不确定性信息具有重要意义。

现阶段风电功率概率区间预测方法包括参数法和非参数法^[6]。参数法是假定预测误差服从某种分布形式,如高斯分布^[7,8]、贝塔分布^[9]等,由历史误差数据统计拟合,确定假设分布模型参数后预测。参数法对先验模型依赖性强,先验模型误差分布偏离将对预测精度产生较大影响;非参数法无需对样本数据做任何分布假设,稳健性好。文献[10]通过支持向量机选取回归函数,采用分位点回归技术

对风电功率波动区间进行分析。文献[11]基于稀疏贝叶斯学习理论构建预测模型,采用分量预测方式实现风场功率的短期预测。文献[12]提出用核密度估计法对风场历史数据进行拟合。上述非参数方法通常在大样本下效果较好。

由于风电功率预测误差序列的分布模型未知,理论解析式难以推导,本文提出一种基于滑动分块百分位数 Bootstrap (moving block percentile Bootstrap, MBPB)的非参数风电功率概率区间预测法。该方法无需对总体分布做先验假设,并且充分考虑到风场功率数据内部相依性,采用滑动分块 Bootstrap 法(moving block Bootstrap, MBB)在不同功率区间抽取预测误差数据构成伪样本;然后利用百分位数法求取单个伪样本置信区间上下限;最后采用四分位法对多个伪样本置信区间上下限进行综合,得到风电功率概率性预测区间。选择区间覆盖率和区间平均带宽两项指标值为评价标准,与其他方法对比仿真表明,该方法在预测准确度以及精度方面具有优越性。

收稿日期: 2016-10-13

基金项目: 国家自然科学基金(51677067); 中央高校基本科研业务费专项(2015MS32)

通信作者: 杨锡运(1973—), 女, 博士、教授, 主要从事新能源发电控制方面的研究。yangxiyun916@sohu.com

1 风电功率特性分析

1.1 风电功率预测误差分布

采用中国西北某风电场运行的风功率预测系统短期功率预测数据和风电场实际有功功率数据,统计其功率预测误差分布如图 1 所示。风场装机容量为 199.5 MW,数据统计时间长度为 2014 年下半年,采样间隔为 15 min。由图 1 可看出,在不同

功率段内的预测误差概率分布存在差异,而且部分差异较大。误差分布存在多种特性,如:有偏性、重尾性、多峰性。例如当风场预测功率小于 40 MW 时,误差的分布存在重尾现象;而当风场功率在 80~120 MW 范围内时,误差的分布呈现近似正态分布状态,部分分布向左偏移并倾斜。可见,不同功率段内误差概率密度分布较难用一个确定性分布模型描述。如果使用传统参数统计方法假定服从某先

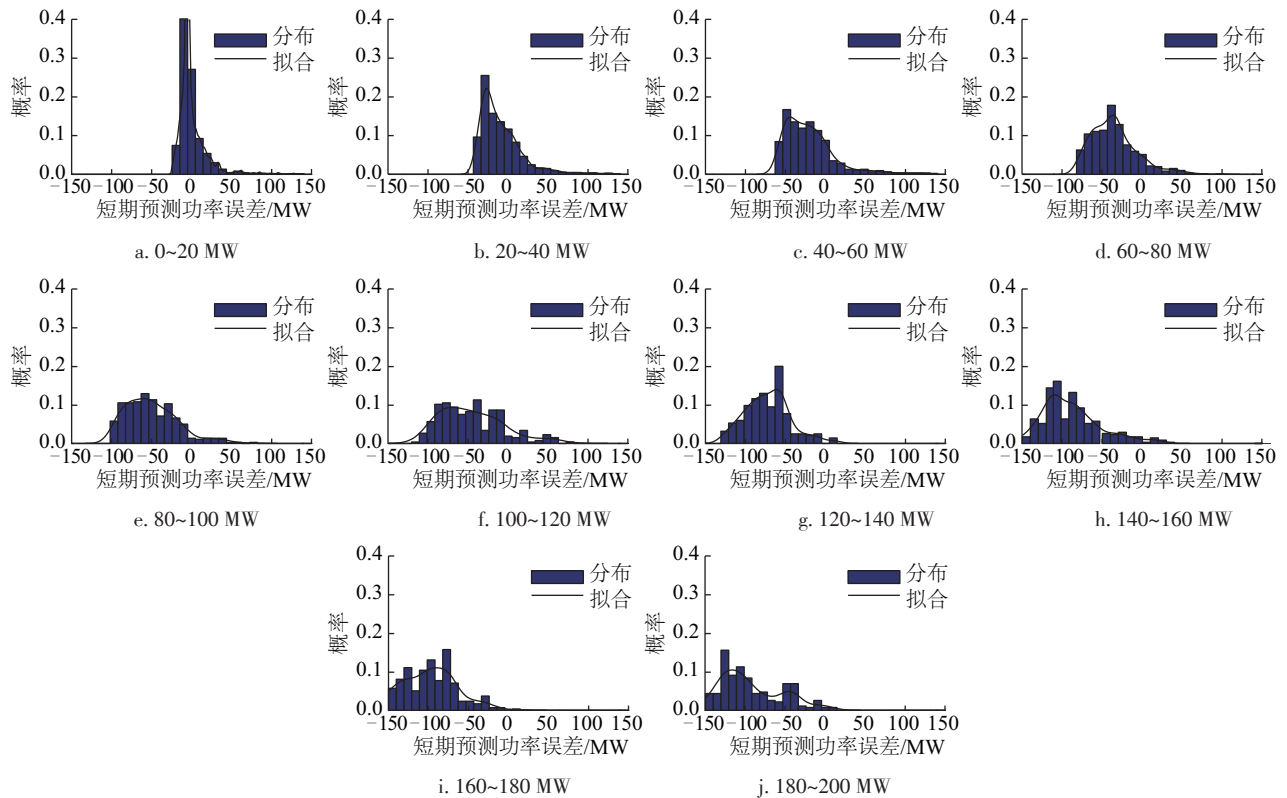


图 1 不同功率段下预测功率误差分布及概率密度函数拟合

Fig. 1 Prediction error distribution of wind power and probability density function estimation under different power segments

验分布必将带来较大误差,有必要使用非参数 Bootstrap 统计法,通过多次抽样获得 Bootstrap 样本来提高模型的参数估计精度,故本文提出基于非参数滑动分块百分位数 Bootstrap 法的风电功率区间预测方法。

1.2 风功率时间相依性

由大量相关文献可知风速序列在时间尺度下具有明显自相关特性,图 2 为某天 96 个风速点自相关序列图,采样间隔 15 min。由图 2 可得风速序列自相关性显著,而风速是衡量风力发电功率的重要因素,为了保留由此带来的风电功率时间序列内在

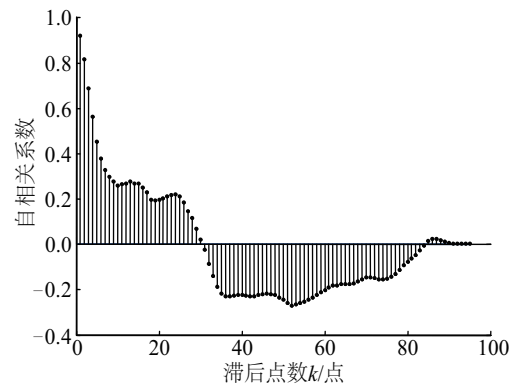


图 2 风速自相关序列图

Fig. 2 Autocorrelation sequence diagram of wind speed

的固有特性,本文采用滑动分块的 Bootstrap 法,在原有放回重复抽样的同时,利用“分块”的优势兼顾了风场数据的时间关联性。

2 滑动分块 Bootstrap 法相关理论

2.1 滑动分块 Bootstrap 法重采样技术

Bootstrap 方法是一种重采样统计方法,也称为自助法,其优势在于直接利用原始样本数据复制观测信息来进行统计推断,无需假设总体分布或做任何限制。若样本总体的分布函数形式 F 未知,可采用非参数 Bootstrap 法估计其分布的统计参数。非参数滑动分块 Bootstrap 法把 Bootstrap 法适用范围延伸到观测值有内部相依性的情况下^[13]。风场功率预测数据内部具有时间相依性,若对单个数据点进行重采样,就破坏了这种相依结构,所以可将同一“滑块”(Block)内的数据以单元形式进行抽取,采用滑动分块 Bootstrap 法可取得更接近于真实分布的预测效果。

滑动分块 Bootstrap 法有交叠滑块 Bootstrap 和不交叠滑块 Bootstrap 这 2 种形式。交叠滑块 Bootstrap 法在时间序列的分析与应用方面更精确有效^[14]。本文采用交叠滑块 Bootstrap 法建立伪样本集。具体操作步骤为:选择滑块长度为 l ,把 N 个样本点按顺序划分成 $N-l+1$ 个部分交叠的滑块,用 $X_{l,i}=(x_i, \dots, x_{i+l-1})$, $i=1, \dots, N-l+1$ 表示。然后随机有放回选取 r 个部分重叠的滑块 $X_{l,i}^*$, $i=1, \dots, r$,将这些滑块放置在一起构成长度为 N 的序列。若 N/l 不是整数,需要对一个或多个滑块缩短长度来达到样本长度为 N 的要求。整个过程重复 B 次形成 B 个伪样本数据集。最后将伪样本的统计量估计作为总体统计量估计。

2.2 滑块最优长度估计

滑动分块 Bootstrap 法中最优滑块长度的选取对相依时间序列的分析至关重要。滑块长度太大使得自助法变成简单的样本重置,统计量取值范围变小,估计误差增大;而滑块长度太小在保留原始样本相依性的同时不可避免要降低统计量的精确性。文献[14]给出了一种当样本数为 n 时滑块最优长度选取的算法。先设定滑块长度初始值为 b ,以置信度 $1-\alpha$ 下误差预测区间的上限估计值 $\hat{\varepsilon}_u$

为例:

1) 适当选取 m, κ 满足 $m < n, \kappa < b$,对 $n-m+1$ 个可交叠不同样的序列 $\{\varepsilon_j, \dots, \varepsilon_{j+m-1}\}$, $j=1, \dots, n-m+1$,分别计算 $\hat{\varepsilon}_{u,j}(m, \kappa)$,得到样本量为 m ,滑块长度为 κ 的 $\hat{\varepsilon}_u(m, \kappa)$ 的均方误差估计:

$$MSE(m, \kappa) = \frac{1}{n-m+1} \cdot \sum_{j=1}^{n-m+1} [\hat{\varepsilon}_{u,j}(m, \kappa) - \hat{\varepsilon}_u(n, b)]^2 \quad (1)$$

式中, $\hat{\varepsilon}_{u,j}(m, \kappa)$ ——第 j 个样本量为 m ,滑块长度为 κ 的误差预测区间上限估计值; $\hat{\varepsilon}_u(n, b)$ ——样本量为 n ,滑块长度为 b 的误差预测区间上限估计值。

2) m 值不变,取不同的 κ 值重复上述过程。找到使 $MSE(m, \kappa)$ 最小的 $\hat{\kappa}$,就可得到样本量为 n 的滑块最优长度。

$$\hat{b} = \hat{\kappa} \times (n/m)^{1/3} \quad (2)$$

同理可得到求下限估计值 $\hat{\varepsilon}_l$ 的 MBPB 模型的滑块最优长度。

3 基于滑动分块百分位数 Bootstrap 法风功率区间预测模型

3.1 风功率预测区间的划分

不同功率段的预测误差分布相差较大,为了准确推断误差分布规律需对功率区间进行划分,本文采取等间隔对功率区间划分。功率取值范围是 $[P_l, P_h]$,假设功率段长为 ΔP ,则所分区间为:

$$D_i = [P_l + (i-1)\Delta P, P_l + i\Delta P] \quad (3)$$

其中, $i=1, 2, \dots, N$ 。

N 是区段数:

$$N = (P_h - P_l) / \Delta P \quad (4)$$

功率区间划分越小,误差分布规律描述越详细,但在样本数据量一定下,划分区间段数量增多,可能导致某些区间样本数量较少,无法由此分析预测误差统计规律。应用非参数滑动分块 Bootstrap 方法,一般情况下,若采用大样本,重复抽样分布与总体分布只有微小差别,效果较好;而当采用小样本时,可能因为某些点未抽到而效果变差。因此对小样本的功率区间需进行二次划分,将区间内点数较少的相邻区间合并(一般为高风速段),以获得更高预测效果。

3.2 基于四分位法的滑动分块百分位数 Bootstrap 法置信区间求取

3.2.1 滑动分块 Bootstrap 单个样本百分位数法置信区间估计

设 $X=(X_1, X_2, \dots, X_n)$ 是来自总体 F 分布未知, 容量为 n 的样本。 $x=(x_1, x_2, \dots, x_n)$ 是一个已知样本值, F 中含有未知参数 θ , $\hat{\theta}=\hat{\theta}(X_1, X_2, \dots, X_n)$ 是 θ 的估计量。非参数滑动分块百分位数 Bootstrap 法求取 θ 的置信水平为 $1-\alpha$ 的置信区间估计过程如下:

采用 2.1 节中的滑动分块 Bootstrap 法, 依次独立地从观测值 $x=(x_1, x_2, \dots, x_n)$ 中抽出 B 个容量为 n 的 Bootstrap 伪样本, $x^*=(x_1^*, x_2^*, \dots, x_n^*)$, $i=1, \dots, B$ 。对于第 i 个 Bootstrap 样本, 求出 θ 的 Bootstrap 估计: $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_n^*$, 将它们按升序排列:

$$\hat{\theta}_{(1)}^* \leq \hat{\theta}_{(2)}^* \leq \dots \leq \hat{\theta}_{(n)}^* \quad (5)$$

式中, $\hat{\theta}_{(n)}^*$ ——升序排列后的第 n 个 $\hat{\theta}^*$ 值。

取 $R(x)=\hat{\theta}$, 用对应的 $R(x^*)=\hat{\theta}^*$ 的分布作为 $R(x)$ 的近似分布, 求出 $R(x^*)$ 的分布近似分位点 $\hat{\theta}_{(\alpha/2)}^*$ 和 $\hat{\theta}_{(1-\alpha/2)}^*$, 使:

$$P\{\hat{\theta}_{(\alpha/2)}^* < \hat{\theta} < \hat{\theta}_{(1-\alpha/2)}^*\} = 1 - \alpha \quad (6)$$

近似可得:

$$P\{\hat{\theta}_{(\alpha/2)}^* < \theta < \hat{\theta}_{(1-\alpha/2)}^*\} = 1 - \alpha \quad (7)$$

由式(7)即可得到 θ 在置信度为 $1-\alpha$ 下的近似置信区间:

$$(\hat{\theta}_{(\alpha/2)}^*, \hat{\theta}_{(1-\alpha/2)}^*) \quad (8)$$

此区间称为第 i 个 Bootstrap 样本在 $1-\alpha$ 置信度下 θ 的百分位数 Bootstrap 法置信区间。

3.2.2 基于四分位法的滑动分块 Bootstrap 样本置信区间估计

以功率区间 $[P_1, P_1 + \Delta P]$ 为例, 置信区间求取过程如下:

首先对该段功率区间的误差数据 ε 按照 2.2 节确定最优滑块长度, 进行滑动分块 Bootstrap 法重采样, 得到 B 个伪样本数据; 然后将预测误差 ε 对应于 3.2.1 节中的 $\hat{\theta}$, 按 3.2.1 节中方法求取置信水平 $1-\alpha$ 下每个伪样本数据的置信区间, B 个伪样本将得到 B 个置信区间。

此 B 个置信区间哪个更代表真实待求置信区

间是关键难点。本文提出基于四分位法的 B 个置信区间综合法来获得目标置信区间。其方法如下:

四分位法是统计学中分析数据集分布特征的重要方法, 原理如图 3 所示。是指将数据集中的全部数据按大小顺序依次排列并分为 4 等份, 处于分割点位置的 3 个数值 Q_1 、 M 、 Q_3 就是四分位数。

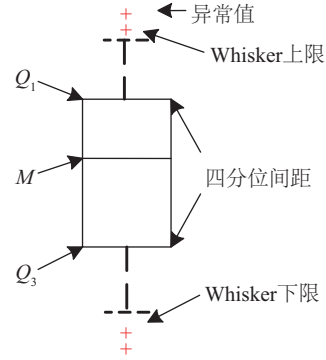


图3 四分位法示意图

Fig. 3 Schematic diagram of quartile method

将 B 个置信区间的下限按从小到大顺序排列, 图 3 中所示的特征数据求取方法参见文献[15], 这里不详述。将位于 Q_1 到 Q_3 之间的所有下限数据取平均作为待求置信区间的下限。同理将 B 个置信区间的上限按从小到大顺序排列, 遵循同样原理, 可得到待求置信区间的上限。上述方法求得的下限和上限即为功率区间 $[P_1, P_1 + \Delta P]$ 在 $1-\alpha$ 置信度下的置信区间。

3.3 风功率区间预测模型预测步骤

采用四分位非参数滑动分块百分位数 Bootstrap 法进行概率性区间预测的具体步骤如下:

1) 将历史数据进行必要的剔除坏点等预处理, 计算功率预测误差序列, 按照预测功率等间隔划分, 得到不同功率区间对应的预测误差序列;

2) 根据各个区间内的点数进行二次划分, 将点数较少的相邻区间合并;

3) 确定滑块最优长度, 采用 3.2.1 节非参数滑动分块百分位 Bootstrap 法计算每个功率段 B 个样本在某置信水平下误差的置信区间上下限;

4) 执行完 3) 后, 对应每个功率段, 会分别得到 B 组置信区间上限。将这 B 组置信区间从小到大排列, 采用四分位法处理, 选取位于 Q_1 到 Q_3 之间的数据求取平均值, 作为这一功率段置信区间的上

限,同理将 B 组置信区间下限,按上述方法处理,计算获得这一功率段置信区间的下限;

5)判断未来的确定性点预测值对应的功率区段,确定对应的置信区间上下限,转化为预测值的置信区间上下限;

6)将每个时刻的预测值的置信区间上限和下限分别联结即可形成上下 2 条包络线,得到概率性区间预测。图 4 为该方法的流程示意图。



图4 非参数滑动分块Bootstrap法风电功率
概率性区间预测流程图

Fig. 4 Flow chart of wind power probability interval
prediction for non-parametric moving block
Bootstrap method

4 算例分析

以西北某风场为例,该风场总装机容量为 199.5 MW,时间分辨率为 15 min,取风场 2014 年 9~11 月间从现场采集到的风电功率数据,包括实际功率数据和预测功率数据,仿真验证本文方法的可行性,并且与百分位数法、百分数 Bootstrap 法预测结果作对比,验证本方法的优越性。

4.1 预测误差区间划分

统计 2014 年 9~10 月份的预测误差数据,近 5500 个点,去除重复点及明显异常点(如功率为负值点),剩余有效点数为 4788 点。由于在不同功率段预测误差分布差异较大,首先按预测功率划分功率区段,按经验选取区间跨度为 20 MW,共计 10 个区段,统计各段内的预测误差点数,如表 1 所示。

表1 数据分段情况

Table 1 Data segmentation

功率/MW	点数	功率/MW	点数
0~20	2221	100~120	161
20~40	809	120~140	109
40~60	471	140~160	118
60~80	336	160~180	47
80~100	231	180~200	38

从表 1 中数据看出,高功率段数据点明显减少,此时采用滑动分块 Bootstrap 法对其重抽样将会有较大误差,得到的结果将不能反映误差分布的真实情况,所以将 160~200 MW 功率段合并,合并后的区段 160~200 MW 点数为 85。

4.2 滑块最优长度选取

以 85%的置信度下, 0~20 MW 功率段的预测误差上限估计值 $\hat{\varepsilon}_u$ 为例,详述最优长度的选取过程。

1) 0~20 MW 功率段内共有 2221 个误差点数,设滑块初始长度为 9,重采样次数 $B=5000$,计算得误差上限估计值 $\hat{\varepsilon}_u(2221,9)$;

2) 取 $m=2213$, $\kappa=2$,对 8 个可交叠不同样序列 $\{\varepsilon_j, \dots, \varepsilon_{j+2212}\}$, $j=1, \dots, 8$,分别计算 $\hat{\varepsilon}_u(m,2)$,得到样本量为 2213,滑块长度为 2 的 $\hat{\varepsilon}_u(2213,2)$ 的均方误差估计 $MSE(2213,2)$;

3) m 值不变,再分别求出 $\kappa=3, \dots, 8$ 时的均方误差 $MSE(2213, \kappa)$,由表 2 可知使 $MSE(m, \kappa)$ 最小的 κ 值为 3,故最优长度估计为 $l=3 \times (2221/2213)^{1/3} \approx 3$ 。

表2 不同 κ 值下的均方误差

Table 2 Mean square error under different κ values

κ	2	3	4	5
$MSE(m, \kappa)$	8.07	2.67	18.47	61.50
κ	6	7	8	—
$MSE(m, \kappa)$	89.09	93.71	98.62	—

求取其区间误差上下限的滑块自助法模型的最优长度均由此法得出。

4.3 仿真结果与分析

利用非参数 Bootstrap 法在样本容量扩充方面的优势,并综合考虑数据量的丰富度和运算效率等

多方面因素,采用滑动分块 Bootstrap 法对 9~10 月份 2 个月的误差样本数据进行统计分析获取各功率段下误差区间上下限。然后对 11 月份的 1~20 日数据进行概率性预测,获取满足误差分布规律的置信区间,将区间上下边界点连接,得到的 2 条包络线就是某置信水平下的概率性区间预测结果。图 5 为 80%、85%、90% 置信水平下的风电功率概率区间预测效果。

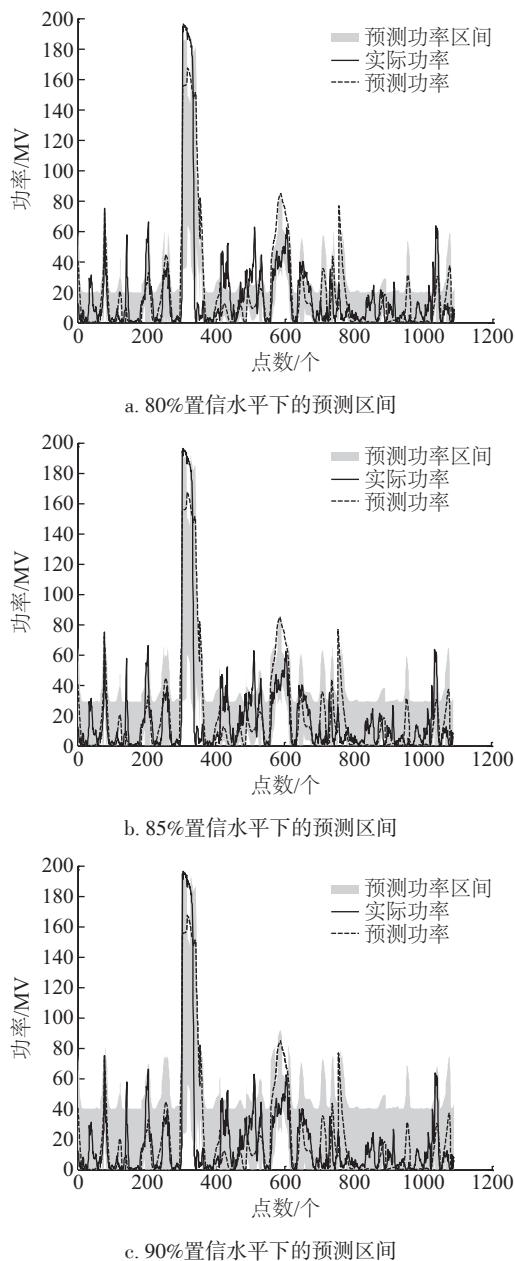


图5 3种不同置信水平下的预测区间

Fig. 5 Three prediction intervals under different confidence levels

从图 5a~图 5c 中可看出:1)利用非参数滑动分块 Bootstrap 法构建的风电功率概率预测区间能够反映对实际功率值的覆盖概率。与传统预测结果不同,本文所求出的置信区间包络线由预测误差的概率分布计算得到,并非关于预测值上下等间隔浮动;2)对于功率值相对较低及较高的区段,包络线间范围较大,中间功率段,置信区间宽度较窄。原因在于当功率值较小时,预测功率变化程度剧烈,预测误差较大;中间功率段的预测功率波动范围较小,预测较稳定;高功率段样本点数较少,预测功率波动剧烈,再加上区段合并造成误差较大;3)置信区间的宽度随着置信度的增加而增大,因为置信区间越大包含实际功率值的概率就越大,这与理论估计结果相吻合。

4.4 对比验证

选取同样的训练样本和测试样本,分别采用百分位数法和百分位 Bootstrap 法进行风电功率预测区间的估计。以区间覆盖率 (PI coverage probability, PICP) 和区间平均带宽 (PI normalized average width, PINAW) 为区间质量评估指标, P 代表百分位数法, PB 代表百分位数 Bootstrap 法, MBPB 代表本文提出的四分位滑动分块百分位数 Bootstrap 法。3 种算法在不同置信度下的 PICP、PINAW 指标值如表 3 所示。

表3 不同算法下性能指标比较

Table 3 Performance index comparison of different algorithms

置信度/%		PICP/%	PINAW
80	P	81.70	43.25
	PB	82.18	39.99
	MBPB	81.07	34.72
85	P	84.21	53.32
	PB	85.01	53.00
	MBPB	85.14	39.33
90	P	87.97	61.41
	PB	89.51	60.18
	MBPB	90.73	56.84

表 3 中:

$$PICP = \frac{1}{N_t} \sum_{i=1}^{N_t} k_i^{(\alpha)} \quad (9)$$

式中, N_t —— 预测样本数; $k_i^{(\alpha)}$ —— 布尔量, 若预测

目标值落入预测区间值为 1, 否则为 0。

$$PINAW = \frac{1}{N} \sum_{i=1}^N [U_i^{(\alpha)}(x_i) - L_i^{(\alpha)}(x_i)] \quad (10)$$

式中, $U_i^{(\alpha)}$ ——设定置信度下对应预测区间上限;
 $L_i^{(\alpha)}$ ——设定置信度下预测区间下限。

$PICP$ 一定时, $PINAW$ 越小表明区间范围越窄, 预测精度越高。

由表 3 可看出同一置信度下: 1) $PICP_{MBPB} > PICP_{PB} > PICP_P$, 尤其在置信度为 90% 时, 预测功率值以设定置信水平只落入了滑动分块 Bootstrap 法构造的预测区间内, 百分位法和百分数 Bootstrap 法失去预测可信度。说明对于区间准确度和有效性而言, Bootstrap 法构造的预测模型优于百分位法, 而基于四分位滑动分块的百分位 Bootstrap 法明显优于百分数 Bootstrap; 2) $PINAW_{MBPB} < PINAW_{PB} < PINAW_P$, 区间平均带宽越小意味着预测精度越高, 不确定性程度越小。综合对比之下, 滑动分块百分位 Bootstrap 法能得到更优良的预测效果。

5 结 论

本文提出的基于滑动分块 Bootstrap 的风电功率概率区间预测法有以下特点:

1) 对不同功率区间内采用非参数滑动分块 Bootstrap 法对样本进行误差统计量的重采样, 产生足够多的收敛于真实样本分布的伪样本序列, 解决了风电功率预测误差序列的分布模型未知给求取概率预测区间带来的难点。

2) 通过确定滑块最佳长度优化 Bootstrap 重采样样本, 并采用四分位法综合多个伪样本百分位数 Bootstrap 置信区间信息, 构建最终的功率区间预测模型, 从而得到区间覆盖率更高, 平均带宽更窄的风电功率预测区间。

3) 采用包含预测区间覆盖率和区间平均带宽的评价指标, 区间覆盖率量化了预测区间的准确度, 区间平均带宽可在保证区间准确度的基础之上, 对区间精度进行对比评估。最后通过与百分位 Bootstrap 法和常用的百分位数法对比彰显本文方法更优越的区间预测效果。

[参考文献]

[1] Miller N W, Guru D, Clark K. Wind generation [J]. IEEE Industry Applications Magazine, 2009, 15 (2):

54—61.

- [2] Khosravi A, Nahavandi S. Combined nonparametric prediction intervals for wind power generation [J]. IEEE Transactions on Sustainable Energy, 2013, 4 (4): 849—856.
- [3] Shi Jing, Guo Jinmei, Zheng Songtao. Evaluation of hybrid forecasting approaches for wind speed and power generation time series [J]. Renewable & Sustainable Energy Reviews, 2012, 16(5): 3471—3480.
- [4] Conejo A J, Plazas M A, Espinola R, et al. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models [J]. IEEE Transactions on Power Systems, 2005, 20(2): 1035—1042.
- [5] Huang G B, Wang D H, Lan Y. Extreme learning machines: A survey [J]. International Journal of Machine Learning & Cybernetics, 2011, 2 (2): 107—122.
- [6] 王晓曦. 基于非参数方法的短期风电功率预测研究 [D]. 大连: 大连理工大学, 2014.
- [6] Wang Xiaoxi. Short-term wind power prediction based on nonparametric method [D]. Dalian: Dalian University of Technology, 2014.
- [7] Yang Ming, Fan Shu, Lee Weijen. Short-term wind power forecast using componential sparse Bayesian learning [J]. IEEE Transactions on Industry Application, 2013, 49(6): 2783—2792.
- [8] Chen Niya, Qian Zheng, Nabney I T, et al. Wind power forecasts using Gaussian processes and numerical weather prediction [J]. IEEE Transactions on Power Systems, 2014, 29(2): 656—665.
- [9] 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): 656—665.
- [9] Liu Xingjie, Xie Chunyu. Prediction of wind power fluctuation intervals based on Beta distribution [J]. Electric Power Automation Equipment, 2014, 34(12): 656—665.
- [10] 李 智, 韩学山, 杨 明, 等. 基于分位点回归的风电功率波动区间分析 [J]. 电力系统自动化, 2011, 35 (3): 83—87.
- [10] Li Zhi, Han Xueshan, Yang Ming, et al. Analysis of wind power fluctuation intervals based on quantile regression [J]. Automation of Electric Power Systems, 2011, 35(3): 83—87.
- [11] 杨 明, 范 澍, 韩学山, 等. 基于分量稀疏贝叶斯学习的风电场输出功率概率预测方法 [J]. 电力系统自

- 动化, 2012, 36(14): 125—130.
- [11] Yang Ming, Fan Shu, Han Xueshan, et al. Probabilistic forecasting of wind power generation based on componential sparse Bayesian learning [J]. Automation of Electric Power Systems, 2012, 36(14): 125—130.
- [12] Bessa R J, Miranda V, Botterud A, et al. Time adaptive conditional kernel density estimation for wind power forecasting [J]. IEEE Transactions on Sustainable Energy, 2012, 3(4): 660—669.
- [13] Khosravi A, Nahavandi S, Creighton D. Prediction intervals for short-term wind farm power generation forecasts [J]. IEEE Transactions on Sustainable Energy, 2013, 4(3): 602—610.
- [14] 沈秀娟. 滑动区组 Bootstrap 方法的一致性及其最优步长估计的研究 [D]. 昆明: 云南大学, 2012.
- [14] Shen Xiujuan. The uniformity and optimum block length estimation of the moving block bootstrap method [D]. Kunming: Yunnan University, 2012.
- [15] 赵永宁, 叶林, 朱倩雯. 风电场弃风异常数据簇的特征及处理方法 [J]. 电力系统自动化, 2014, 38(21): 39—46.
- [15] Zhao Yongning, Ye Lin, Zhu Qianwen. Characteristics and processing method of abnormal data clusters caused by wind curtailments in wind farms [J]. Automation of Electric Power Systems, 2014, 38(21): 39—46.

PROBABILISTIC INTERVALS FORECASTING OF WIND POWER BASED ON MOVING BLOCK PERCENTILE BOOTSTRAP METHOD

Yang Xiyun¹, Zhang Huang¹, Guan Wenyuan², Dong Dehua¹

(1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China;

2. CSSC Electronics Technology Co., Ltd., Beijing 100070, China)

Abstract: A probabilistic wind generation forecasting intervals was established based on non-parametric Moving Block Percentile Bootstrap (MBPB) method. Due to the temporal dependence in power data, firstly, prediction power is divided equidistantly. Then moving block bootstrap method is applied to reshape new samples using forecasting errors of each power section. Combining MBPB and Quartile together, the lower upper bounds estimation under given confidence level can be obtained by utilizing the combination method for deduction. Finally, probabilistic forecasting interval of corresponding power section is achieved, and evidently superior to both Percentile and Percentile Bootstrap (PB) methods in terms of Prediction Interval Coverage Probability and PI Normalized Average Width which are introduced to assess the quality of interval. Results show that the proposed approach ensures a higher PICP and a narrower PINAW simultaneously, indicating a more satisfactory performance.

Keywords: wind power; prediction; probabilistic interval; moving block percentile Bootstrap method; quartile