

考虑时空不确定性的风电出力场景生成方法

孙 锴¹, 张大海¹, 李亚平², 严嘉豪²

(1. 北京交通大学 电气工程学院, 北京 100044; 2. 中国电力科学研究院有限公司(南京), 江苏 南京 210003)

摘要:为准确描述风电出力的不确定性及时空相关性,提出一种考虑时空不确定性的风电出力场景生成方法。将生成对抗网络作为风电出力的场景生成模型,将卷积神经网络作为模型生成器与判别器以实现时间特征的提取,采用特征工程方式实现不同风电场间出力空间相关性的量化;通过格拉姆角场方式进行特征变换,并合理设置网络结构及参数进行网络训练,得到生成器输入与输出场景间的映射关系。采用实测数据对所提方法的有效性进行对比验证,实验结果表明所提方法具有较强的风电出力不确定性表示能力。

关键词:场景生成;时空特性;特征工程;不确定性;生成对抗网络

中图分类号:TM73

文献标志码:A

DOI:10.16081/j.epae.202312016

0 引言

场景分析方法是一种描述可再生能源不确定性的有效方法,是通过对不确定性因素的量化将不确定性场景转换为确定性场景的建模方法^[1-3],在降低问题求解复杂度方面有着优异的性能,是求解电力系统优化运行问题、提升系统新能源消纳能力的有效途径^[4]。

传统场景分析方法如 Markov 链^[5]、场景树^[6]等方法是基于某种概率密度函数的统计假设,但确定性假设概率模型难以全面表征可再生能源出力的特征^[7]。随着人工智能算法^[8]的兴起与应用,人工神经网络算法^[9]被广泛应用于场景分析领域中。变分自编码器^[10](variational auto-encoder, VAE)与生成对抗网络^[11](generative adversarial network, GAN)是2种典型的生成式神经网络模型。VAE模型需要假设先验分布^[12],这导致其鲁棒性受到限制。GAN模型基于神经网络的“黑盒”特性对可再生能源出力的不确定性进行表征,能够通过构造不同形式的生成器与判别器网络完成场景生成任务^[13-14],由于避免了极大似然的计算过程,该模型不但能更全面地表征可再生能源出力的不确定性,而且极大简化了建模过程,降低了建模难度。

风资源的传播过程具有较强的惯性,不同风电场的出力在时间和空间维度上存在一定的相关性。时间相关性表示的是风电出力在时序上满足相近时刻出力的特定相关关系;而空间位置相近的风电场在相同或相近时刻的出力也表现出某种相关关系。这类时空上的相关关系是不断变化的,表现出一定

的不确定性。在时间方面,风速等自然条件随机性强,造成风电出力也包含较强的随机性;在空间方面,若简单假设各风电区域彼此之间相互独立,则生成的场景无法描述这类特性。因此,需要对时空不确定性进行综合考虑才能实现风电功率的准确模拟。现有的矩阵表示或相关性函数的风电出力时空相关性建模方法对多个风电出力数据间存在的线性或非线性相关性表示能力不足。

综合以上分析,本文考虑可再生能源出力的时空相关性,提出一种考虑风电出力时空关系的风电出力场景生成方法。采用基于卷积神经网络^[14-15](convolutional neural network, CNN)的GAN作为生成模型。在时间特征方面,分析可再生能源出力的时间相关性,将CNN作为判别器,并合理设置CNN卷积核尺寸,实现时间特征的有效处理;在空间特征方面,采用空间相关关系的特征工程方法对风电出力的空间相关性进行表示,并联合噪声数据作为生成器的输入特征,实现空间相关特征的提取。此外,采用格拉姆角场(Gramian angular field, GAF)方法进行特征变换,充分发挥CNN在特征图处理方面的优势。合理设置网络结构,在网络训练达到纳什均衡后,输出场景生成结果。将所提模型应用于实测数据,并将区间覆盖率(interval coverage probability, ICP)与区间平均宽度(interval average width, IAW)作为衡量标准,判别生成场景的有效性和准确性。

1 基于CNN的GAN

GAN是由Goodfellow I^[16]等提出的一种无监督学习模型,该模型包含生成器 G 和判别器 D , G 和 D 结构相互独立。在训练过程中,给定生成器 G 一个噪声向量 z (将其分布记为 $p_z(z)$),生成器 G 的任务是根据 z 得到输出,并使该输出尽可能逼近真实样本,判别器 D 负责判断样本是来自真实样本(真)还是来

收稿日期:2023-01-04;修回日期:2023-06-30

在线出版日期:2023-12-18

基金项目:国家重点研发计划项目(2022YFB2403400)

Project supported by the National Key R&D Program of China(2022YFB2403400)

自生成器 G (假)。

由于判别器 D 输出的是关于输入数据是否为真实样本的概率,可以通过最大化判别器期望 $E_{x \sim p_r(x)}[\ln D(x)]$ (x 为输入样本, $p_r(x)$ 为输入样本 x 的分布, $D(x)$ 为判别器 D 的输出),确保判别器 D 尽可能识别出真实数据。生成器 G 和判别器 D 在训练过程中不断交替进行参数调整,当网络达到纳什均衡后收敛,此时生成器 G 的输出能够使判别器 D 无法正确判断。GAN的训练过程可以由一个价值函数 $V(G, D)$ 进行评估, $V(G, D)$ 中包含生成器 G 和判别器 D 中需要调整的参数。GAN模型的极大极小训练目标为:

$$\min_G \max_D V(G, D) \quad (1)$$

GAN需要分别设置生成器 G 和判别器 D 的模型结构。以CNN、长短期记忆网络^[15]等为代表的深度学习算法在处理欧氏空间数据时具有良好的性能,有强大的非结构化数据处理能力。CNN能够通过设置不同尺寸的卷积核进行输入数据的滑动遍历处理,以特征值的方式实现输入数据的特征表示^[16]。因此,本文选择CNN作为生成器 G 和判别器 D 的网络结构,对风电出力数据的时空特性进行表示。融合CNN的GAN模型结构如图1所示。

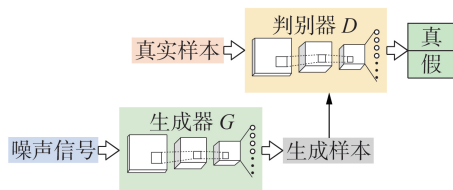


图1 融合CNN的GAN模型结构

Fig.1 Structure of GAN model integrated with CNN

CNN的计算方式及流程可参考文献^[17],此处不再赘述。为CNN设置不同卷积核数目及尺寸可实现输入数据的不同特征处理,本身特性决定了CNN具有平移、交替、缩放等“数据增强”功能,因此,可进一步提高模型的鲁棒性能。

2 时空特征分析

本文采用西班牙 Artelonía 地区 2020 年风电出力数据集(包含 4 座风电场的实测与日前预测数据)说明风电场间的空间相关性。4 座风电场的装机容量均为 150 MW,数据采样时间为 2020 年全年,采样间隔为 1 h,共包含 8 760 条风电及风速数据。4 座风电场编号为 a—d,其地理位置分布示意图如附录 A 图 A1 所示。

2.1 时间特征分析

现有文献大多利用时序卷积网络来提取可再生能源出力的长时序信息。事实上, t 时刻的风电出力

情况只与该时刻有较小时间间隔的几个时刻相关,而与较远时刻的相关性较小,符合“近大远小”的规律。

本文采用自相关系数^[18] $R(k)$ (k 为提前时间)表示风电出力的时间相关性, $R(k)$ 的计算公式为:

$$R(k) = \frac{\sum_{t=1}^{N-k} [(s_t - \mu)(s_{t+k} - \mu)]}{\sigma^2} \quad (2)$$

式中: N 为序列长度; s_t, s_{t+k} 分别为 t 时刻、 $t+k$ 时刻的均值为 μ 、方差为 σ 的生成场景或真实数据样本。 $R(k)$ 越接近1(-1),表示正(负)相关性越高,为0则表示不相关。

图2为不同 k 值下4座风电场出力实测值与预测值间的 $R(k)$ 计算结果。由图可知,随着时间间隔的增大,风电出力实测值与预测值间的相关性呈现出由大到小的变化趋势。因此,风电出力在时间层面表现出的相关性符合实际规律。

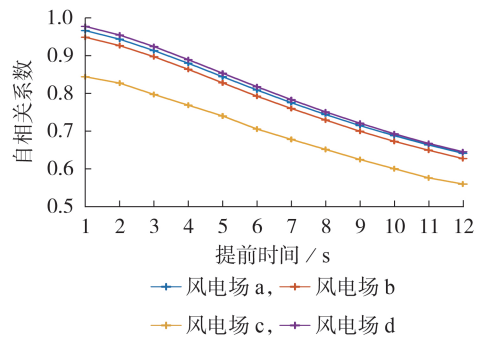


图2 自相关系数计算结果

Fig.2 Calculation result of autocorrelation coefficient

2.2 空间特征分析

空间相关性是指在一定空间范围内不同风电场之间表现出的相关性。风电出力受空间风资源的影响远超过其他因素,而风资源在一定空间范围内具有较高的相似性,这使风电出力也表现出一定的空间相关性。

以图A1中的风电场b为例,由于地理位置分布较接近,若在西北风向下,则风电场b与风电场a、d间均可能表现出一定的空间相关性,但与风电场c间的相关性可能相对较弱;而若在西南风向下,则风电场b与风电场c间的风电资源空间相关性相对较强。因此,除了考虑不同风电场间是否存在空间相关性外,还需要考虑不同时刻空间风资源造成的相关性强弱程度。

对于回归分析问题而言,若考虑不同对象间存在的相关性,则会提升模型精度,而现有文献大多是对整体风电出力数据进行空间关系分析,而没有考虑不同时刻的相关性强弱程度。此外,传统的GAN以噪声值为输入,在对数据信息的有效利用方面有

所欠缺。基于此,本文参考文献[19]的方法,利用特征工程的方式对不同风电场间的空间相关性进行表示,从而形成GAN的输入。

空间相关性提取步骤如下。

1)将不同的风电场出力数据视为不同的预测对象,并将出力数据按照时间顺序进行排列,对不同风电场间的风电出力数据进行相关性分析,即:

$$\mathbf{R} = \begin{bmatrix} W_{a,1} & W_{b,1} & W_{c,1} & W_{d,1} \\ W_{a,2} & W_{b,2} & W_{c,2} & W_{d,2} \\ \vdots & \vdots & \vdots & \vdots \\ W_{a,N} & W_{b,N} & W_{c,N} & W_{d,N} \end{bmatrix} \quad (3)$$

式中: \mathbf{R} 为由不同风电场出力数据组成的矩阵; $W_{i,n}$ ($i=a,b,c,d;n=1,2,\dots,N$)为风电场*i*在*n*点的出力数据。

2)由于不同风电场间表现出的相关性不仅包含线性关系,还包含非线性关系,因此,基于原始数据 \mathbf{R} ,以高次幂形式对*n*点数据进行扩展,联合不同风电场扩展数据组成扩展向量 $\mathbf{E}_{i,n}$,即:

$$\mathbf{E}_{i,n} = [W_{i,n}^1 \ W_{i,n}^2 \ \dots \ W_{i,n}^L] \quad (4)$$

式中: L 为数据幂次; $W_{i,n}^l$ ($l=1,2,\dots,L;i=a,b,c,d$)为风电场*i*在*n*点的扩展数据。

3)由前述分析可知,不同风电场之间可能会表现出不同的相关性特征,较低相关度的相关性特征会对模型性能产生不利影响,因此,需要对扩展数据进行相关度分析并剔除低相关度的数据。本文将所有向量 $\mathbf{E}_{i,n}$ 联合组成所有风电场在所有采样点下的扩展矩阵 \mathbf{E} ,并计算 \mathbf{E} 中列向量间的Pearson相关度^[20](Pearson系数的计算方式可参考文献[20]),然后删除相关度低于0.6的弱数据。

4)基于相关系数计算结果,将筛选后的矩阵 \mathbf{E} 按照式(5)、(6)形式组成数据耦合矩阵。

$$\mathbf{C}_{i,\text{in}} = \begin{bmatrix} c_{11}^{\text{in}} & \dots & c_{1L}^{\text{in}} \\ \vdots & & \vdots \\ c_{L1}^{\text{in}} & \dots & c_{LL}^{\text{in}} \end{bmatrix}_{L \times L} \quad (5)$$

$$\mathbf{C}_{i,\text{out}} = \begin{bmatrix} c_{11_1}^{\text{out}} & \dots & c_{1L_1}^{\text{out}} & \dots & c_{11_{j-1}}^{\text{out}} & \dots & c_{1L_{j-1}}^{\text{out}} \\ \vdots & & \vdots & & \vdots & & \vdots \\ c_{L1_1}^{\text{out}} & \dots & c_{LL_1}^{\text{out}} & \dots & c_{L1_{j-1}}^{\text{out}} & \dots & c_{LL_{j-1}}^{\text{out}} \end{bmatrix}_{L \times (j-1)L} \quad (6)$$

式中: $\mathbf{C}_{i,\text{in}}$ 、 $\mathbf{C}_{i,\text{out}}$ 分别为风电场*i*的内耦合和互耦合矩阵,表示同一风电场数据不同幂次间以及与不同风电场数据间的相关关系; c_{lg}^{in} ($l,g=1,2,\dots,L$)为同一风电场出力数据间不同次幂的相关系数计算结果; $c_{lg_s}^{\text{out}}$ ($l,g=1,2,\dots,L;s=1,2,\dots,j-1$)为与不同风电场出力数据间相关系数计算结果, $j=4$ 为数据阶数, $j-1$ 表示其余3座不同的风电场。

5)在得到耦合矩阵后,按照式(7)对其进行量化,得到相关关系的特征表示结果。

$$\mathbf{F}_{i,n} = \mathbf{E}_{i,n} \odot \mathbf{w} \mathbf{C}_{i,\text{in}}^T + \mathbf{E}_{r,n} \odot [\mathbf{w}, \mathbf{w}, \dots, \mathbf{w}]_{1 \times (j-1)L} \mathbf{C}_{i,\text{out}}^T \quad (7)$$

式中: $\mathbf{F}_{i,n}$ 为风电场*i*在*n*点量化后的耦合特征矩阵;“ \odot ”表示哈达玛积; $\mathbf{w}=[1/1! \ 1/2! \ \dots \ 1/L!]$ 为权重; $\mathbf{E}_{r,n}$ 为其余风电场在*n*点*L*次幂下的扩展数据矩阵。

通过式(3)—(7)计算得到风电场*i*的序列中所有点的量化矩阵组成的最终耦合特征矩阵后,将其作为输入特征,联合噪声数据作为GAN模型的输入。通过生成器与判别器的交替迭代训练可实现风电场时空相关特征的提取。

经以上特征工程,可将风电场间空间相关关系转变为确定性表示。如前文所述,由于空间地理位置不同,不同风电场出力间具有不同强弱程度的空间相关性,本文基于Pearson相关系数计算对应时序关系的内耦合及互耦合矩阵,通过相关系数的大小表示风电场出力间空间相关性的强弱程度。

在实际电网调度任务中,需要考虑未来24 h的新能源出力情况,须在日前得到新能源出力场景。经分析可知,与时间相关性类似,风电出力的预测值数据与实测值数据间也表现出相同的空间特性,因此,上述特征工程是基于风电出力的预测值数据进行分析的。

3 算例分析

采用西班牙 Artelonia 地区 2020 年全年 8760 点风电数据。实验平台硬件配置为 Intel(R) Core(TM) i5-8400 CPU 处理器,8 GB 内存;软件为 Windows 10 环境下安装的 Pytorch 库及 MATLAB 2020a,编程语言为 Python 及 m。

3.1 数据预处理

风电出力容易受到计划外停机及风速传感器故障或失灵等原因的影响,这使出力数据中包含类似“数据跳变”、空值数据、(近)零值等坏数据,因此,需先对数据集进行判别修正,避免坏数据对任务效果产生影响。本文采用文献[21]中的数据修正方式对原始数据进行清洗,再对数据进行归一化处理,从而消除数据量纲的影响,归一化计算公式如附录 A 式(A1)所示。

3.2 特征变换

为了充分发挥 CNN 在处理多维特征图方面的优势,通常采用特征工程中的特征变换方式对数据特征进行进一步处理。特征变换的目的是在不改变原始特征信息的基础上,进一步提取和扩展特征信息。常用的特征变换方式是主成分分析系列算法、小波系算法^[22]和 GAF 方法^[23]。其中,GAF 方法采用图谱式特征变换,能够在保留原始数据特性的基础上进一步挖掘和提取特征信息,该方法的本质是降低原始数据中某一范围内(如极值附近)的稀疏性,从而提高算法的数据处理能力,因此,本文采用 GAF 方法对输入数据特征进行变换。

GAF特征变换的过程如下。

1)数据标准化处理:采用3.1节方式对数据进行预处理,得到标准化数据。

2)极坐标转化:通过式(8)、(9)将归一化后的特征数据转换为极坐标形式。

$$\phi_n = \arccos \tilde{x}_n \quad (8)$$

$$r_n = t_n / N \quad (9)$$

式中: ϕ_n 为归一化后的序列值 \tilde{x}_n 映射到极坐标下的角度值; r_n 为极坐标下的半径; $t_n=1,2,\dots,N$ 为采样点 n 的时间戳,表示将区间 $[0,1]$ 进行 N 等份。

3)生成类格拉姆矩阵:标准化特征数据的格拉姆矩阵形式如式(10)所示。

$$\mathbf{G} = \begin{bmatrix} \cos \phi_{11} & \cos \phi_{12} & \cdots & \cos \phi_{1N} \\ \cos \phi_{21} & \cos \phi_{22} & \cdots & \cos \phi_{2N} \\ \vdots & \vdots & & \vdots \\ \cos \phi_{N1} & \cos \phi_{N2} & \cdots & \cos \phi_{NN} \end{bmatrix} \quad (10)$$

式中: \mathbf{G} 为格拉姆矩阵; $\phi_{ij}(i,j=1,2,\dots,N)$ 取值为 $\phi_i - \phi_j$ 或 $\phi_i + \phi_j$,分别表示差分GAF(difference GAF, GADF)及求和GAF(summation GAF, GASF)。GASF变换生成的是对称矩阵,更有利于网络学习,因此本文选择的特征变换方式为GASF。

需要指出的是,数据置换区间须在 $[0,1]$ 区间内,GAF变换的反变换才会对应唯一值。

3.3 模型参数设置

本文将CNN作为时间特征的提取单元,由于风电出力的时间相关性随着提前时间的增长而减小,因此CNN的卷积核尺寸不宜设置得过大。

此外,在GAN模型的训练过程中,选取随机梯度下降(stochastic gradient descent,SGD)作为优化器对初始学习率、批处理数量、最大训练次数等超参数进行优化,其余超参数通过迭代追踪的方式进行选取。综合考虑网络计算速度与结果精度,设置生成器 G 和判别器 D 网络及GAN模型训练参数,如附录A表A1所示。生成器与判别器网络参数设置分别如附录A表A2、A3所示。

3.4 评价指标

生成的场景集需要尽可能全面地包络次日预测/实际出力数据,才能有效表示出力不确定性,本文选择区间预测方法中常用的ICP和IAW^[24]来评价预测结果的准确性和有效性。此外,为检验生成样本与实际样本是否具有相同的时间统计特性和空间相关性,本文分别采用自相关系数和Pearson相关系数进行评价。

1)ICP。

ICP是指生成场景对次日出力数据的覆盖情况,其计算公式为:

$$I^{ICP} = \frac{\text{sum}(n^{\text{inner}})}{N} \times 100\% \quad (11)$$

式中: I^{ICP} 为ICP指标计算结果; $\text{sum}(\cdot)$ 为求和函数; n^{inner} 为实测出力数据落在区间内部的采样点数。 I^{ICP} 越高表示方法准确性越高。

2)IAW。

IAW表示功率区间的宽度均值,其计算公式为:

$$I^{IAW} = \frac{1}{N} \sum_{n=1}^N (p_n^{\text{max}} - p_n^{\text{min}}) \quad (12)$$

式中: I^{IAW} 为IAW指标计算结果; p_n^{max} 、 p_n^{min} 分别为采样点 n 的最大、最小出力值。 I^{IAW} 越小表示方法的有效性越高。

3)时空相关性评价。

在时间相关性方面,本文采用自相关系数 $R(k)$ 来评估生成场景对风电出力时间相关性的模拟效果。由于自相关系数会随着提前时间的增长而减小,因此,可以利用生成场景自相关系数的变化特征来衡量场景在时间方面的有效性。

在空间相关性方面,本文将Pearson相关系数作为评价标准。若生成场景之间的Pearson相关系数范围能够涵盖不同区域风电出力间的Pearson相关系数,则说明生成的场景能够有效描述风电出力间的空间相关性特征。

3.5 风电出力场景生成

在空间相关性表示部分,需要确定数据幂次 L , L 越大则方法的相关性表示能力越强。文献[19]指出, L 超过4以后对结果精确性的影响趋于恒定,因此,本文将 L 设置为4。网络参数设置完成后,采用风电场实测数据及预测数据对网络进行训练,网络训练流程如附录A图A2所示。在网络训练过程中,采用GPU并行计算加速,以提高网络训练效率和计算能力。

算例输入数据长度为24(代表一天中的24个时刻),以1h为步长循环截取数据,从而形成实验数据集,并分别按照95%与5%的比例将数据集划分为训练集与测试集。网络在CPU与GPU上的训练时长分别为244.92s和75.93s,可见GPU计算单元可以显著缩短网络的训练时长。训练过程中产生的生成器与判别器损失如附录B图B1所示。由图可知,生成器与判别器损失随着迭代次数的增加均趋于稳定,这说明网络参数设置合理,未出现过拟合等网络无法收敛的情况。

1)生成场景。

以12月31日风电出力数据为例,利用训练完成的生成器模型生成风电出力场景。经GAF反变换得到风电场a的100组场景生成结果,如图3所示。风电场b—d的场景生成结果如附录B图B2所示。结合图3与图B2可知,4座风电场的生成场景不存在明显的波动性,生成场景能够较好地包含实测值。

2)场景有效性。

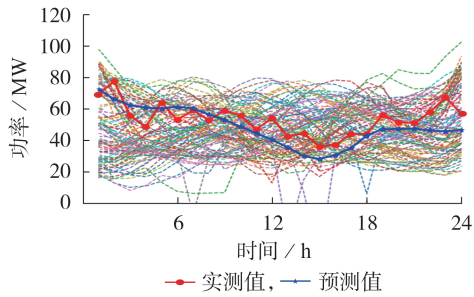


图3 风电场a的生成场景

Fig.3 Generation scenarios of Wind Farm a

ICP 与 IAW 评价指标的计算结果如表 1 所示。ICP 指标结果表明风电场 a、b、d 的生成场景实现了对实测值 100% 的覆盖,而风电场 c 对实测值也有超过 95% 的覆盖率。

表 1 生成场景的评价指标

Table 1 Evaluation indexes of generation scenarios

风电场	ICP / %	IAW
a	100	62.93
b	100	69.46
c	95.83	42.09
d	100	104.40

此外,本文方法得到的生成场景能够在较小的区间宽度上对风电出力进行较好地表征。

3) 场景空间相关性。

表示不同风电场间空间相关性的 Pearson 相关系数计算结果如图 4 所示。图中,“a-b”表示风电场 a 与风电场 b 之间,其他类似。

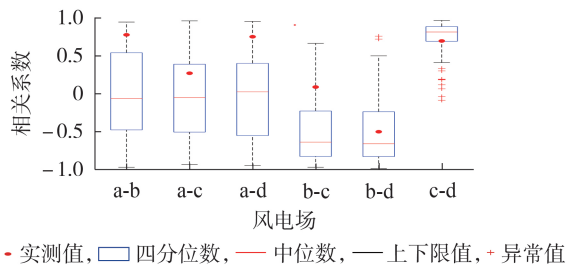


图 4 Pearson 相关系数计算结果

Fig.4 Calculation results of Pearson correlation coefficient

由图 4 可知,生成场景的 Pearson 相关系数能够完全包络各区域间实测值的 Pearson 相关系数,即场景生成方法能够有效刻画各区域风电出力间的空间相关关系。

4) 场景时间相关性。

在不同提前时间下风电场 a 的自相关系数计算结果如图 5 所示。风电场 b—d 的自相关系数计算结果如附录 B 图 B3 所示。

结合图 5 与图 B3 可知,各风电场生成场景的自相关系数分布均能够包络实测风电出力的自相关性

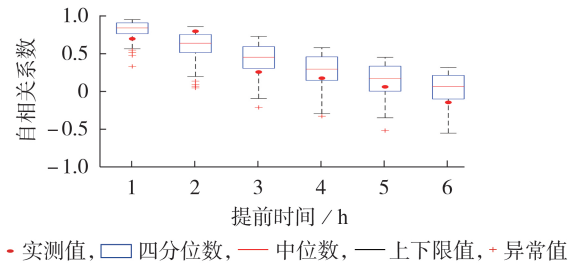


图 5 风电场 a 的自相关系数

Fig.5 Autocorrelation coefficient of Wind Farm a

系数,并且生成场景的自相关系数变化趋势与实测风电出力的自相关性的变化趋势基本一致。

5) 方法对比。

为了验证本文方法的优越性,采用以下 3 种方案进行对比分析:方案 1,利用本文方法生成场景集;方案 2,利用不考虑时空相关性特征的 GAN 模型生成场景集;方案 3,采用文献[25]方法的场景生成结果;方案 4,利用 VAE 算法生成场景集。其中:方案 1、2 可以用于验证考虑时空相关性的必要性;方案 1、3、4 可以用于验证不同方法的性能。

针对风电场 a 的数据,4 种方案的生成场景结果如附录 B 图 B4 所示。不同方案的评价指标计算结果如表 2 所示。

表 2 不同方案的评价指标

Table 2 Evaluation indexes of different schemes

方案	ICP / %	IAW
1	100	62.93
2	100	93.51
3	94.97	63.82
4	91.67	78.21

结合图 B4 与表 2 可知:在不考虑时空相关性的条件下,方案 2 的场景生成结果虽然可以实现风电场场景 100% 的覆盖,但 IAW 指标明显大于方案 1;方案 3 的 IAW 指标比方案 1 的大,且未能实现风电场场景的全部覆盖;方案 4 生成结果的整体波动性比其他方案的小,但方案 4 的误差指标表明,在相同的实验条件下,GAN 比 VAE 算法的效果更好;方案 1 能够在 100 组场景中实现 100% 覆盖的基础上得到较小的 IAW。

风电场 a 各方案场景生成结果的空间 Pearson 相关系数与自相关系数的计算结果分别如附录 B 图 B5、B6 所示。相较于方案 1,方案 2、3 所生成场景的自相关系数极值相对较多,方案 4 结果的极值较少,极值越少的场景生成结果的波动性越小,这与图 B3 的场景生成结果相吻合。

综合以上实验可知,本文方法生成的 100 组场景能够对风电出力时空相关性进行准确表征,从而在一定程度上避免了生成场景的冗余,有利于对考

虑风电不确定性的随机优化调度问题的研究。

4 结论

本文基于GAN生成模型,提出一种考虑时空相关性的风电出力场景生成方法。以特征工程方式联合CNN实现风电出力的时空相关性表征和提取,并建立场景生成模型。通过实验对比得到以下结论。

1)基于空间相关性的表征,联合噪声数据作为生成器模型的输入数据,在保证生成场景多样性的同时,通过CNN的卷积计算实现空间相关性的表征和提取。

2)采用GAF特征变换方式实现数据特征的进一步处理,降低了原始数据的稀疏性,更有利于深度学习模型的学习。

3)通过合理设置融合了CNN单元的GAN模型结构实现了不同时空相关性的有效提取,这使得网络更加灵活。此外,GAN生成模型属于黑盒模型,不涉及任何显式的特性分析,这进一步提高了模型的普适性。

4)在考虑数据间时空相关性的前提下,GAN模型能够更好地利用数据间的相关关系。多次实验对比结果表明,本文方法能够有效解决生成模型的过拟合问题。

风电场规模及间隔等其他因素也会对风电出力的相关性产生影响,但由于受到条件限制,对于该部分数据的获取有一定难度,因此本文中未对其进行考虑,未来可在本文方法的基础上进一步研究其他相关因素的影响。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 朱西平,江强,刘明航,等. 计及源-荷不确定性的综合能源系统两阶段优化调度[J]. 电力自动化设备,2023,43(8):9-16,32. ZHU Xiping, JIANG Qiang, LIU Minghang, et al. Two-stage optimal scheduling of integrated energy system considering source-load uncertainty[J]. Electric Power Automation Equipment, 2023, 43(8): 9-16, 32.
- [2] 陈志颖,温步瀛,朱振山. 计及风电相关性的区域综合能源系统多时间尺度优化调度[J]. 电力自动化设备,2023,43(8):25-32. CHEN Zhiying, WEN Buyang, ZHU Zhenshan. Multi-time scale optimal scheduling of regional integrated energy system considering wind power correlation[J]. Electric Power Automation Equipment, 2023, 43(8): 25-32.
- [3] 彭春华,于蓉,孙惠娟. 基于K-均值聚类多场景时序特性分析的分布式电源多目标规划[J]. 电力自动化设备,2015,35(10):58-65. PENG Chunhua, YU Rong, SUN Huijuan. Multi-objective DG planning based on K-means clustering and multi-scenario timing characteristics analysis[J]. Electric Power Automation Equipment, 2015, 35(10): 58-65.
- [4] 要金铭,赵书强,韦子瑜,等. 基于场景分析的电力系统日前调度及其快速求解方法[J]. 电力自动化设备,2022,42(9):102-110. YAO Jinming, ZHAO Shuqiang, WEI Ziyu, et al. Day-ahead dispatch and its fast solution method of power system based on scenario analysis[J]. Electric Power Automation Equipment, 2022, 42(9): 102-110.
- [5] 孟天骄. 基于马尔科夫链-多场景技术的主动配电网优化调度[D]. 北京:华北电力大学,2018. MENG Tianjiao. Uncertain scheduling based on the Markov chains-multiple scenarios technique in active distribution network[D]. Beijing: North China Electric Power University, 2018.
- [6] 胡锦涛. 基于场景分析的多随机变量规划方法的研究[D]. 沈阳:东北大学,2018. HU Jinxing. Studies on multi-stochastic variable programming methods based on scenario analysis[D]. Shenyang: Northeastern University, 2018.
- [7] 曾爽,丁屹峰,李香龙,等. 基于条件时序生成对抗网络的楼宇空调负荷场景生成方法[J]. 电网技术,2022,46(7):2448-2456. ZENG Shuang, DING Yifeng, LI Xianglong, et al. Scenario generation of air conditioning loads in buildings based on conditional TimeGAN[J]. Power System Technology, 2022, 46(7): 2448-2456.
- [8] 李博,孙建军,余攀,等. 基于负荷聚类与网络等效的配电网多维典型场景生成方法[J]. 中国电机工程学报,2021,41(8):2661-2671. LI Bo, SUN Jianjun, YU Pan, et al. A multi-dimensional typical scenarios generation algorithm for distribution network based on load clustering and network structure equivalence[J]. Proceedings of the CSEE, 2021, 41(8): 2661-2671.
- [9] 王阳,刘希喆. 基于GRU-MPC的光储充电站日前-日内两阶段优化控制[J]. 电力自动化设备,2022,42(10):177-183. WANG Yang, LIU Xizhe. Day-ahead and intra-day two-stage optimal control of photovoltaic-energy storage charging station based on GRU-MPC[J]. Electric Power Automation Equipment, 2022, 42(10): 177-183.
- [10] 王守相,陈海文,李小平,等. 风电和光伏随机场景生成的条件变分自动编码器方法[J]. 电网技术,2018,42(6):1860-1869. WANG Shouxiang, CHEN Haiwen, LI Xiaoping, et al. Conditional variational automatic encoder method for stochastic scenario generation of wind power and photovoltaic system[J]. Power System Technology, 2018, 42(6): 1860-1869.
- [11] 赵冬梅,谢家康,杜泽航,等. 基于统计信息系数和Wasserstein生成对抗网络的风火系统暂态特征选择与两阶段稳定评估[J]. 电力自动化设备,2023,43(4):106-113. ZHAO Dongmei, XIE Jiakang, DU Zehang, et al. Transient feature selection and two-stage stability assessment of wind-fire system based on uniform information coefficient and Wasserstein-generative adversarial network[J]. Electric Power Automation Equipment, 2023, 43(4): 106-113.
- [12] 王新迎,李焯,董骁狮,等. 基于变分自编码器的主动配电网多源-荷场景生成方法[J]. 电网技术,2021,45(8):2962-2969. WANG Xinying, LI Ye, DONG Xiaochong, et al. Multi-source-load scenario generation of active distribution network based on variational autoencoder[J]. Power System Technology, 2021, 45(8): 2962-2969.
- [13] 贾丽丽. 基于生成式对抗网络的特定场景生成技术及应用研究[D]. 郑州:郑州大学,2019. JIA Lili. Research and application on particular scene generation based on generative adversarial network[D]. Zhengzhou: Zhengzhou University, 2019.
- [14] 钱涛,任孟极,邵成成,等. 基于深度学习考虑出行模式的电动汽车充电负荷场景生成[J]. 电力系统自动化,2022,46(12):67-75. QIAN Tao, REN Mengji, SHAO Chengcheng, et al. Scenario generation of electric vehicle charging load based on deep learning considering travel mode[J]. Power System Automation, 2022, 46(12): 67-75.

- QIAN Tao, REN Mengji, SHAO Chengcheng, et al. Deep-learning-based electric vehicle charging load scenario generation considering travel mode[J]. Automation of Electric Power Systems, 2022, 46(12): 67-75.
- [15] 刘亚琿, 赵倩. 基于聚类经验模态分解的CNN-LSTM超短期电力负荷预测[J]. 电网技术, 2021, 45(11): 4444-4451.
- LIU Yahui, ZHAO Qian. Ultra-short-term power load forecasting based on cluster empirical mode decomposition of CNN-LSTM[J]. Power System Technology, 2021, 45(11): 4444-4451.
- [16] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of Advances in Neural Information Processing Systems. Montréal, Canada: MIT Press, 2014: 2672-2680.
- [17] 梁海峰, 袁芑, 高亚静. 基于CNN-Bi-LSTM网络的锂离子电池剩余使用寿命预测[J]. 电力自动化设备, 2021, 41(10): 213-219.
- LIANG Haifeng, YUAN Peng, GAO Yajing. Remaining useful life prediction of lithium-ion battery based on CNN-Bi-LSTM network[J]. Electric Power Automation Equipment, 2021, 41(10): 213-219.
- [18] 黎静华, 文劲宇, 程时杰, 等. 考虑多风电场出力Copula相关关系的场景生成方法[J]. 中国电机工程学报, 2013, 33(16): 30-36.
- LI Jinghua, WEN Jinyu, CHENG Shijie, et al. A scene generation method considering copula correlation relationship of multi-wind farms power[J]. Proceedings of the CSEE, 2013, 33(16): 30-36.
- [19] WANG C, SHE Z, CAO L. Coupled attribute analysis on numerical data[C]//International Joint Conference on Artificial Intelligence. Beijing, China: AAAI Press, 2013: 1736-1742.
- [20] 纪德洋, 金锋, 冬雷, 等. 基于皮尔逊相关系数的光伏电站数据修复[J]. 中国电机工程学报, 2022, 42(4): 1514-1523.
- Ji Deyang, JIN Feng, DONG Lei, et al. Data repairing of photovoltaic power plant based on Pearson correlation coefficient[J]. Proceedings of the CSEE, 2022, 42(4): 1514-1523.
- [21] 沈小军, 付雪姣, 周冲成, 等. 风电机组风速-功率异常运行数据特征及清洗方法[J]. 电工技术学报, 2018, 33(14): 3353-3361.
- SHEN Xiaojun, FU Xuejiao, ZHOU Chongcheng, et al. Characteristics of outliers in wind speed-power operation data of wind turbines and its cleaning method[J]. Transactions of China Electrotechnical Society, 2018, 33(14): 3353-3361.
- [22] 苏坤林. 储能技术在大规模新能源并网中的运用研究[D]. 北京: 华北电力大学, 2019.
- SU Kunlin. Research on application of energy storage technology in large-scale new energy integration[D]. Beijing: North China Electric Power University, 2019.
- [23] 贺才郡, 李开成, 杨王旺, 等. 基于双通道GAF和深度残差网络的电能质量复合扰动识别[J]. 电网技术, 2023, 47(1): 369-379.
- HE Caijun, LI Kaicheng, YANG Wangwang, et al. Power quality compound disturbance identification based on dual channel GAF and depth residual network[J]. Power System Technology, 2023, 47(1): 369-379.
- [24] 梅飞, 顾佳琪, 裴鑫, 等. 基于自适应滚动匹配预测修正模式的光伏区间预测[J]. 电力自动化设备, 2022, 42(2): 92-98.
- MEI Fei, GU Jiaqi, PEI Xin, et al. Photovoltaic interval prediction based on adaptive rolling matching prediction correction mode[J]. Electric Power Automation Equipment, 2022, 42(2): 92-98.
- [25] 董晓骞, 孙英云, 蒲天骄. 基于条件生成对抗网络的可再生能源日前场景生成方法[J]. 中国电机工程学报, 2020, 40(17): 5527-5536.
- DONG Xiaochong, SUN Yingyun, PU Tianjiao. Day-ahead scenario generation of renewable energy based on conditional GAN[J]. Proceedings of the CSEE, 2020, 40(17): 5527-5536.

作者简介:

孙 锴(1995—), 男, 博士研究生, 主要研究方向为人工智能方法在电网调度中的应用(**E-mail**: 21117014@bjtu.edu.cn);

张大海(1973—), 男, 副教授, 博士, 通信作者, 研究方向为电力系统监控、继电保护与电能质量(**E-mail**: dhzhang1@bjtu.edu.cn);

李亚平(1982—), 女, 高级工程师, 博士, 研究方向为电力系统优化与控制、人工智能技术应用(**E-mail**: liyaping@epri.sgcc.com.cn)。

(编辑 王锦秀)

Scenario generation method of wind power output considering spatiotemporal uncertainty

SUN Kai¹, ZHANG Dahai¹, LI Yaping², YAN Jiahao²

(1. School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. China Electric Power Research Institute(Nanjing), Nanjing 210003, China)

Abstract: In order to accurately describe the uncertainty and spatiotemporal correlation characteristic of wind power output, a scenario generation method of wind power output considering the spatiotemporal uncertainty is proposed. The generative adversarial network is taken as the scenario generation model of wind power output, the convolution neural network is taken as the model generator and discriminator to extract the time feature, and the quantization of spatiotemporal correlation characteristic between the outputs of different wind farms is realized by the mode of feature engineering. The feature transformation is carried out through the mode of Gramian angular field, the network structure and parameters are reasonably set for network training, and the mapping relationship between input and output scenarios of generator is obtained. The measured data is adopted for comparison and verification of the proposed method, and the experimental results show that the proposed method has strong ability to express the uncertainty of wind power output.

Key words: scenario generation; spatiotemporal characteristic; feature engineering; uncertainty; generative adversarial network

附录 A:

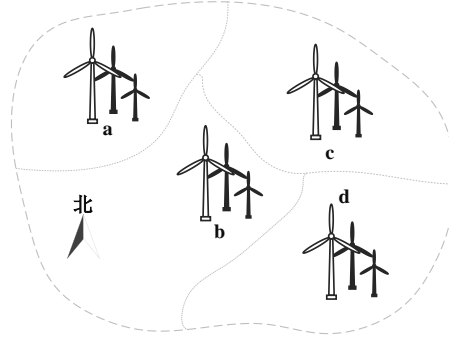


图 A1 风电场地理位置分布
Fig.A1 Geographical location distribution of wind farms

本文采用的归一化计算方式如下:

$$\tilde{x}_n = \frac{x_n - x_{\min}}{x_{\max} - x_{\min}} \quad (\text{A1})$$

式中: x_n 与 \tilde{x}_n 分别表示 t 时刻归一化前后的序列; x_{\min} 与 x_{\max} 分别为序列 x 中的最小值与最大值。

表 A1 模型结构及参数
Table A1 Model structure and parameters

参数项	参数设置
网络单元	CNN ReLU
激活函数	Leaky ReLU Tanh Sigmoid
优化器	SGD
损失函数	BCELoss
网络层数	5
初始学习率	0.01
学习率调整	0.1/20 次迭代
迭代次数	100
批处理	64

表 A2 生成器网络参数
Table A2 Parameters of generator

层	计算单元	输入/输出特征	kernel_size/stride/padding
1	ConvTranspose2d	5/64	5/1/2
	BatchNorm2d	64/64	—
	ReLU	—	—
2	ConvTranspose2d	64/32	5/1/2
	BatchNorm2d	32/32	—
	ReLU	—	—
3	ConvTranspose2d	32/16	5/1/2
	BatchNorm2d	16/16	—
	ReLU	—	—
4	ConvTranspose2d	16/8	5/1/2
	BatchNorm2d	8/8	—
	ReLU	—	—
5	ConvTranspose2d	8/1	5/1/2
	Tanh	—	—

表 A3 判别器网络参数
Table A3 Parameters of discriminator

层	计算单元	输入/输出特征	参数	kernel_size/stride/padding
1	Conv2d	1/8	—	5/1/2
	LeakyReLU	—	0.2	—
	Dropout	—	0.5	—
2	Conv2d	8/16	—	5/1/2
	BatchNorm2d	16/16	—	—
	LeakyReLU	—	0.2	—
3	Conv2d	16/32	—	5/1/2
	BatchNorm2d	32/32	—	—
	LeakyReLU	—	0.2	—
4	Dropout	—	0.5	—
	Conv2d	32/64	—	5/1/2
	BatchNorm2d	64/64	—	—
5	LeakyReLU	—	0.2	—
	Dropout	—	0.5	—
	Sigmoid	64/1	—	8/1/2

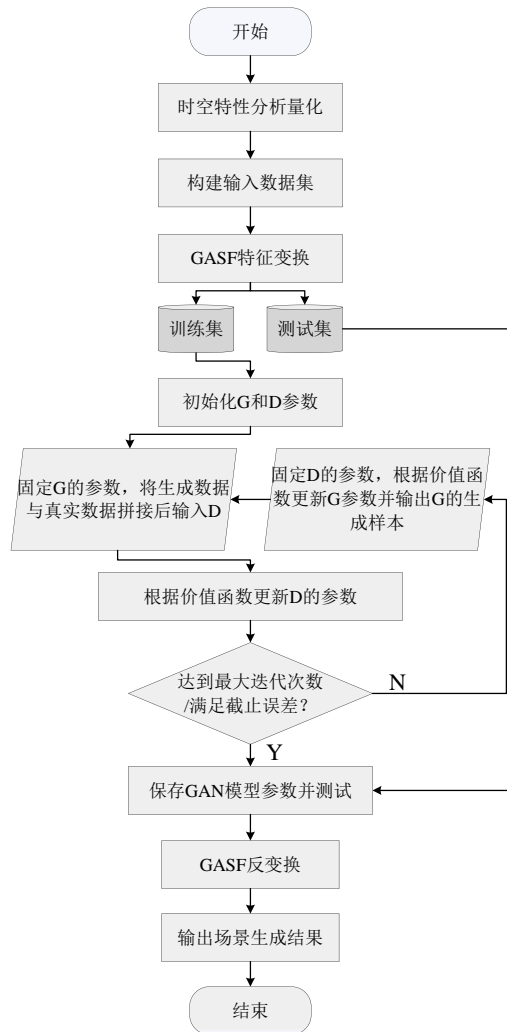


图 A2 场景生成流程

Fig.A2 Process of scenario generation

附录 B:

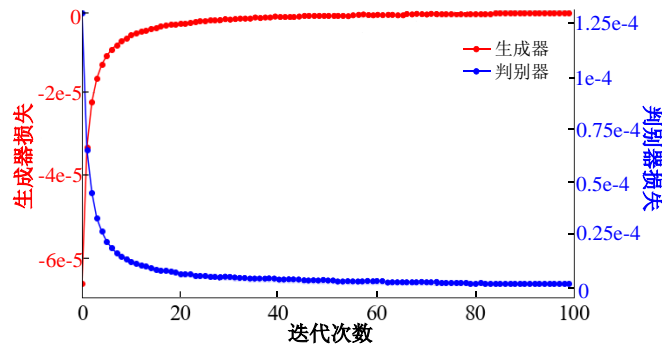
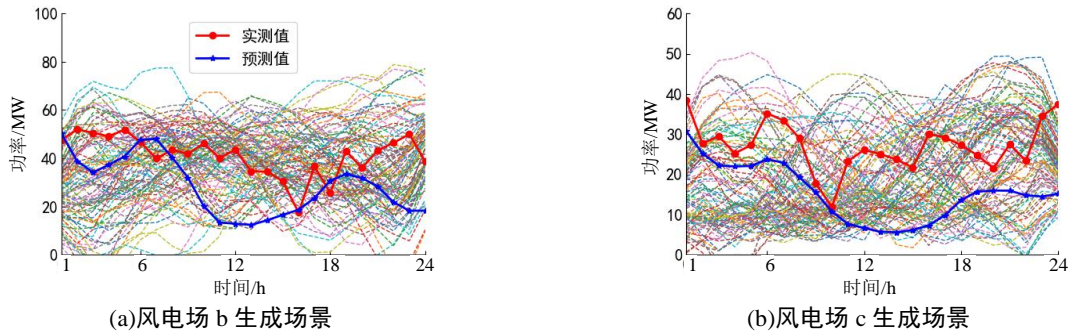


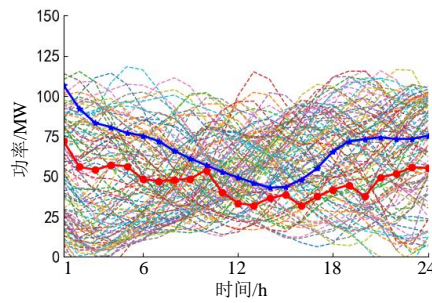
图 B1 生成器与判别器训练损失

Fig.B1 Training loss of generator and discriminator



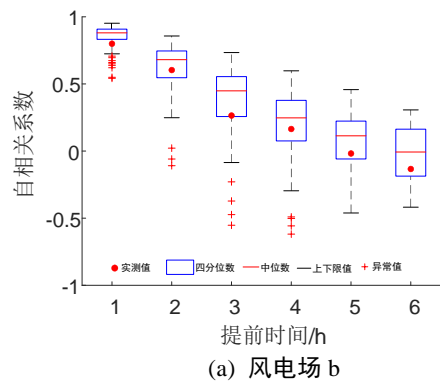
(a)风电场 b 生成场景

(b)风电场 c 生成场景

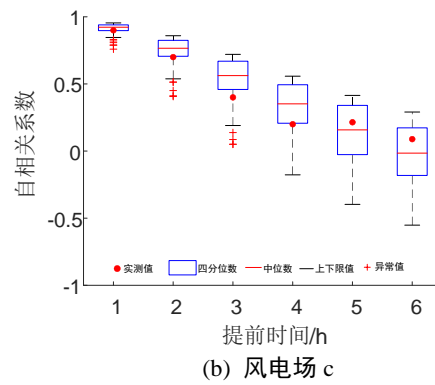


(c)风电场 d 生成场景

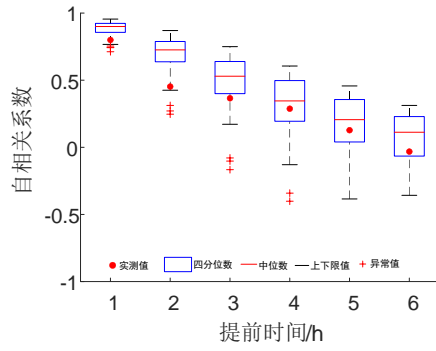
图 B2 3 座风电场出力生成场景
Fig.B2 Three wind farms output generation scenarios



(a) 风电场 b



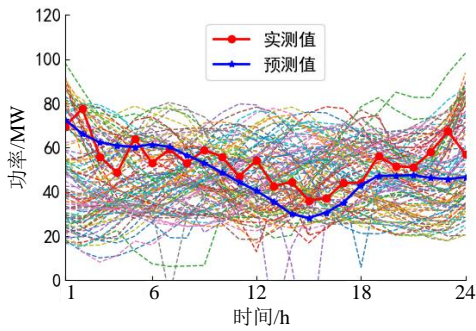
(b) 风电场 c



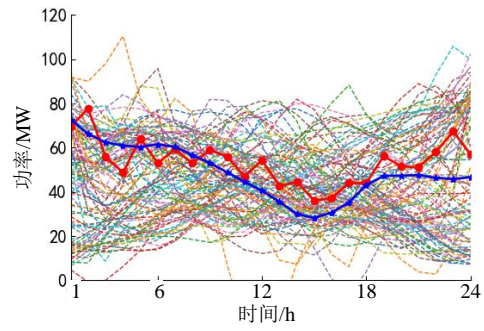
(c) 风电场 d

图 B3 3 座风电场自相关系数计算结果

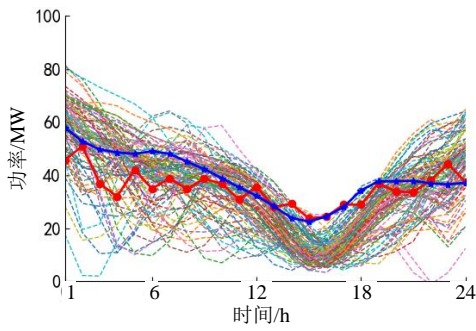
Fig.B3 Autocorrelation coefficient calculation result of three wind farms



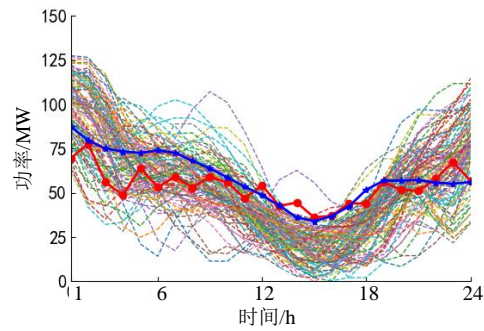
(a) 方案 1 生成结果



(b) 方案 2 生成结果



(c) 方案 3 生成结果



(d) 方案 4 生成结果

图 B4 不同方案的生成结果对比

Fig.B4 Comparison of generation results of different schemes

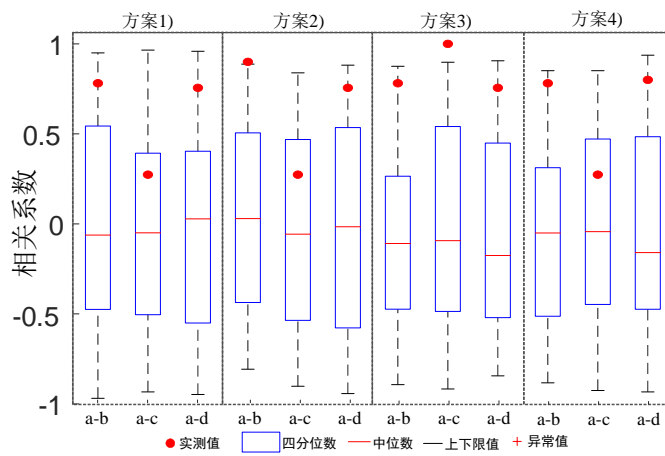


图 B5 不同方案 Pearson 系数计算结果

Fig.B5 Pearson correlation calculation result of different schemes

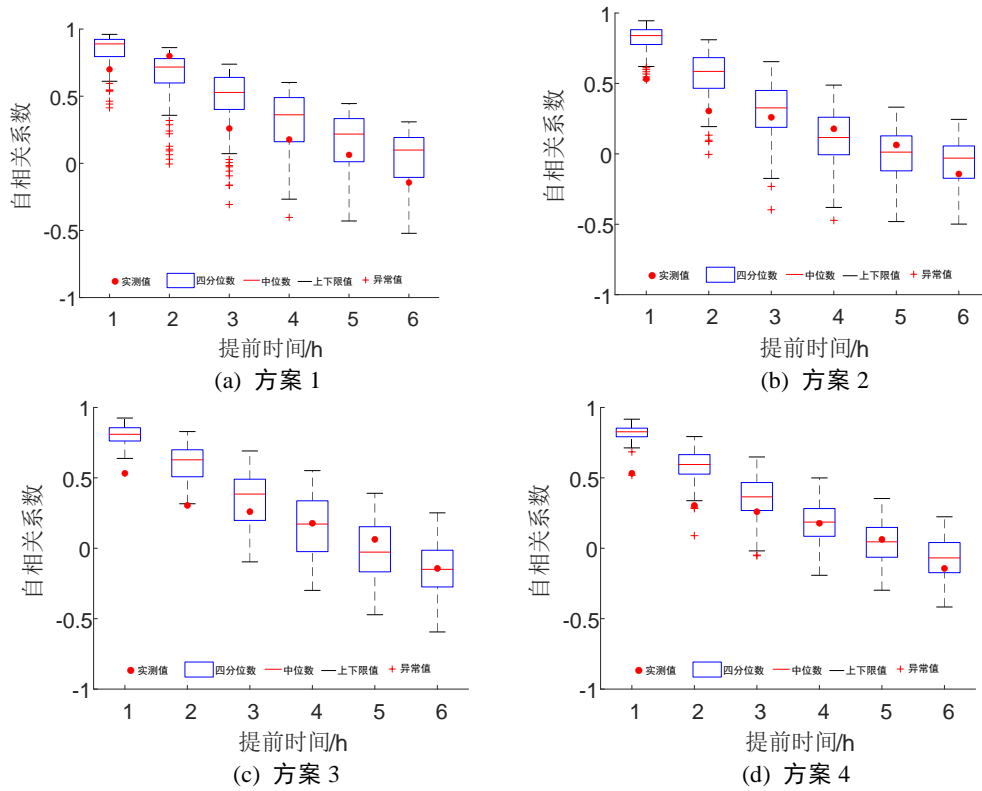


图 B6 不同方案自相关系数计算结果
 Fig.B6 Autocorrelation coefficient calculation result of different schemes