

DOI: 10.19783/j.cnki.pspc.231546

## 基于改进 MMD-GAN 的可再生能源随机场景生成

吴艳梅<sup>1,2</sup>, 陈红坤<sup>1,2</sup>, 陈磊<sup>1,2</sup>, 褚昱麟<sup>1,2</sup>, 高鹏<sup>1,2</sup>, 吴海涛<sup>3</sup>

(1. 交直流智能配电网湖北省工程中心(武汉大学), 湖北 武汉 430072; 2. 武汉大学电气与自动化学院, 湖北 武汉 430072; 3. 国网武汉供电公司, 湖北 武汉 430050)

**摘要:** 针对可再生能源出力不确定性的准确表征问题, 提出了一种基于改进的最大均值差异生成对抗网络(maximum mean discrepancy generative adversarial networks, MMD-GAN)的可再生能源随机场景生成方法。首先, 阐述了 GAN 及 MMD-GAN 的基本原理, 提出了 MMD-GAN 的改进方案, 即在 MMD-GAN 的基础上改进鉴别器损失函数, 并采用谱归一化和有界高斯核提升生成器和鉴别器的训练稳定性。然后, 设计了基于改进 MMD-GAN 的可再生能源随机场景生成流程。最后, 分析了所提方法在可再生能源随机场景生成中的效果, 比较了改进 MMD-GAN 方法与 MMD-GAN 方法及典型 GAN 方法的性能差异。结果表明, 改进 MMD-GAN 方法在生成分布和真实分布的 Wasserstein 距离上较对比方法降低超过 50%, 生成的场景精度得到有效提升。

**关键词:** 场景生成; 最大均值差异; 生成对抗网络; 可再生能源; 数据驱动

### Stochastic scenario generation for renewable energy based on improved MMD-GAN

WU Yanmei<sup>1,2</sup>, CHEN Hongkun<sup>1,2</sup>, CHEN Lei<sup>1,2</sup>, CHU Yulin<sup>1,2</sup>, GAO Peng<sup>1,2</sup>, WU Haitao<sup>3</sup>

(1. Hubei Engineering and Technology Research Center for AC/DC Intelligent Distribution Network, Wuhan University, Wuhan 430072, China; 2. School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China; 3. State Grid Wuhan Power Supply Company, Wuhan 430050, China)

**Abstract:** It is difficult to obtain an accurate characterization of the uncertainty of renewable energy output. Thus an approach for generating stochastic scenarios of renewable energy based on improved maximum mean discrepancy generative adversarial networks (MMD-GAN) is proposed. First, the fundamental principles of GAN and MMD-GAN are described, and an improved scheme of MMD-GAN is proposed, one which enhances the discriminator's loss function on the basis of MMD-GAN and uses spectral normalization and the bounded Gaussian kernel to improve the training stability of the generator and discriminator. Then, the process of stochastic scenario generation for renewable energy based on the improved MMD-GAN is designed. Finally, the effects of the proposed methods are analyzed. The performance of improved MMD-GAN, MMD-GAN, and typical GAN are compared. The results indicate that the improved MMD-GAN method can reduce the Wasserstein distance between the generated distribution and the real distribution by more than 50% contrasted with the comparison method, and the generated scenario accuracy can be effectively improved.

This work is supported by the Joint Fund of National Natural Science Foundation of China (No. U23B20117).

**Key words:** scenario generation; maximum mean discrepancy (MMD); generative adversarial networks (GAN); renewable energy; data-driven

## 0 引言

随着能源转型的推进, 风电、光伏等可再生能源在电力系统中的渗透率持续增长, 电力系统的安全稳定运行、调度规划及电能质量等方面正面临着

巨大挑战<sup>[1-2]</sup>。如何准确、可靠地对可再生能源输出特性的不确定性建模是应对上述挑战的关键所在。场景生成是广泛用于可再生能源不确定性建模的一种方法, 即在历史数据的基础上生成大量符合可再生能源出力特征的新时间序列表征其不确定性<sup>[3-4]</sup>, 从而在考虑可再生能源不确定性的前提下进行电力系统调度运行的最优决策, 例如随机经济调度<sup>[5-6]</sup>、

**基金项目:** 国家自然科学基金联合基金项目资助(U23B20117)

电力系统规划<sup>[7-8]</sup>、调峰调度<sup>[9]</sup>、机组组合<sup>[10]</sup>等。

已有研究提出了诸多基于模型的方法生成可再生能源随机场景。文献[11-12]将马尔可夫链蒙特卡洛法(Markov Chain Monto Carlo, MCMC)引入场景生成领域,发现其能较好地拟合可再生能源出力的概率密度函数和自相关系数。文献[13-14]使用自回归移动平均模型(autoregressive moving average model, ARMA)生成考虑空间相关性的多个风电场的出力场景集。文献[15-17]使用 copula 方法对可再生能源出力的相关参数化分布进行建模,通过拉丁超立方抽样(Latin hypercube sampling, LHS)或蒙特卡洛抽样(Monte Carlo sampling, MCS)得到场景。然而,上述基于模型的方法依赖于统计假设,难以对可再生能源输出的动态特性进行灵活建模;此外,不同地区可再生能源的出力分布存在差异,这限制了基于模型的方法进一步推广和应用。

随着人工智能技术的发展,以生成对抗网络(generative adversarial networks, GAN)为代表的深度学习模型为解决场景生成问题提供了新的思路。文献[18]应用 Wasserstein 生成对抗网络(Wasserstein generative adversarial networks, WGAN)学习可再生能源出力的时空相关性,生成具有统计相似性和模式多样性的风光随机场景;除 WGAN 外,GAN 的其他变体<sup>[19-22]</sup>也被引入到场景生成领域中。GAN 是一种数据驱动的方法,不需要对真实数据服从的概率分布进行任何假设,仅基于历史数据即可生成大量符合可再生能源出力特征的场景。然而,尽管 GAN 已成功应用于场景生成领域中,其模型的训练依然存在训练不稳定、梯度消失及收敛困难等问题,在以往研究中未能得到有效解决。

改进最大均值差异生成对抗网络(maximum mean discrepancy generative adversarial networks, MMD-GAN)采用最大均值差异度量生成样本和真实样本之间的距离,训练过程中梯度不会消失,并且可生成高质量的样本。目前,改进 MMD-GAN 在图像生成及修复领域取得了较好的应用效果<sup>[23]</sup>,但在可再生能源随机场景生成领域中的应用仍处于空白阶段,如何将其应用于可再生能源随机场景生成领域,从而解决可再生能源不确定性表征问题亟待进一步研究。

针对上述问题,本文提出了一种基于改进 MMD-GAN 的可再生能源随机场景生成方法。首先,阐述 GAN 及 MMD-GAN 的基本原理,为后续改进和应用提供理论依据;其次,提出 MMD-GAN 的改进方案,实现模型性能的进一步提升;然后,设计基于改进 MMD-GAN 的可再生能源随机场景

生成流程;最后,通过算例分析验证所提随机场景生成方法的有效性。

## 1 改进 MMD-GAN

### 1.1 生成对抗网络

生成对抗网络(generative adversarial networks, GAN)是由文献[24]提出的一种基于零和博弈的生成模型,该模型由生成器 G 和鉴别器 D 两个在结构上彼此独立的神经网络组成,其基本结构如图 1 所示。定义  $G(\cdot)$  和  $D(\cdot)$  分别为生成器和鉴别器映射函数。

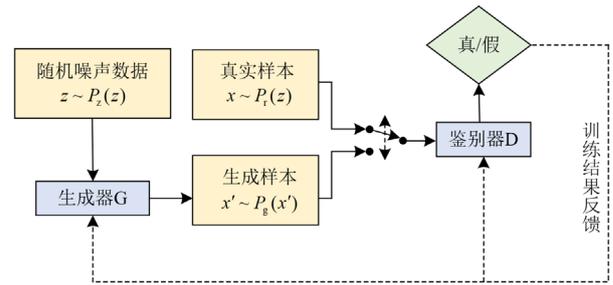


图 1 GAN 基本结构

Fig. 1 Basic structure of GAN

生成器的作用是学习真实样本  $x$  的分布规律并按此规律生成样本,其输入为服从先验分布  $P_z(z)$  的随机噪声数据  $z$ , 输出为生成样本  $x'$ , 其中  $x' = G(z)$ , 且  $x'$  服从生成分布  $P_g(x')$ 。鉴别器的作用是正确区分其输入来源,其输入为  $x'$  或  $x$ , 其中  $x$  服从真实分布  $P_r(x)$ , 输出为一个标量  $D(x')$  或  $D(x)$ , 该标量表示其输入服从  $P_r(x)$  的概率, 该值越大表明其输入为真实样本的可能性越高。不同的生成样本和真实样本分别构成了生成样本集  $X$  和真实样本集  $X'$ 。

在训练过程中,生成器和鉴别器依次根据鉴别器对其输入的判别结果更新自身参数,不断迭代,生成器生成样本的真实性和鉴别器的判别能力均不断提高。最终,当鉴别器无法区分出其输入来源时,对所有的输入都判断输出为 0.5,网络达到纳什均衡状态,即可认为  $x$  和  $x'$  服从同一概率分布。

根据生成器和鉴别器作用可知,对给定鉴别器,生成器要使  $D(x')$  尽可能接近 1;对给定生成器,鉴别器要在最大化  $D(x)$  的同时最小化  $D(x')$ 。因此,生成器和鉴别器的损失函数可分别定义为

$$L_G = E_{x' \sim P_g(x')} [\log(1 - D(x'))] \quad (1)$$

$$L_D = -E_{x \sim P_r(x)} [\log(D(x))] - E_{x' \sim P_g(x')} [\log(1 - D(x'))] \quad (2)$$

式中:  $L_G$  和  $L_D$  分别为生成器和鉴别器的损失函数,

$E$  为样本期望。

综合上述分析, GAN 的基本原理是一个极大极小优化问题, 其目标函数见附录 A 式(A1)。

### 1.2 MMD-GAN

传统 GAN 的最终目标是极小化  $P_g(x')$  和  $P_r(x)$  之间的 JS 距离<sup>[25]</sup>。但当  $P_g(x')$  和  $P_r(x)$  无重叠部分或重叠部分可忽略时, 两个分布之间的 JS 距离恒等于常数<sup>[26]</sup>, 此时梯度消失导致无法继续进行网络训练。

针对上述问题, MMD-GAN 将最大均值差异 MMD 引入 GAN 的目标函数中, 通过 MMD 损失的对抗性训练提升 GAN 的训练性能<sup>[27]</sup>。两个概率分布  $P$  和  $Q$  之间 MMD 距离的平方  $M_k^2(P, Q)$  的表达式见附录 A 式(A2)。

文献[28]证明了以下结论: 对于特征核变换  $k$ ,  $M_k^2(P, Q) \geq 0$  当且仅当  $P=Q$  时等号成立。因此两个分布越相似, 两者间的 MMD 距离越小。MMD-GAN 通过极小化  $P_r(x)$  和  $P_g(x')$  的 MMD 距离生成样本, 其将生成器和鉴别器的损失函数分别定义为

$$L_G = M_{k_D}^2(P_r(x), P_g(x')) = \frac{1}{m(m-1)} \sum_{i=1}^m \sum_{j \neq i}^m k_D(x_i, x_j) + \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i}^n k_D(x'_i, x'_j) - \frac{2}{mn} \sum_{i=1}^m \sum_{j=1}^n k_D(x_i, x'_j) \quad (3)$$

$$L_D = -M_{k_D}^2(P_r(x), P_g(x')) = -\frac{1}{m(m-1)} \sum_{i=1}^m \sum_{j \neq i}^m k_D(x_i, x_j) - \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i}^n k_D(x'_i, x'_j) + \frac{2}{mn} \sum_{i=1}^m \sum_{j=1}^n k_D(x_i, x'_j) \quad (4)$$

式中:  $k_D(\cdot)$  为核函数;  $x_i$  和  $x_j$  分别为  $X$  中第  $i$  个和第  $j$  个真实样本, 如附录 A 式(A3)所示;  $x'_i$  和  $x'_j$  分别为  $X'$  中第  $i$  个和第  $j$  个生成样本;  $m$  和  $n$  分别为  $X$  和  $X'$  的样本个数。

### 1.3 改进 MMD-GAN

根据式(4)可知, MMD-GAN 中, 鉴别器更加关注各个真实样本之间的相似性, 因此现有的 MMD 损失可能会忽略对真实样本中细节的学习<sup>[28]</sup>。

针对上述问题, 在 MMD-GAN 的基础上, 对鉴别器的损失函数进行改进<sup>[23]</sup>, 改进后鉴别器的损失函数为

$$L_D = \frac{1}{m(m-1)} \sum_{i=1}^m \sum_{j \neq i}^m k_D(x_i, x_j) - \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i}^n k_D(x'_i, x'_j) \quad (5)$$

生成器的损失函数仍采用式(3), 经改进后, 通

过极小化  $L_D$  的第一项, 即增大各个真实样本在鉴别器中的输出  $D(x)$  的差异, 鉴别器达到学习真实样本细节的目的, 为生成器提供更有意义的梯度, 因此整个网络对各个真实样本的差异区分更加明确。同时改进后鉴别器的损失函数较改进前简单, 训练时整个网络的计算量减少, 计算效率有所提高。

在训练过程中, 对鉴别器应用谱归一化以保证网络梯度满足  $K$ -Lipschitz 连续条件。权重矩阵的谱范数计算过程如图 2 所示, 其中: “\*” 表示卷积运算;  $W_c$  表示权重矩阵;  $v$  表示与  $W_c$  进行卷积运算的向量;  $u$  表示归一化  $W_c$  和  $v$  的卷积运算结果后的向量;  $\sigma(W_c)$  表示  $W_c$  的谱范数;  $c$  表示计算次数;  $N_{\max}$  表示最大计算次数。

同时对鉴别器  $D$  采用有界高斯核提升网络的训练稳定性, 有界高斯核表达式为

$$\begin{cases} k(D(x'_i), D(x'_j)) = \exp\left(-\frac{1}{2\zeta^2} \max\left(\|D(x'_i) - D(x'_j)\|^2, b_1\right)\right) \\ k(D(x_i), D(x_j)) = \exp\left(-\frac{1}{2\zeta^2} \min\left(\|D(x_i) - D(x_j)\|^2, b_u\right)\right) \end{cases} \quad (6)$$

式中:  $b_u$  和  $b_1$  分别为上下界;  $\zeta$  为窗宽。

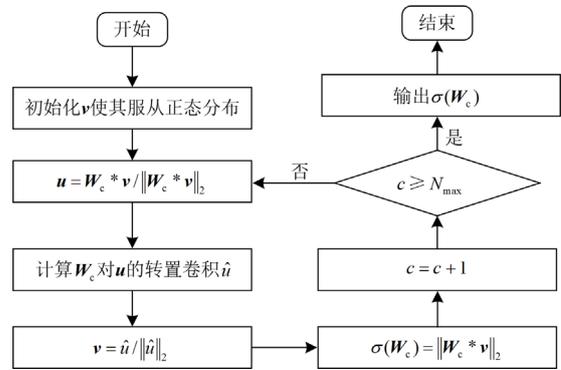


图 2 权重矩阵的谱范数计算过程

Fig. 2 Spectral norm calculation process of weight matrix

## 2 基于改进 MMD-GAN 的随机场景生成

### 2.1 基本思路

随机场景生成通过在历史数据的基础上生成大量符合可再生能源出力特征的新时间序列来表征其不确定性, 已有研究将 GAN 应用于可再生能源随机场景生成领域, 但未能有效解决 GAN 存在的梯度消失、训练失稳等问题。为生成更高质量的可再生能源随机场景并解决 GAN 训练失稳、梯度消

失等问题, 本文将改进 MMD-GAN 方法引入可再生能源随机场景生成领域, 并采用谱归一化及有界高斯核提升网络训练稳定性。

在改进 MMD-GAN 方法应用于随机场景生成任务中, 首先, 设计基于改进 MMD-GAN 的可再生能源随机场景生成流程, 接着, 基于可再生能源的出力特征设计适用于生成高质量随机场景的改进 MMD-GAN 网络结构, 最后, 进行生成器 G 和鉴别器 D 的对抗训练, 使生成器 G 可学习到可再生能源出力的分布规律, 在训练达到纳什平衡状态后, 生成器 G 可生成和历史数据服从同一概率分布的可再生能源随机场景。需要说明的是, 在该任务中,  $x$  对应为可再生能源的历史功率场景,  $x'$  对应为生成的可再生能源随机场景,  $P_r(x)$  和  $P_g(x')$  分别对应为可再生能源出力服从的真实分布和生成分布。

### 2.2 训练流程

基于改进 MMD-GAN 的可再生能源随机场景生成流程如图 3 所示, 训练步骤如下所述。

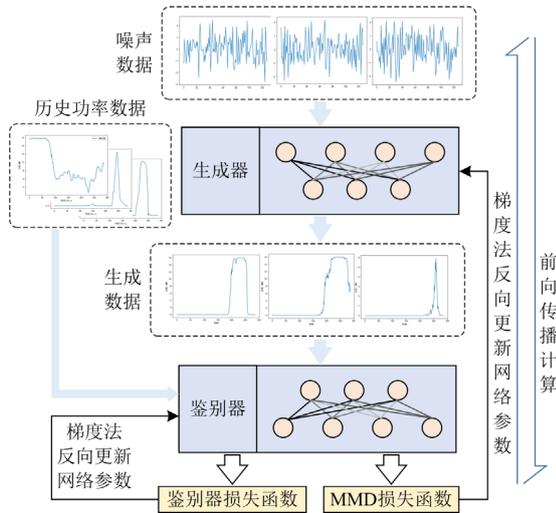


图 3 随机场景生成流程

Fig. 3 Process of generating stochastic scenarios

步骤 1 根据可再生能源的出力特征及时间分辨率, 设计适用于可再生能源随机场景生成的网络结构和参数, 并进行初始化。

步骤 2 将可再生能源的出力数据按式(7)进行归一化处理, 并将其重塑为与鉴别器输入维度大小一致的矩阵。

$$x'' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (7)$$

式中:  $x''$  为归一化后的真实样本;  $x_{\max}$ 、 $x_{\min}$  分别为可再生能源出力的最大值和最小值。

步骤 3 分别对  $P_r(x)$  和  $P_z(z)$  抽样得到  $x$  和  $z$ ,

将  $z$  作为生成器的输入进行前向传播得到  $x'$ , 并获得  $x$  和  $x'$  在鉴别器中的输出值。

步骤 4 按式(5)计算鉴别器的损失函数, 采用 Adam 优化器更新鉴别器的网络参数。

步骤 5 重复步骤 3, 根据步骤 3 的结果, 按式(3)计算生成器的损失函数, 采用 Adam 优化器更新生成器的网络参数。判断是否达到设定训练次数, 若未达到设定训练次数, 则返回步骤 3, 开始下一轮训练。

步骤 6 达到设定训练次数时结束训练, 提取改进 MMD-GAN 中的生成器网络。将服从先验分布的噪声数据作为输入, 生成器可生成与可再生能源真实场景服从相同规律的随机场景, 进行反归一化即可构成出力场景集。

## 3 算例分析

### 3.1 算例设置

本文采用美国可再生能源实验室的数据集进行验证分析, 考虑到生成风光随机场景的流程一致, 且风电的波动性更强, 生成风电随机场景的代表性更强<sup>[29]</sup>, 因此算例主要对风电随机场景生成结果进行分析, 光伏随机场景生成结果在附录 A 中给出。

以美国华盛顿州某地理位置分别相近的 52 个风电场和 32 个光伏场站 1 年的历史数据作为数据集<sup>[18]</sup>, 原始数据的时间分辨率为 5 min, 即每个出力样本的采样点数为 288 个。随机选择样本总数的 80% 作为训练数据, 用于训练改进 MMD-GAN 模型的网络权重参数, 其余 20% 作为测试数据, 用于验证训练所得模型的有效性和准确性。

算例中改进 MMD-GAN 模型通过深度学习框架 TensorFlow 搭建, 首先根据风电和光伏出力特征分别设计适用于生成高质量随机场景的网络结构并进行训练, 训练完成后从生成的随机场景、概率分布及相关性三方面验证改进 MMD-GAN 方法的有效性和准确性, 并将改进 MMD-GAN 与 MMD-GAN 及 WGAN 进行仿真实验对比。

### 3.2 网络结构

可再生能源的出力存在时间相关性, 为使生成器可充分捕捉到历史数据的分布规律, 本文选择卷积神经网络作为生成器和鉴别器的主要结构。

针对风电和光伏的出力特征分别设计适用于风电和光伏随机场景生成的网络结构, 经过反复验证, 找到其最佳的超参数组合。针对风电和光伏随机场景生成, 分别设计如图 4 及附录 A 图 A1 所示的网络结构和参数。其中,  $d$  表示全连接层的神经元个数,  $K$  表示卷积核的尺寸,  $N$  表示卷积核的个数,

$s$  表示卷积核滑动步长。例如  $K=5, s=2, N=64$  表示该卷积层或转置卷积层含有 64 个尺寸为  $5 \times 5$  的卷积核, 且其滑动步长为 2。

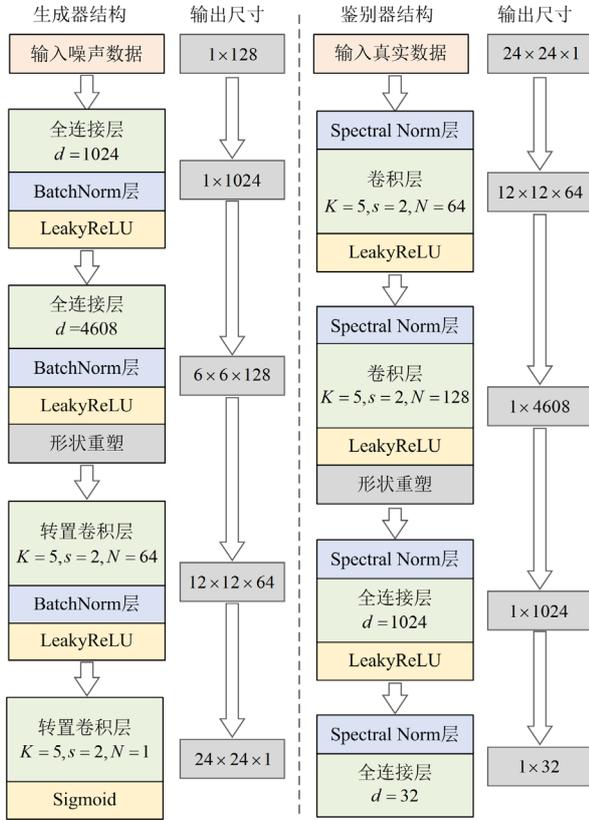


图 4 生成风电随机场景的网络结构和参数

Fig. 4 Network architecture and parameters for generating wind power stochastic scenarios

在生成器的输入层和隐藏层中均设置正则化层, 避免数据经全连接层和转置卷积层后落入激活函数的不敏感区使梯度消失, 而在生成器的输出层则不设置正则化层以避免生成场景偏向于服从标准正态分布。根据数据的预处理过程可知, 生成器生成场景的功率大小应落在  $[0,1]$  p.u. 区间上, 因此使用 Sigmoid 函数作为生成器输出层的激活函数。鉴别器的结构与生成器基本对称, 主要区别在于鉴别器采用谱归一化代替正则化提升网络训练稳定性。

### 3.3 模型训练

分别按照图 4 结构和参数搭建生成风电随机场景的模型并进行训练, 训练初期, 由于生成器未学习到  $P_r(x)$  的相关信息, 生成样本和真实样本的差异较大,  $P_r(x)$  和  $P_g(x')$  之间的 MMD 距离较大, 因此, 鉴别器能正确区分其输入来源。随着训练的进行, 生成器逐渐学习到  $P_r(x)$  的相关信息,  $P_g(x')$  和  $P_r(x)$  间的重合度增大, 两者的 MMD 距离减小, 当

MMD 距离收敛到 0 附近时, 整个网络达到纳什均衡状态, 此时生成样本和真实样本服从同一概率分布。图 5 展示了训练过程中  $P_g(x')$  与  $P_r(x)$  两者的 MMD 距离随着训练迭代次数的增加而衰减的过程。由图 5 可知,  $P_g(x')$  与  $P_r(x)$  两者的 MMD 距离随着迭代次数的增加而迅速衰减, 当迭代次数大于 2000 次时, 该值收敛至 0 附近, 表明网络已达到纳什均衡状态。为确保 MMD 距离收敛, 使改进 MMD-GAN 继续迭代至 4000 次后结束训练。

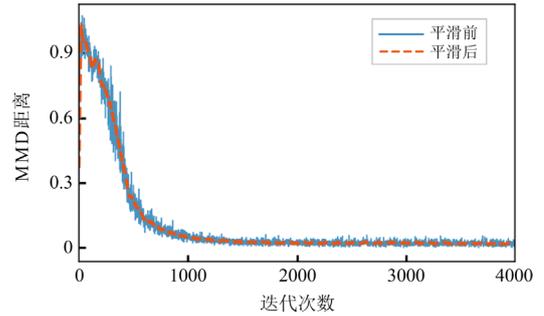


图 5 MMD 距离随训练过程的变化

Fig. 5 Change of MMD distance with training process

### 3.4 性能评价

训练完成后, 提取风电和光伏随机场景生成模型中的生成器, 将服从高维正态分布的 1600 个随机噪声数据作为所提取生成器的输入, 分别生成 3200 个风电和光伏出力样本用于后续的性能评价。

#### 1) 时间序列

随机选择测试集中的部分真实样本, 并从  $X'$  中筛选出与所选真实样本具有相似视觉属性的样本进行可视化处理, 按式(8)计算所选样本对的平均绝对误差(mean absolute error, MAE), 部分风电样本对的对比如图 6 所示, 对应 MAE 计算结果如表 1 所示。

$$M(x_i, x'_i) = \frac{1}{T} \sum_{t=1}^T |x_{i,t} - x'_{i,t}| \quad (8)$$

式中:  $M$  表示样本对的 MAE;  $T$  表示采样点个数;  $x_{i,t}$  和  $x'_{i,t}$  分别表示  $x_i$  和  $x'_i$  在第  $t$  个采样点上的功率。

图 6 中, 第 1 行和第 2 行分别为风电测试集真实样本和生成样本, 可看出两者具有相似的形状轮廓, 同时, 表 1 中具有相似视觉属性的样本对的 MAE 低至 0.0164 p.u., 表明本文所提方法可准确学习可再生能源出力的动态特性。在训练过程中并未使用测试集中的样本, 但训练好的生成器却能生成与其模式相似的样本, 说明本文所提方法并非简单记住历史数据, 而是根据历史数据的内在特征生成与其具有相同内在特征的数据。

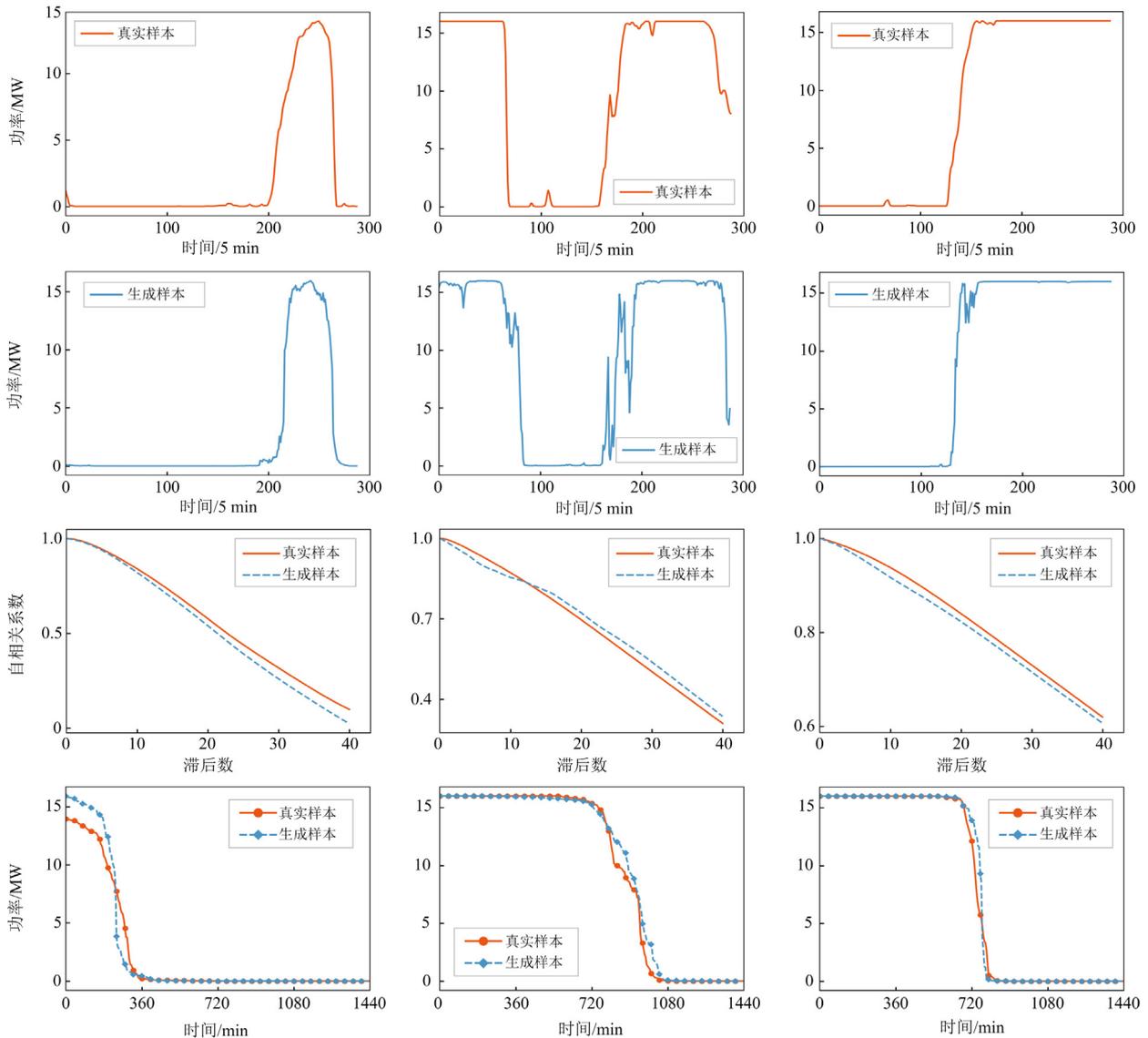


图 6 风电场景对比

Fig. 6 Comparison of wind power scenarios

表 1 风电样本对的 MAE

Table 1 MAE of wind power sample pairs

	样本对 1	样本对 2	样本对 3
MAE/p.u	0.0348	0.0889	0.0164

为检验生成样本和真实样本在时间统计特性上的一致性，分别按式(9)计算其自相关系数  $R(\tau)$  [22]。

$$R(\tau) = \frac{E[(S_t - \mu)(S_{t+\tau} - \mu)]}{\sigma^2} \quad (9)$$

式中： $S_t$  和  $S_{t+\tau}$  分别为  $S$  中  $t$  时刻和  $t + \tau$  时刻的值，其中  $S$  为一组时序数据； $\mu$  和  $\sigma$  分别为该组时序数据的均值与方差； $\tau$  为时间间隔。

图6的第3行展示了风电样本对的自相关系数，

对比可知，生成样本较好地拟合了真实样本自相关系数的变化趋势，表明本文所提方法可体现风电出力在时间上的相关性。

图6的第4行对风电样本对的功率持续时间曲线进行了可视化处理，类比负荷持续时间曲线，可再生能源持续时间曲线下的面积代表风电场或者光伏电站一天内产生的所有能量[30]。对比可知，两者具有高度相似性，且与横轴、纵轴围成区域的面积基本相等，表明使用改进 MMD-GAN 方法生成的场景在 1 天内的发电量与真实场景的发电量基本一致。

通过上述分析可知，改进 MMD-GAN 方法不仅能准确学习可再生能源出力的动态特性，如功率爬坡、峰谷变化等，还能兼顾可再生能源出力的时间

相关性和对外输出电量。

### 2) 概率分布

对于任一实数  $y$ , 可使用核密度估计方法按式 (10) 对其概率密度函数  $\hat{f}_h(y)$  进行估算<sup>[17]</sup>, 核密度估计结果接近其实际分布的概率密度函数。

$$\hat{f}_h(y) = \frac{1}{Bh} \sum_{i=1}^B K\left(\frac{y-y_i}{h}\right) \quad (10)$$

式中:  $y_i$  为来自未知分布的随机样本;  $B$  为样本数量;  $K(\cdot)$  为核平滑函数;  $h$  为带宽。

为进一步分析生成样本是否与真实样本具有相同的概率分布规律, 统计并绘制某时间剖面下风电的频率分布直方图及相应的核密度估计曲线, 结果如图 7 所示。

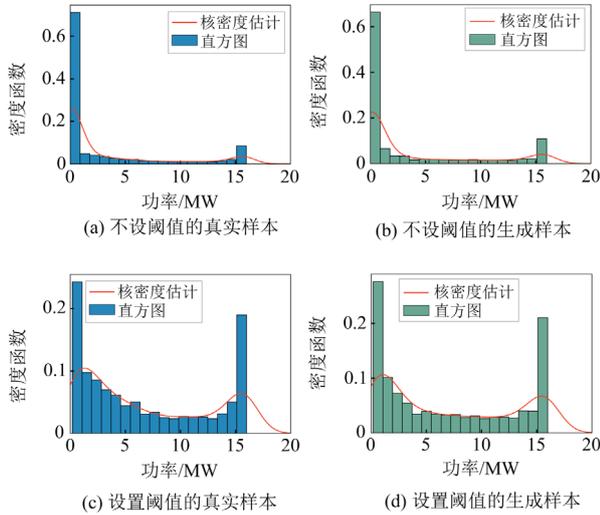


图 7 概率密度函数(风电)

Fig. 7 Probability density function (wind power)

将该时间剖面下的所有样本纳入统计范围时, 真实出力及生成出力的频率分布直方图、核密度估计曲线分别如图 7(a)和图 7(b)所示。观察可知, 此时风电分布强烈偏向一侧, 表明存在大量零功率样本, 这是由于风电在有效风速范围外无法输出功率<sup>[31]</sup>。然而, 这些零功率样本对概率分布特征的分析并无实际意义, 为进一步对比生成样本和真实样本在非零出力下统计特征的相似性, 设置风电出力阈值为 0.2 MW, 对两者的频率分布直方图及核密度估计曲线进行统计分析, 结果如图 7(c)和图 7(d)所示。

由图 7 可以看出, 使用改进 MMD-GAN 方法生成的风电分布与真实分布高度相似, 表明所提方法可学习到真实样本的统计分布规律, 可认为生成样本与真实样本服从同一概率分布。

### 3) 相关性

Pearson 相关系数  $\rho_{X,Y}$  是评价连续变量之间关系的常用指标之一, 常用于衡量可再生能源出力在不同预测时间的线性关系<sup>[3]</sup>, 其定义如式(11)所示。

$$\rho_{X,Y} = \frac{E[(X - \bar{X})(Y - \bar{Y})]}{\sigma_X \sigma_Y} \quad (11)$$

式中:  $X$  和  $Y$  为随机变量;  $\bar{X}$  和  $\sigma_X$ 、 $\bar{Y}$  和  $\sigma_Y$  分别为  $X$ 、 $Y$  的期望和标准差。

为验证使用改进 MMD-GAN 生成的样本与真实样本具有相同的时间特征, 将真实样本及生成样本的 Pearson 相关矩阵进行可视化处理, 风电结果如图 8 所示。

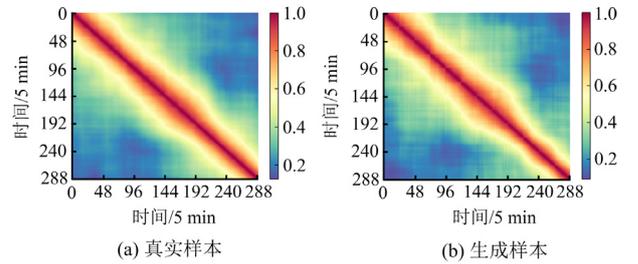


图 8 Pearson 相关矩阵(风电)

Fig. 8 Pearson correlation matrix (wind power)

由图 8 可以看出, 生成样本和真实样本的 Pearson 相关矩阵高度相似, 表明所提方法在训练时不需要对历史数据服从的概率分布进行任何统计假设, 即可准确学习风电出力的时间依赖性。光伏的场景生成结果如附录 A 图 A2—图 A5 所示, 光伏样本对的 MAE 如附录 A 表 A1 所示, 与风电的场景生成分析结果类似。

### 3.5 方法对比

以生成风电随机场景为例, 算例将改进 MMD-GAN 方法与 MMD-GAN、WGAN 方法进行仿真对比, 为避免网络结构对结果的影响, 3 种方法对应的网络结构均按图 4 结构进行设置。使用 3 种方法训练完成后, 分别使用 3 种方法对应的生成器生成 3200 个风电出力样本, 用于后续性能对比。

为直观比较 3 种方法的性能, 随机选择测试集中的部分真实样本, 并分别从 3 种方法的生成样本集中筛选出与所选真实样本具有相同视觉属性的样本进行可视化处理。其中的两个典型真实样本及其分别对应的生成样本如图 9 所示, 其中第 1 行—第 4 行分别为真实样本及使用改进 MMD-GAN、MMD-GAN、WGAN 3 种方法生成的样本, 第 5 行为其自相关系数。同时定量计算真实样本及其分别对应的生成样本的 MAE, 结果如表 2 所示。

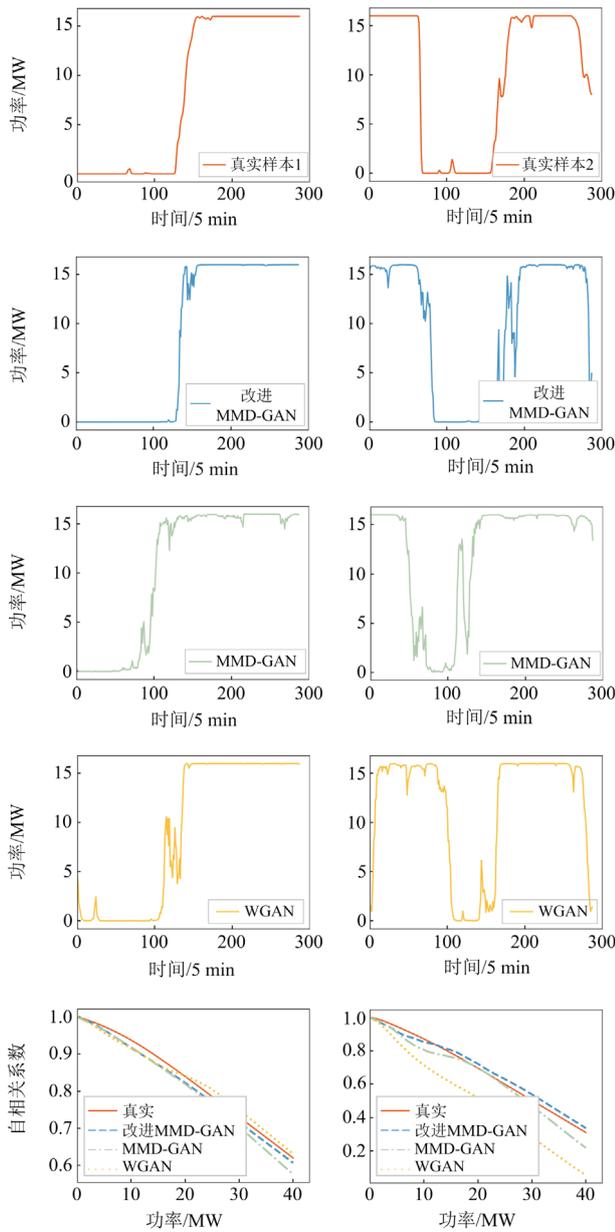


图 9 不同方法生成的场景对比

Fig. 9 Comparison of scenarios generated by different methods

表 2 MAE 对比

Table 2 Comparison of MAE

	改进 MMD-GAN	MMD-GAN	WGAN
真实样本 1	0.0164	0.1410	0.0528
真实样本 2	0.0889	0.2395	0.1929

由图 9 及表 2 可看出, 3 种方法均可学习到风电的出力特征, 但使用改进 MMD-GAN 方法生成的样本与真实样本的 MAE 比 MMD-GAN 及 WGAN 的小, 且在学习真实样本的时间相关性方面优于另外两种方法, 表明改进 MMD-GAN 方法精度更高。

除从视觉上直观比较单个随机场景外, 本文从统计学特征上对 3 种方法进行对比分析, 首先, 统计并绘制风电测试集样本和 3 种方法生成的样本在某个相同时间剖面上的概率密度函数; 另外, 为定量分析不同方法的生成分布与真实分布之间的相似程度, 按式(12)计算各生成分布与真实分布的 Wasserstein 距离<sup>[14]</sup>, Wasserstein 距离越小, 两分布越相似。

$$W(P_r(x), P_g(x')) = \inf_{\pi(x, x')} \int d(x, x') \pi(dx, dx') \quad (12)$$

式中:  $W(P_r(x), P_g(x'))$ 、 $\pi(x, x')$  分别为  $P_r(x)$  与  $P_g(x')$  的 Wasserstein 距离、满足两者边缘分布的联合概率密度分布;  $d(x, x')$  为场景间的距离测度。

图 10 为测试集的真实样本和 3 种方法生成的样本在某个相同时间剖面上的概率密度函数, 表 3 为各生成分布与真实分布的 Wasserstein 距离。通过对比可知, 虽然使用 3 种方法生成的分布总体趋势与真实分布一致, 但改进 MMD-GAN 方法生成的分布与真实分布相似度更高; 在生成分布与真实分布的 Wasserstein 距离上, 改进 MMD-GAN 分别比 MMD-GAN 及 WGAN 降低了 54.8% 及 59.4%, 表明改进 MMD-GAN 方法比 MMD-GAN 与 WGAN 具有更好的概率分布拟合能力。

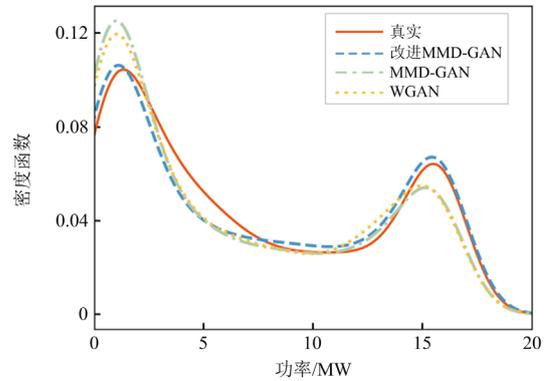


图 10 概率密度函数对比

Fig. 10 Comparison of probability density function

表 3 Wasserstein 距离对比

Table 3 Comparison of Wasserstein distance

方法	改进 MMD-GAN	MMD-GAN	WGAN
Wasserstein 距离	0.3530	0.7805	0.8695

综上所述, 与 MMD-GAN 及 WGAN 方法对比, 本文所提改进 MMD-GAN 方法明显提高了生成场景精度及生成分布与真实分布的相似度, 降低了生成分布与真实分布的 Wasserstein 距离, 能够更加准确地描述可再生能源的出力不确定性。

## 4 结论

针对可再生能源随机场景生成方法过分依赖假设统计及生成场景准确性低等问题, 本文提出了一种基于改进 MMD-GAN 的可再生能源随机场景生成方法。通过详细阐述改进 MMD-GAN 的基本原理及其应用于可再生能源随机场景生成的流程, 并基于可再生能源出力特征, 设计适用于生成高质量随机场景的改进 MMD-GAN 网络结构, 最后结合实际数据仿真分析, 验证了本文所提随机场景生成方法的有效性和优越性。本文具体结论如下所述。

1) 本文将改进 MMD-GAN 方法引入到可再生能源随机场景生成领域, 并引入谱归一化和有界高斯核提升网络的稳定性, 解决了传统 GAN 梯度消失、训练失稳等问题。

2) 基于可再生能源出力特征设计的改进 MMD-GAN 网络结构, 可有效学习到可再生能源的出力特征, 与 MMD-GAN 及 WGAN 方法相比, 改进 MMD-GAN 方法在生成分布和真实分布的 Wasserstein 距离上较两者降低超过 50%, 生成的场景精度得到有效提升。

3) 本文所提方法不需要复杂的概率建模、概率抽样等步骤, 仅基于历史数据即可学习可再生能源的出力特征, 且具备更高的准确度, 为可再生能源随机场景生成提供了一种新的有效工具。

本文对改进的 MMD-GAN 在可再生能源随机场景生成中的应用进行了初步研究, 后续工作将尝试对所提方法的结构进行改进, 使其生成的场景带有标签或同时考虑多个发电场站之间的关联关系。

### 附录 A

生成对抗网络的目标函数为

$$\min_G \max_D V(G, D) = E_{x \sim P_g(x)} [\log(D(x))] + E_{x' \sim P_g(x')} [\log(1 - D(x'))] \quad (A1)$$

式中,  $V(G, D)$  表示二分类的交叉熵函数。

两个概率分布  $P$  和  $Q$  之间 MMD 距离的平方  $M_k^2(P, Q)$  表达式为

$$M_k^2(P, Q) = \frac{1}{n_1(n_1 - 1)} \sum_{i=1}^{n_1} \sum_{j \neq i}^{n_1} k(a_i, a_j) + \frac{1}{n_2(n_2 - 1)} \sum_{i=1}^{n_2} \sum_{j \neq i}^{n_2} k(b_i, b_j) - \frac{2}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} k(a_i, b_j) \quad (A2)$$

式中:  $n_1$  和  $n_2$  分别为  $A$  和  $B$  的样本个数, 其中  $A$  和  $B$  分别为服从概率分布  $P$  和  $Q$  的样本集合;  $a_i$  和  $a_j$

分别为  $A$  中第  $i$  个和第  $j$  个样本;  $b_i$  和  $b_j$  分别为  $B$  中第  $i$  个和第  $j$  个样本;  $k(\cdot)$  表示某种核变换。

$k_D(\cdot)$  表达式为

$$\begin{cases} k_D(x_i, x_j) = k(D(x_i), D(x_j)) \\ k_D(x'_i, x'_j) = k(D(x'_i), D(x'_j)) \\ k_D(x_i, x'_j) = k(D(x_i), D(x'_j)) \end{cases} \quad (A3)$$

式中, 核变换  $k(\cdot)$  采用高斯核变换, 表达式为

$$k(D(x_i), D(x_j)) = \exp\left(-\frac{1}{2\sigma^2} |D(x_i) - D(x_j)|^2\right) \quad (A4)$$

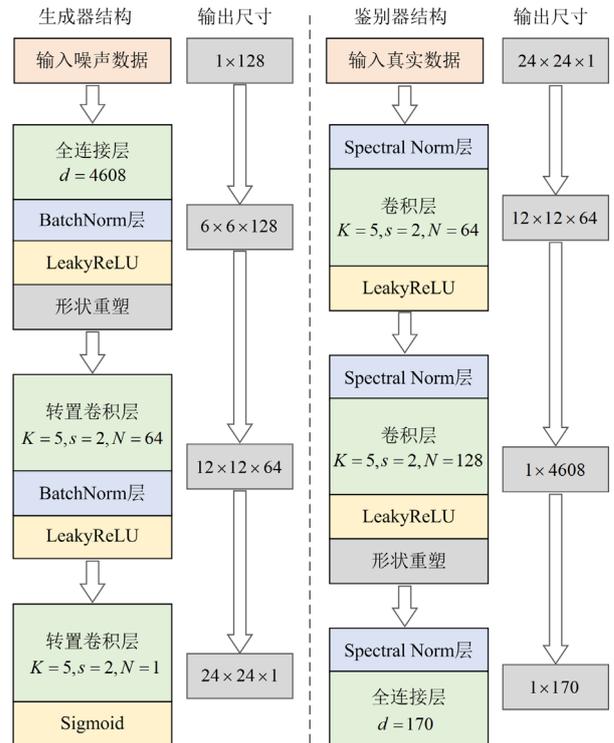


图 A1 生成光伏随机场景的网络结构和参数

Fig. A1 Network architecture and parameters for generating PV power stochastic scenarios

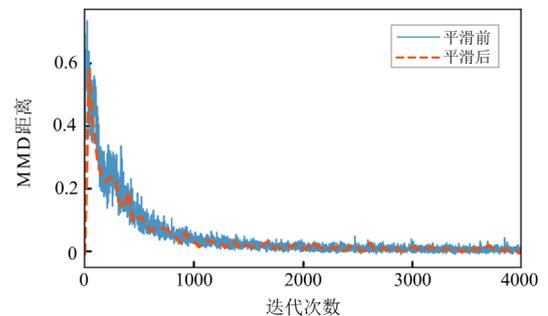


图 A2 MMD 距离随训练过程的变化(光伏)

Fig. A2 Change of MMD distance with training process (PV power)

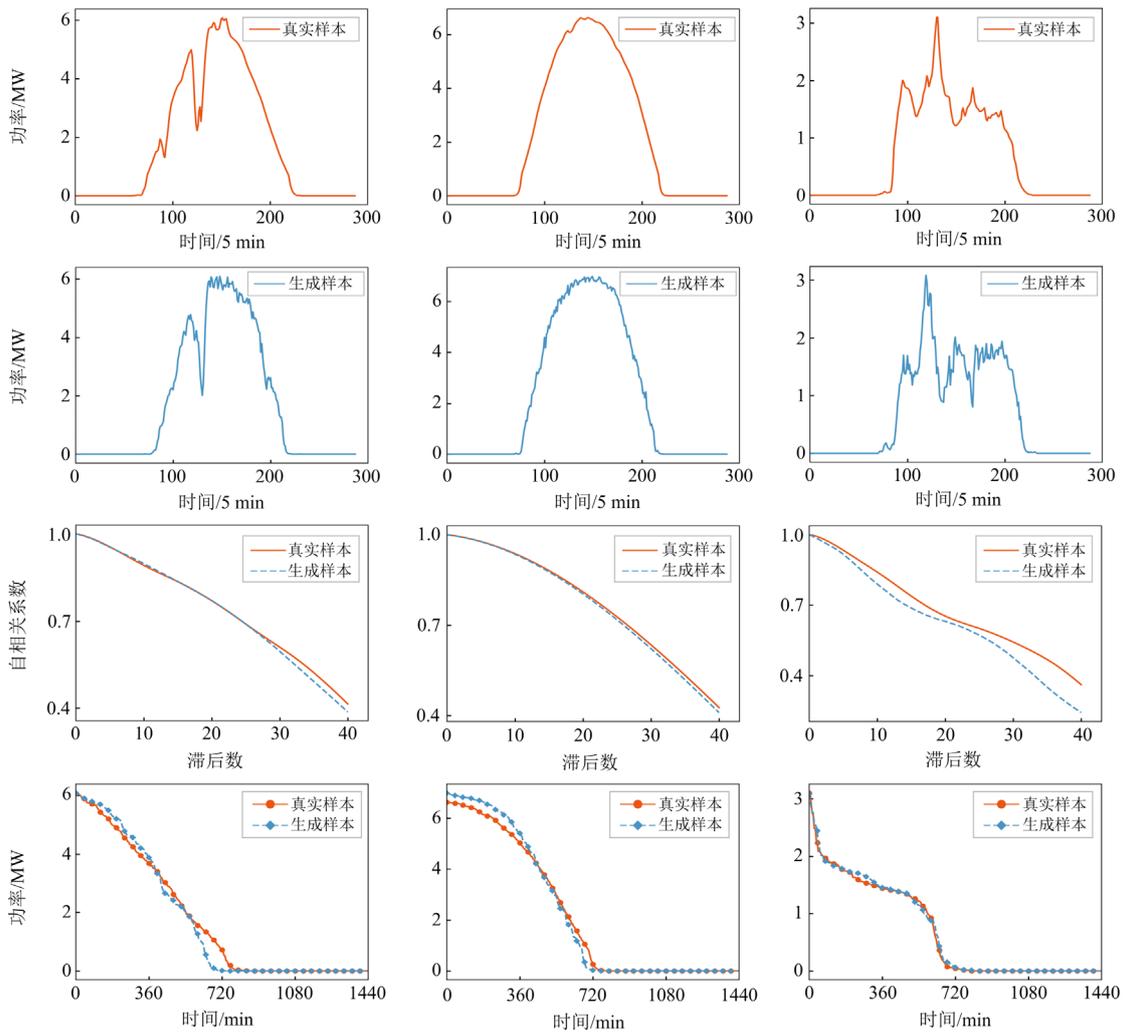


图 A3 光伏场景对比

Fig. A3 Comparison of PV power scenarios

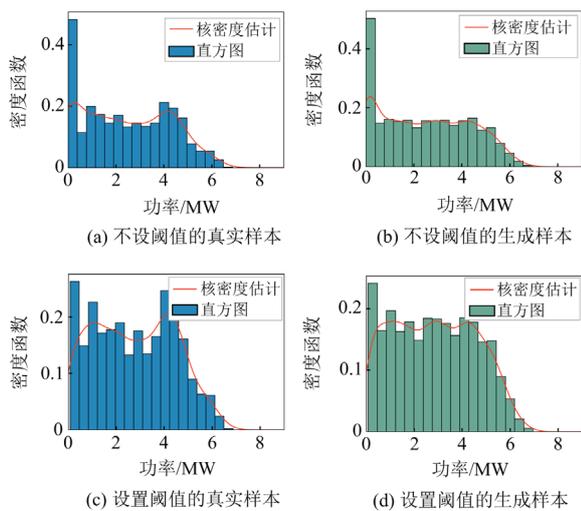


图 A4 概率密度函数(光伏)

Fig. A4 Probability density function (PV power)

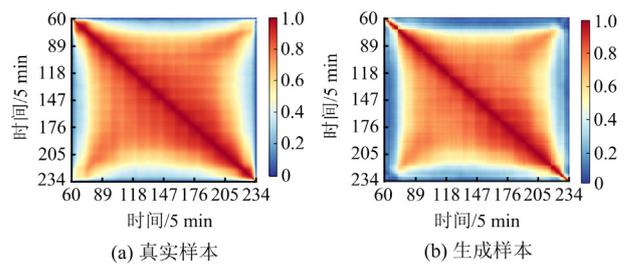


图 A5 Pearson 相关矩阵(光伏)

Fig. A5 Pearson correlation matrix (PV power)

表 A1 光伏样本对的 MAE

Table A1 MAE of PV power sample pairs

	样本对 1	样本对 2	样本对 3
MAE/p.u.	0.0157	0.0117	0.0140

参考文献

[1] 惠慧, 李蕊, 朱逸楠, 等. 含高比例分布式光伏的配电

- 网多目标概率规划方法[J]. 电测与仪表, 2023, 60(11): 2-10.
- HUI Hui, LI Rui, ZHU Yidi, et al. Multi-objective probabilistic planning method for distribution network with high proportion of distributed photovoltaic[J]. *Electrical Measurement & Instrumentation*, 2023, 60(11): 2-10.
- [2] 乔志杰, 马临超. 适用于电力系统稳定性提升的大规模储能控制策略研究[J]. 高压电器, 2022, 58(12): 75-84, 91.
- QIAO Zhijie, MA Linchao. Research on large-scale energy storage control strategy for power system stability improvement[J]. *High Voltage Apparatus*, 2022, 58(12): 75-84, 91.
- [3] LIAO Wenlong, BIRGITTE B J, PILLAI J R, et al. Scenario generations for renewable energy sources and loads based on implicit maximum likelihood estimations[J]. *Journal of Modern Power Systems and Clean Energy*, 2022, 10(6): 1563-1575.
- [4] 鲁肖龙, 潘淼, 鞠立伟, 等. 考虑碳捕集和电转气的热电联合虚拟电厂调度优化模型[J]. 电力建设, 2023, 44(8): 107-117.
- LU Xiaolong, PAN Miao, JU Liwei, et al. Dispatching optimization model of combined heat and power virtual power plant considering carbon capture and power-to-gas[J]. *Electric Power Construction*, 2023, 44(8): 107-117.
- [5] 马国真, 林毓军, 张泽亚, 等. 计及源荷多重不确定性的综合能源系统鲁棒经济调度方法[J]. 电力系统保护与控制, 2021, 49(20): 43-52.
- MA Guozhen, LIN Yujun, ZHANG Zeya, et al. A robust economic dispatch method for an integrated energy system considering multiple uncertainties of source and load[J]. *Power System Protection and Control*, 2021, 49(20): 43-52.
- [6] 闫庆友, 刘达, 李金孟, 等. 基于场景生成与 IGDT 的风光-碳捕集-P2G 虚拟电厂经济调度[J]. 智慧电力, 2023, 51(2): 1-7.
- YAN Qingyou, LIU Da, LI Jinqing, et al. Economic dispatching of wind power-PV-carbon capture-P2G virtual power plant based on scenario generating and IGDT[J]. *Smart Power*, 2023, 51(2): 1-7.
- [7] 金国彬, 刘玉龙, 李国庆, 等. 考虑可靠性的交直流混合配电网网架与分布式电源协同优化规划[J]. 电力系统保护与控制, 2022, 50(22): 59-70.
- JIN Guobin, LIU Yulong, LI Guoqing, et al. Collaborative optimization planning of an AC/DC hybrid distribution network frame and distributed power generation considering reliability[J]. *Power System Protection and Control*, 2022, 50(22): 59-70.
- [8] 李强, 王尧, 胡迎迎, 等. 考虑碳交易成本的输储多阶段协调规划方法[J]. 中国电力, 2023, 56(12): 199-205.
- LI Qiang, WANG Yao, HU Yingying, et al. Multi-stage coordinated planning method for transmission network and energy storage considering carbon trading cost[J]. *Electric Power*, 2023, 56(12): 199-205.
- [9] 郝文波, 景菲, 颜庆宇, 等. 数据驱动下基于风电场景的多时间尺度调峰调度研究[J]. 电力系统保护与控制, 2023, 51(16): 115-126.
- HAO Wenbo, JING Fei, YAN Qingyu, et al. A multi-time scale peak shaving scheduling strategy based on wind power scenario using a data-driven method[J]. *Power System Protection and Control*, 2023, 51(16): 115-126.
- [10] 雷宇, 杨明, 韩学山. 基于场景分析的含风电系统机组组合的两阶段随机优化[J]. 电力系统保护与控制, 2012, 40(23): 58-67.
- LEI Yu, YANG Ming, HAN Xueshan. A two-stage stochastic optimization of unit commitment considering wind power based on scenario analysis[J]. *Power System Protection and Control*, 2012, 40(23): 58-67.
- [11] PAPAETHYMIOS G, KLOCKL B. MCMC for wind power simulation[J]. *IEEE Transactions on Energy Conversion*, 2008, 23(1): 234-240.
- [12] 罗钢, 石东源, 陈金富, 等. 风光发电功率时间序列模拟的 MCMC 方法[J]. 电网技术, 2014, 38(2): 321-327.
- LUO Gang, SHI Dongyuan, CHEN Jinfu, et al. A Markov chain monte carl method for simulation of wind and solar power time series[J]. *Power System Technology*, 2014, 38(2): 321-327.
- [13] MORALES J M, MINGUEZ R, CONEJO A J. A methodology to generate statistically dependent wind speed scenarios[J]. *Applied Energy*, 2010, 87(3): 843-855.
- [14] 邹金, 朱继忠, 赖旭, 等. 基于时空自回归移动平均模型的风电出力序列模拟[J]. 电力系统自动化, 2019, 43(3): 101-107.
- ZOU Jin, ZHU Jizhong, LAI Xu, et al. Simulation of wind power output series based on space-time auto-regressive moving average model[J]. *Automation of Electric Power Systems*, 2019, 43(3): 101-107.
- [15] 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.
- [16] 马喜平, 何世恩, 姚寅, 等. 计及风速不确定性及相关性的风电场分区虚拟惯量估计[J]. 电力系统保护与控制, 2022, 50(10): 123-131.
- MA Xiping, HE Shien, YAO Yin, et al. Virtual inertia estimation of wind farm zones with wind speed uncertainty

- and correlation[J]. *Power System Protection and Control*, 2022, 50(10): 123-131.
- [17] 刘蓉晖, 王乐凯, 孙改平, 等. 考虑不确定性的风-光-储合作联盟参与含需求响应的主辅联合市场的竞价交易模型[J]. *电力系统保护与控制*, 2023, 51(11): 96-107. LIU Ronghui, WANG Lekai, SUN Gaiping, et al. Bidding transaction model of wind-solar-storage cooperative alliance participating in the main and auxiliary joint market with demand response considering uncertainty[J]. *Power System Protection and Control*, 2023, 51(11): 96-107.
- [18] CHEN Yize, WANG Yishen, KIRSCHEN D, et al. Model-free renewable scenario generation using generative adversarial networks[J]. *IEEE Transactions on Power Systems*, 2018, 33(3): 3265-3275.
- [19] 董晓翀, 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.
- [20] 李辉, 任洲洋, 胡博, 等. 基于时序生成对抗网络的月度风光发电功率场景分析方法[J]. *中国电机工程学报*, 2022, 42(2): 537-548. LI Hui, REN Zhouyang, HU Bo, et al. A sequential generative adversarial network based monthly scenario analysis method for wind and photovoltaic power[J]. *Proceedings of the CSEE*, 2022, 42(2): 537-548.
- [21] 贾梦瑶, 王玉玮, 宋明浩. 基于时间生成对抗网络的风电随机场景预测[J]. *智慧电力*, 2023, 51(8): 59-66. JIA Mengyao, WANG Yuwei, SONG Minghao. Stochastic scenario prediction of wind power based on time generative adversarial networks[J]. *Smart Power*, 2023, 51(8): 59-66.
- [22] 胡石峰, 朱瑞金, 唐波. 含有梯度惩罚的 WGAN 光伏功率场景生成方法[J]. *电力系统及其自动化学报*, 2022, 34(11): 109-115. HU Shifeng, ZHU Ruijin, TANG Bo. PV power scenario generation method based on Wasserstein generative adversarial networks with gradient penalty[J]. *Proceedings of the CSU-EPSA*, 2022, 34(11): 109-115.
- [23] WANG Wei, SUN Yuan, HALGAMUGE S K. Improving MMD-GAN training with repulsive loss function[C] // 7th International Conference on Learning Representations, May 6-9, 2019, New Orleans, USA: 1-11.
- [24] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C] // 28th International Conference on Neural Information Processing Systems, December 8-13, 2014, Montreal, Canada: 2672-2680.
- [25] 邵振国, 张承圣, 陈飞雄, 等. 生成对抗网络及其在电力系统中的应用综述[J]. *中国电机工程学报*, 2023, 43(3): 987-1004. SHAO Zhenguo, ZHANG Chengsheng, CHEN Feixiong, et al. A review on generative adversarial networks for power system applications[J]. *Proceedings of the CSEE*, 2023, 43(3): 987-1004.
- [26] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C] // 34th International Conference on Machine Learning, August 6-11, 2017, Sydney, Australia: 214-223.
- [27] LI Chunliang, CHANG Weicheng, CHENG Yu, et al. MMD GAN: towards deeper understanding of moment matching network[C] // 31st Conference on Neural Information Processing Systems, December 4-9, 2017, Long Beach, USA: 1-11.
- [28] GRETTON A, BORGWARDT K M, RASCH M J, et al. A kernel two-sample test[J]. *Journal of Machine Learning Research*, 2012, 13: 723-773.
- [29] 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.
- [30] 朱瑞金, 廖文龙, 王玥珑, 等. 基于生成匹配网络的光伏和风电随机场景生成[J]. *高电压技术*, 2022, 48(1): 374-385. ZHU Ruijin, LIAO Wenlong, WANG Yuelong, et al. Stochastic scenarios generation for wind power and photovoltaic system based on generative moment matching network[J]. *High Voltage Engineering*, 2022, 48(1): 374-385.
- [31] DONG Wei, CHEN Xianqing, YANG Qiang. Data-driven scenario generation of renewable energy production based on controllable generative adversarial networks with interpretability[J]. *Applied Energy*, 2022, 308: 1-13.

收稿日期: 2023-12-06; 修回日期: 2024-04-09

作者简介:

吴艳梅(2001—), 女, 硕士研究生, 研究方向为电力系统规划与分析; E-mail: wu\_yanmei@whu.edu.cn

陈红坤(1967—), 男, 通信作者, 博士, 教授, 博士生导师, 研究方向为电能质量分析、电力系统优化规划。E-mail: chkinsz@163.com

(编辑 周金梅)