

基于相似日与BA-WNN的分布式光伏数据虚拟采集方法

葛磊蛟¹, 秦羽飞², 刘嘉恒¹, 白星振²

(1. 天津大学 智能电网教育部重点实验室, 天津 300072; 2. 山东科技大学 电气与自动化工程学院, 山东 青岛 266590)

摘要:我国分布式光伏电站具有点多面广、分散无序等特点,这使得光伏运维数据采集面临投资成本巨大且实时采集任务繁重等难题。提出一种基于相似日、蝙蝠算法与小波神经网络(BA-WNN)相结合的分布式光伏数据虚拟采集方法,实现在同一区域内仅1座分布式光伏电站安装完备的数据采集装置、其他分布式光伏电站安装价格较低的电流采集终端的场景下,完成区域范围内所有电站运维数据的虚拟采集。利用灰色关联度与余弦相似度构成相似性综合指标进行相似日的选取并建立相似日样本集;利用蝙蝠算法对小波神经网络的权值、伸缩因子与平移因子进行不断调整优化,并利用训练好的BA-WNN模型进行区域范围内分布式光伏电站出力数据的“实时+虚拟”采集。以区域范围内9座分布式光伏电站为例,验证了所提方法的可行性与有效性。

关键词:分布式光伏;相似日;相似性综合指标;蝙蝠算法;小波神经网络;虚拟采集

中图分类号:TM 615

文献标志码:A

DOI:10.16081/j.epae.202106011

0 引言

在当前全球能源危机、环境污染加剧的背景下,光伏作为一种高效、清洁的可再生能源,得到世界各国的大力支持并得以迅速发展。2019年,我国光伏发电量为 $2.2426 \times 10^{11} \text{ kW} \cdot \text{h}$, 占有能源全年总发电量的3.1%^[1-2]。光伏电站主要有集中式和分布式2种类型。其中,分布式光伏电站大多具有点多面广、分散无序、数量庞大的特点,导致分布式光伏电站监测数据点数量庞大,仅依靠增加传感器数量以及提高采集频次等方法数据采集成本过高,且多数户用型分布式光伏的用户不愿承担该项费用,这制约了分布式光伏行业的发展。为此,开发一种适用于分布式光伏运维的低成本、高效率的新型分布式光伏运维数据采集方案具有重要意义。

针对分布式光伏行业运维数据采集难题,虚拟采集的概念被提出,其核心思想是先利用同区域电站的外部条件与电站内部数据数值和变化趋势的相似性,结合电站部分数据,对分布式光伏电站的数据进行拟合,然后利用区域范围内已安装完整采集终

端的光伏数据,将未安装采集装置或采集装置不完备的分布式光伏电站的数据补充完善,实现光伏运维数据的获取,节约数据采集成本,其本质是一种实时运维数据的预测技术。

目前,国内外对光伏数据虚拟采集的研究相对较少,但对光伏发电功率预测与数据修复相关方法的研究,可为分布式光伏运维数据虚拟采集提供借鉴。在光伏功率预测方面,根据光伏出力预测过程的不同可将其分为直接预测与间接预测^[3],直接预测是将光伏发电的历史数据(光伏出力和气象因素)作为预测模型的输入量来进行预测,如文献[4]将聚类历史日数据与预测日气象数据作为输入,采用神经网络模型对预测日光伏发电功率进行预测。间接预测首先对光照强度和温度等环境因素进行预测,然后根据光伏发电系统发电模型得到输出功率值。文献[5]采用最小二乘法建立预测日和前一天的外部光照强度与表面光照强度的关系,并结合空间建模和人工神经网络技术提出一种局部预测全球水平光照强度的方法,该方法能够帮助电网运营商预测与光伏发电性质相关的问题。直接预测不用预置大量的环境监测设备,成本相对较低,但其需要大量的数据支持,如文献[6]将天气分为平稳与突变2种情况,平稳天气下光伏出力较容易预测,突变天气下结合相似时段筛选和神经网络分时段预测光伏出力,最后按顺序排列各时段的预测结果得到完整的预测值。在光伏数据修复方面,文献[7]提出一种考虑天气和空间相关性的光伏电站输出功率数据修复方法,将待修复电站历史相似数据与空间相关电站的数据结合,利用BP神经网络实现数据的修复,扩展光伏电站输出功率数据修复的应用场景。上述研究仅针对未来功率的预测与历史数据的修复,而未涉

收稿日期:2020-08-31; **修回日期:**2021-04-23

基金项目:国家重点研发计划项目(2018YFB1500800);国家自然科学基金资助项目(51807134);国家电网有限公司总部科技项目(高渗透率分布式光伏接入电网动态特性及稳定运行控制技术研究)

Project supported by the National Key R&D Program of China (2018YFB1500800), the National Natural Science Foundation of China(51807134) and the Science and Technology Project of Headquarters of State Grid Corporation of China (Research on Dynamic Characteristics and Stable Operation Control Technology of Power System with High-permeability Distributed Photovoltaic)

及实时光伏数据的采集与预测。

文献[8-11]利用灰色关联度进行相似日的选取,但其仅考虑了特征向量的数值相关度,而未考虑其变化趋势的相关度,因此本文引入余弦相似度构造综合相似日指标。文献[11]利用相似日与小波神经网络WNN(Wavelet Neural Network)进行短期光伏功率预测,但其存在对小波神经网络初值敏感性不足的问题。此外,电流采集终端的成本较低,且电站内的电流数据能够较好地反映该电站出力的不确定性,因此本文将电流数据作为一项输入因子。本文借鉴以上经验并结合实际项目工程的需求,提出一种基于相似日和蝙蝠算法与小波神经网络BA-WNN(Bat Algorithm and Wavelet Neural Network)相结合的分布式光伏数据虚拟采集方法。为了减少预测模型的复杂度,提高数据采集的精度,本文首先利用灰色关联理论与余弦相似度相结合的指标进行相似日的选取,利用选取相似日的功率、电流数据建立小波神经网络的训练样本;然后,利用蝙蝠算法BA(Bat Algorithm)对小波神经网络模型的权值、伸缩因子与平移因子进行优化,避免其陷入局部收敛,提高虚拟采集精度;最后,通过仿真实验实现分布式光伏电站功率数据的虚拟采集,验证本文所提模型的可行性与有效性。

1 基于灰色关联理论与余弦相似度相结合的相似日选取

分布式光伏发电功率受到气候、地域等多因素影响,但是一个特定地域内光伏发电的历史数据具有一定相似性,即相似日。本文采用灰色关联理论与余弦相似度相结合的方法选取相似日,减少模型需要的历史数据量,简化虚拟采集模型。选取相似日的步骤^[8-10]如下。

(1)构造气象特征向量。

一般而言,光伏发电的主要决定因素为光照强度,光伏电站的输出功率随光照强度的变化而变化。图1为电站在平稳天气与突变天气2种天气特性下的光照强度与光伏出力特征曲线。从图中可以看出,无论天气特征如何,输出功率的曲线特征与光照强度的曲线特征总是极为相似的,二者具有较高的关联度。参考相关文献并结合实际数据,本文将光照强度作为相似日选取的特征向量^[12-15]。

考虑历史数据对光伏发电的影响具有“近大远小”的特点^[16],本文选取近3个月的历史数据作为训练集,采用每个历史日光照强度的平均值与各时刻的光照强度构造特征向量 $\mathbf{X}$:

$$\mathbf{X}=[F_1, F_2, \dots, F_g, \dots, F_{av}] \quad (1)$$

其中, F_g 为第 g 个时刻的光照强度; F_{av} 为光照强度的

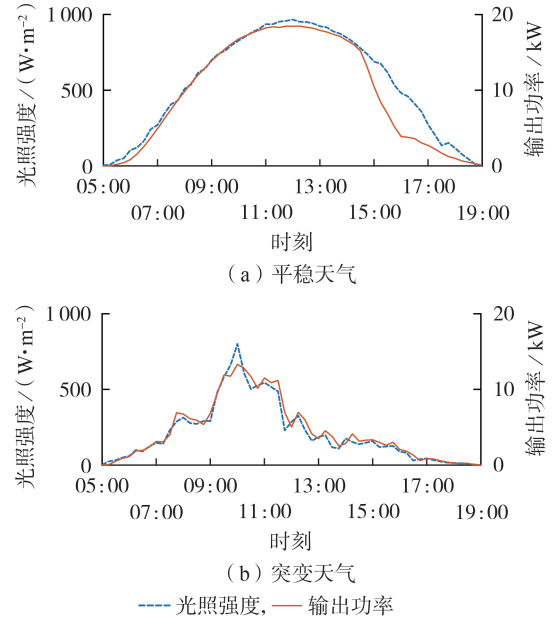


图1 光照强度与输出功率曲线图

Fig.1 Irradiance and output power curves

平均值。

(2)数据归一化。

对光照强度数据进行如下归一化处理:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

其中, $x, x_{\min}, x_{\max}$ 分别为原始数据中的任意值、原始数据中的最小值、原始数据中的最大值; x' 为归一化后的数据。

归一化后的待采集日与历史日的特征向量分别为:

$$\mathbf{x}_0 = [x_0(1), x_0(2), \dots, x_0(n), \dots] \quad (3)$$

$$\mathbf{x}_i = [x_i(1), x_i(2), \dots, x_i(n), \dots] \quad (4)$$

其中, $\mathbf{x}_0$ 为待采集日的特征向量; $\mathbf{x}_i$ 为第 i 个历史日的特征向量; $x_0(n), x_i(n)$ 分别为待采集日、第 i 个历史日特征向量的第 n 个元素。

(3)关联度计算。

分别计算 $\mathbf{x}_0$ 与 $\mathbf{x}_i$ 在第 n 个分量的关联系数:

$$\xi_i(n) = \frac{\min_i \min_n \Delta + r \max_i \max_n \Delta}{\Delta + r \max_i \max_n \Delta} \quad (5)$$

其中, $\xi_i(n)$ 为关联系数; $\Delta = |x_0(n) - x_i(n)|$; r 为分辨系数,一般取为0.5。

由于关联系数很多,信息过于分散,不便于比较,为综合各个分量的关联系数,一般采用求平均值的方式进行处理, $\mathbf{x}_0$ 与 $\mathbf{x}_i$ 的灰色关联度定义为:

$$R_i = \frac{1}{N} \sum_{n=1}^N \xi_i(n) \quad (6)$$

其中, N 为各分量的关联系数总数。

余弦相似度定义为:

$$D_{\cos i} = \frac{\sum_{n=1}^N x_i(n) x_0(n)}{\sqrt{\sum_{n=1}^N x_i^2(n) \sum_{n=1}^N x_0^2(n)}} \quad (7)$$

灰色关联度分析法是一种灰色系统分析方法,其根据数据序列数值之间的相似程度来判断关联程度,是一种衡量因素数值间关联程度的方法。而余弦相似度则是一种衡量数据序列变化趋势间相似性的方法。为避免单一评价方法的局限性,综合考虑数据序列数值的相似性与变化趋势的相似性,本文将灰色关联度 R_i 和余弦相似度 $D_{\cos i}$ 这2个指标组合成1个相似性综合指标 S_i , 表征总体相似性,其计算公式如式(8)所示,其值越接近1表示越相似。

$$S_i = \gamma R_i + (1 - \gamma) D_{\cos i} \quad (8)$$

其中, γ 为经验权重系数,本文取为0.5。

(4)相似日选取。

将相似性综合指标 S_i 按由大到小顺序排列,选取前10 d作为相似日数据集。

2 基于BA-WNN的分布式光伏运维数据的虚拟采集

2.1 小波神经网络原理

小波神经网络是将神经网络的激活函数替换成小波函数,而相应的输入层到隐含层的权值和激活阈值由小波函数的尺度伸缩因子和时间平移因子所代替,这是使用最广泛的紧致性结构,网络结构见图2。图中, $I_i (i=1, 2, \dots, q)$ 为小波神经网络的输入数据, q 为输入层节点总数; $y(k) (k=1, 2, \dots, m)$ 为小波神经网络的输出数据, m 为输出层节点总数; $\omega_{ij} (i=1, 2, \dots, q; j=1, 2, \dots, p)$ 为输入层和隐含层的权值, p 为隐含层节点总数; $\omega_{jk} (j=1, 2, \dots, p; k=1, 2, \dots, m)$ 为隐含层和输出层的权值; $\varphi_j (j=1, 2, \dots, p)$ 为隐含层输出。

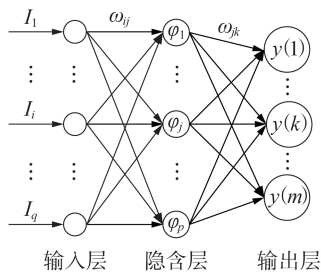


图2 小波神经网络结构

Fig.2 Structure of wavelet neural network

本文选取 Morlet 函数作为小波基函数^[17]:

$$\rho(\kappa) = \cos(1.75\kappa) e^{-\kappa^2/2} \quad (9)$$

其中, κ 为该函数的输入变量。

隐含层输出公式为:

$$\varphi_j = \rho \left(\frac{\sum_{i=1}^q \omega_{ij} I_i - b_j}{a_j} \right) \quad j=1, 2, \dots, p \quad (10)$$

其中, a_j, b_j 分别为隐含层第 j 个节点的伸缩因子、平移因子。

输出层计算公式为:

$$y(k) = \sum_{j=1}^p \omega_{jk} \varphi_j \quad k=1, 2, \dots, m \quad (11)$$

网络误差定义为:

$$e = \sum_{k=1}^m |y'(k) - y(k)| \quad (12)$$

其中, $y'(k)$ 为期望输出值; $y(k)$ 为拟合输出值。

相较于BP神经网络,小波神经网络的基元和整个结构是依据小波分析理论确定的,有效避免了BP神经网络等结构设计上的盲目性。此外,小波理论是全尺度分析,不仅有全局最优解,还保持了局部细节最优解,因此小波神经网络具有相对更强的学习能力与更高的精度。但是随着输入维数增加,训练样本增大,小波神经网络的收敛速度大幅下降。同时,当小波神经网络初始化参数不合理时,将导致整个网络学习过程不能收敛^[18],因此本文对神经网络参数初始化问题采用蝙蝠算法进行优化。

2.2 蝙蝠算法优化小波神经网络

蝙蝠算法是一种全局搜索优化算法,其基本思想是将蝙蝠种群个体映射为空间中的可行解,将搜索过程和优化过程模拟成蝙蝠个体搜寻猎物和移动的过程,蝙蝠所处位置的优劣以求解问题的适应度函数来衡量,将寻优过程类比为个体的优胜劣汰过程^[19]。蝙蝠算法基本原理如下。

(1)设置蝙蝠算法初始值:最大迭代次数 $I_{\max}$ 、种群规模 N_b 、种群维度 D 、脉冲频率范围 $f_i \in [f_{\min}, f_{\max}]$ (f_i 为蝙蝠个体 i 的脉冲频率, $f_{\max}, f_{\min}$ 分别为脉冲频率的上、下限)、初始脉冲发射率 r_0 、初始脉冲响度 A_0 、脉冲响度控制系数 α 、频度控制系数 μ 及精度 ε 。

(2)随机初始化蝙蝠个体位置 $B_i (i=1, 2, \dots)$, 并根据适应度函数 F_{fitness} 的最小值找到最佳位置 B^* 。

$$F_{\text{fitness}} = \frac{1}{2} \sum_{k=1}^m |y'(k) - y(k)| \quad (13)$$

(3)进行个体位置的寻优,每个个体脉冲频率、速度、位置更新如下:

$$f_i = f_{\min} + (f_{\max} - f_{\min}) \beta \quad (14)$$

$$v_i^{t+1} = v_i^t + (p_i^t - p_{\text{best}}) f_i \quad (15)$$

$$p_i^{t+1} = p_i^t + v_i^{t+1} \quad (16)$$

其中, β 为均匀分布在 $[0, 1]$ 上的随机数; v_i^t 和 v_i^{t+1} 分别为蝙蝠个体 i 在 t 和 $t+1$ 时刻的速度; p_{best} 为当前所有蝙蝠个体的最优位置; p_i^{t+1} 和 p_i^t 分别为蝙蝠个体 i

在 $t+1$ 和 t 时刻的位置。

(4)产生随机数 r_{and1} 用于选取当前最优位置,若 $r_{\text{and1}} > r_i$ (r_i 为迭代中的脉冲发射率),则扰动处于最优位置的蝙蝠个体,以扰动后的位置取代当前位置。

(5)产生随机数 r_{and2} ,若 $r_{\text{and2}} < r_i$,且 $F_{\text{itness}}(B_i) < F_{\text{itness}}(B^*)$,则将蝙蝠个体移动到更新后的位置。

(6)当满足步骤(5)时,分别根据式(17)、(18)调整脉冲发射率和响度:

$$r_i^{t+1} = r_0(1 - e^{-\mu}) \quad (17)$$

$$A_i^{t+1} = \alpha A_i^t \quad (18)$$

其中, r_i^{t+1} 为 $t+1$ 时刻的脉冲发射率; A_i^{t+1} 和 A_i^t 分别为蝙蝠个体 i 在 $t+1$ 和 t 时刻的脉冲响度; α 和 μ 为常量,通常均取为0.9。脉冲响度和脉冲发射率只在解进化时更新,这意味着蝙蝠算法只朝着最优解方向进行。

(7)根据蝙蝠群体适应度值,找出优化后蝙蝠个体的最佳位置。

(8)若达到最大搜索次数或者满足搜索精度,则转入步骤(9),否则转入步骤(2)进行搜索。

(9)输出全局最优解,算法结束^[20]。

2.3 虚拟采集模型的建立

蝙蝠算法具有较好的收敛性和全局搜索能力且模型简单、参数少、通用性强。若将其与小波神经网络相结合,则可利用蝙蝠算法较强的全局收敛能力对小波神经网络中的权值、伸缩因子与平移因子进行不断调整优化,使小波神经网络的性能得到明显提高。

本文建立基于相似日和BA-WNN相结合的分布式光伏运维数据虚拟采集模型,如附录中图A1所示。

3 算例分析

为验证本文虚拟采集方法的有效性与可行性,选取同区域内9座分布式光伏电站作为验证对象,其中1号电站为具有完整数据采集系统的标杆电站,2—9号电站为只安装了电流采集终端的虚拟采集电站。利用1号电站的部分数据与2—9号电站的电流数据,实现区域范围内分布式光伏电站的“实时+虚拟”采集。

3.1 虚拟采集前提条件的判断

虚拟采集本质上是利用1座具有完备数据的分布式光伏站与自身站内的电流数据,实现自身站内功率数据的预测技术。因此,虚拟采集的实现要满足2个前提条件:区域内电站的出力具有相似性;电站内的电流数据能够较完整地反映电站输出功率的变化趋势。下面对这2个前提条件进行验证。

分析各电站光伏出力情况之间的关系,选取同区域内9座电站连续3 d(05:00—19:00)的有效出力

曲线如图3所示。从图中可以看出,同一区域内同时段各电站的光伏出力曲线趋势基本相似。

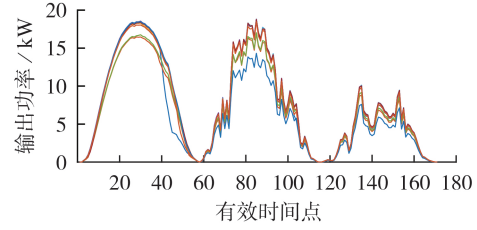


图3 9座电站出力曲线

Fig.3 Output curves of nine power stations

另外,从图4中1号电站输出功率与电流曲线可以看出,电流能够较完整地反映电站输出功率的变化趋势,可以引入各电站电流数据来调整各电站输出功率之间的差异。

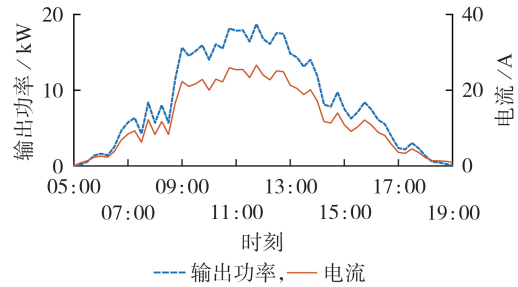


图4 1号电站输出功率与电流曲线

Fig.4 Output power and current curves of No.1 power station

因此,对于本文中选取的同区域9座分布式光伏电站,可以利用1座数据获取完备的电站实现其余电站数据的虚拟采集。

3.2 虚拟采集算例结果与分析

为了验证本文虚拟采集方法的适用性,分别选取平稳天气与突变天气进行验证实验,首先根据第1节分别选取各待采集日的相似日作为神经网络的训练样本。具体选取方法为将 S_i 按由大到小顺序排列,选取前10 d作为相似日数据集,综合相似日选取结果如图5所示。

由图5中可以看出,本文所建立的相似日选取模型效果较好,各相似日的光照强度无论在数值上还是变化趋势上都具有较强的相似性,可以为后续神经网络的训练提供强有力的支撑。

利用相似日的功率数据、电流数据作为小波神经网络的训练样本,将数据归一化后分别对稳定天气与突变天气进行神经网络的训练。小波神经网络相关参数设置为:训练次数为100次,权重因子学习效率为0.01,网络结构为3层,输入层节点总数为2,输出层节点总数为1。待采集数据电站的输出功率隐含层节点数选取的经验公式如式(19)所示,经过多次实验比对结果,本文选取隐含层节点总数为4。

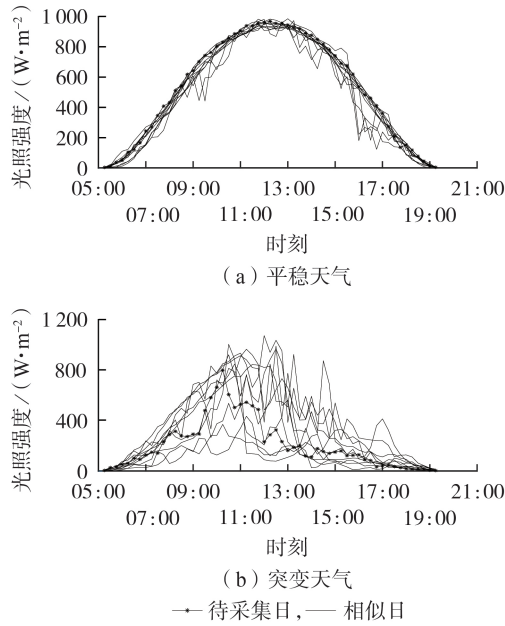


图5 相似日选取结果

Fig.5 Results of similar day selection

$$p = \sqrt{q+m} + a \quad (19)$$

其中, a 为 1~10 之间的常数。

蝙蝠算法相关参数设置为: 最大迭代次数为 100 次, 种群规模为 25, 需要优化的权值、平移因子、伸缩因子的总数量为 $2 \times 4 + 4 \times 1 + 4 + 4 = 20$, 因此蝙蝠算法解空间的维度为 20, 蝙蝠脉冲频率范围为 $[f_{\min}, f_{\max}] = [0, 2]$, 初始脉冲响度为 0.9, 初始脉冲发射率为 0.5, 脉冲响度控制参数为 0.98, 频度控制参数为 0.98, 精度为 0.01。设置完毕后开始训练模型并进行数据的虚拟采集, 虚拟采集结果如附录中图 A2 与图 A3 所示。

本文以均方误差 R_{MSE} 与平均绝对误差 M_{AE} 评价虚拟采集结果, 计算方法分别如式 (20) 与式 (21) 所示。

$$R_{\text{MSE}} = \sqrt{\frac{\sum_{i=1}^{N_p} (P_f^i - P_a^i)^2}{N_p}} \quad (20)$$

$$M_{\text{AE}} = \frac{1}{N_p} \sum_{i=1}^{N_p} |P_f^i - P_a^i| \quad (21)$$

其中, N_p 为所采集数据点的总数; P_f^i 为虚拟采集的光伏输出功率数据; P_a^i 为实际值。

平稳天气与突变天气下本文方法的误差指标如表 1 所示。由图 A2 与图 A3 中可以看出, 虚拟采集模型的实时数据预测值十分接近真实值。由表 1 可以看出, 本文虚拟采集模型的实时数据预测结果与真实值之间误差较小, 虚拟采集模型采集效果较好。为进一步验证本文方法的优越性, 与 BP 神经网络、小波神经网络、基于历史日的 BA-WNN 进行比较,

结果如图 6 所示。各方法的误差如表 2 所示。

表 1 不同天气条件下本文方法的误差指标

Table 1 Error indexes of proposed method under different weather conditions

电站 编号	平稳天气		突变天气	
	$R_{\text{MSE}} / \text{kW}$	$M_{\text{AE}} / \text{kW}$	$R_{\text{MSE}} / \text{kW}$	$M_{\text{AE}} / \text{kW}$
2	0.1940	0.1502	0.1118	0.0783
3	0.2481	0.1903	0.1575	0.1351
4	0.2450	0.1694	0.1725	0.1050
5	0.2266	0.2015	0.0914	0.0691
6	0.1795	0.1509	0.1845	0.1579
7	0.1992	0.1739	0.1775	0.1389
8	0.2468	0.2166	0.1546	0.1350
9	0.1772	0.1503	0.1068	0.0872

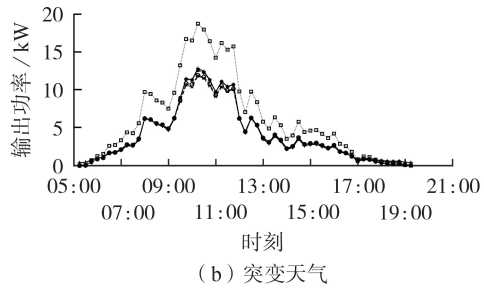
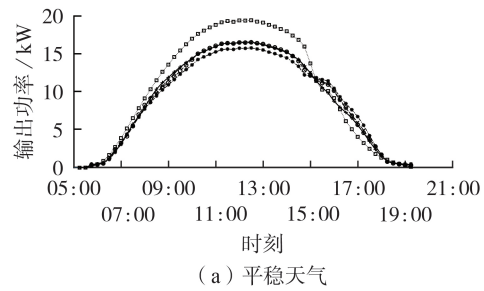


图6 各方法的虚拟采集结果

Fig.6 Virtual acquisition results of each method

表 2 各方法的误差指标

Table 2 Error indexes of each method

方法	平稳天气		突变天气	
	$R_{\text{MSE}} / \text{kW}$	$M_{\text{AE}} / \text{kW}$	$R_{\text{MSE}} / \text{kW}$	$M_{\text{AE}} / \text{kW}$
BP 神经网络	1.9527	1.6275	2.9659	2.2133
小波神经网络	0.5492	0.4673	0.3121	0.1981
基于历史日的 BA-WNN	0.2232	0.1940	0.3856	0.2937
本文方法	0.1940	0.1502	0.1118	0.0783

由图 6 和表 2 可以得出以下结论:

(1) 基于 BP 神经网络虚拟采集得到的数据误差较大, 但其训练过程较好, 说明其存在明显的过拟合状态, 不能很好地拟合需要的数据;

(2) 无论在平稳天气还是突变天气, 相比于基于小波神经网络的虚拟采集方法和基于历史日的 BA-WNN 虚拟采集方法, 本文方法都能够更好地反映光伏输出功率的情况, 相似日的选取与算法的优化都在一定程度上降低了虚拟采集模型的误差, 验

证了本文方法的有效性。

综上所述,对于区域范围内分布式电站光伏出力数据的虚拟采集,本文方法相较于其他方法具有更加优越的性能。

4 结论

本文提出一种基于相似日与BA-WNN相结合的分布式光伏数据虚拟采集方法,并通过算例验证了本文方法的有效性与可行性,可以得到以下结论:

(1)通过灰色关联与余弦相似度相结合建立相似日样本数据,使得虚拟采集的相似日光照强度在数值和变化趋势上都具有较强的相似性,避免了单一方法进行相似度判断的局限性,为后续神经网络的训练提供了强有力的支撑;

(2)本文方法无论在平稳天气还是突变天气,均能获取较为精准的光伏出力采集数据,而且相较于BP神经网络、小波神经网络、基于历史日的BA-WNN具有更加优越的性能,证明了本文虚拟采集技术的可行性;

(3)相较于传统仅依靠传感器设备进行采集的方法,本文方法不仅可大幅节约投入开发的成本,还可避免传输数据量过大造成的传输通道堵塞以及节约存储成本,具有较强的实用性,对我国分布式光伏工程的成本节约具有一定的参考价值。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 江华,金艳梅,叶幸,等. 中国光伏产业2019年回顾与2020年展望[J]. 太阳能,2020(3):14-23.
JIANG Hua, JIN Yanmei, YE Xing, et al. Review of China's PV industry in 2019 and prospect in 2020[J]. Solar Energy, 2020(3):14-23.
- [2] 王贤,刘文颖,夏鹏,等. 光伏电站参与电网主动调压的无功优化控制方法[J]. 电力自动化设备,2020,40(7):76-83.
WANG Xian, LIU Wenying, XIA Peng, et al. Reactive power optimization control method for PV station participating in active voltage regulation of power grid[J]. Electric Power Automation Equipment, 2020, 40(7):76-83.
- [3] 王洪坤,葛磊蛟,李宏伟,等. 分布式光伏发电的特性分析与预测方法综述[J]. 电力建设,2017,38(7):1-9.
WANG Hongkun, GE Leijiao, LI Hongwei, et al. A review on characteristic analysis and prediction method of distributed PV[J]. Electric Power Construction, 2017, 38(7):1-9.
- [4] 李燕青,杜莹莹. 基于双维度顺序填补框架与改进Kohonen天气聚类的光伏发电短期预测[J]. 电力自动化设备,2019,39(1):60-65.
LI Yanqing, DU Yingying. Short-term photovoltaic power forecasting based on double-dimensional sequential imputation framework and improved Kohonen clustering[J]. Electric Power Automation Equipment, 2019, 39(1):60-65.
- [5] AMROUCHE B, LE PIVERT X. Artificial neural network based daily local forecasting for global solar radiation[J]. Applied Energy, 2014, 130:333-341.
- [6] 李建文,焦衡,刘凤梧,等. 基于相似时段的分时段光伏出力短

期预测[J]. 电力自动化设备,2018,38(8):183-188.

- LI Jianwen, JIAO Heng, LIU Fengwu, et al. Short-time segmented photovoltaic output forecasting based on similar period[J]. Electric Power Automation Equipment, 2018, 38(8):183-188.
- [7] 于若英,陈宁,苗森,等. 考虑天气和空间相关性的光伏电站输出功率修复方法[J]. 电网技术,2017,41(7):2229-2236.
YU Ruoying, CHEN Ning, MIAO Miao, et al. A repair method for PV power station output data considering weather and spatial correlations[J]. Power System Technology, 2017, 41(7):2229-2236.
- [8] 梁彩霞,高赵亮. 基于相似日和GA-DBN神经网络的光伏发电短期功率预测[J]. 电气应用,2019,38(3):97-102.
- [9] LUO P, ZHU S C, HAN L J, et al. Short-term photovoltaic generation forecasting based on similar day selection and extreme learning machine[C]//2017 IEEE Power & Energy Society General Meeting. Chicago, IL, USA: IEEE, 2017:1-5.
- [10] CUI Y, ZHANG J, ZHONG W. Short-term photovoltaic output prediction method based on similar day selection with grey relational theory[C]//2019 IEEE Innovative Smart Grid Technologies-Asia(ISGT Asia). [S.l.]: IEEE, 2019:792-797.
- [11] 祝喧懿,姚李孝. 基于相似日和小波神经网络的光伏短期功率预测[J]. 电网与清洁能源,2019,35(3):75-78.
ZHU Xuanyi, YAO Lixiao. Solar power plant short-term power forecast based on similar days and WNN[J]. Power System and Clean Energy, 2019, 35(3):75-78.
- [12] 韦航宇. 基于GA-BP神经网络的光伏电站短期发电功率预测[D]. 南宁:广西大学,2019.
WEI Hangyu. Short-term generation power forecast of photovoltaic power station based on GA-BP neural network[D]. Nanning: Guangxi University, 2019.
- [13] 黎静华,赖昌伟. 考虑气象因素的短期光伏出力预测的奇异谱分析方法[J]. 电力自动化设备,2018,38(5):50-57,76.
LI Jinghua, LAI Changwei. Singular spectrum analysis method for short-term photovoltaic output prediction considering meteorological factors[J]. Electric Power Automation Equipment, 2018, 38(5):50-57, 76.
- [14] 许寅,李佳旭,王颖,等. 考虑光伏出力不确定性的园区配电网日前运行计划[J]. 电力自动化设备,2020,40(5):85-94,106.
XU Yin, LI Jiaxu, WANG Ying, et al. Day-ahead operation plan for campus distribution network considering uncertainty of photovoltaic output[J]. Electric Power Automation Equipment, 2020, 40(5):85-94, 106.
- [15] 王育飞,付玉超,薛花. 计及太阳辐射和混沌特征提取的光伏发电功率DMCS-WNN预测法[J]. 中国电机工程学报,2019,39(增刊1):63-71.
WANG Yufei, FU Yuchao, XUE Hua. DMCS-WNN prediction method of photovoltaic power generation by considering solar radiation and chaotic feature extraction[J]. Proceedings of the CSEE, 2019, 39(Supplement1):63-71.
- [16] 程伟. 基于神经网络的微电网光伏发电及负荷短期预测研究[D]. 济南:山东大学,2019.
CHENG Wei. Research on micro-grid photovoltaic power generation and load short-term forecasting based on neural network[D]. Jinan: Shandong University, 2019.
- [17] 杜杰,彭丽霞,刘玉宝,等. 风电场风机测量风速缺值的组合填充模型[J]. 电力自动化设备,2015,35(9):125-129.
DU Jie, PENG Lixia, LIU Yubao, et al. Combined interpolation model for wind speed measurement missing of wind farm[J]. Electric Power Automation Equipment, 2015, 35(9):125-129.
- [18] 郑金宇. 基于小波神经网络及布谷鸟算法的停车位预测方法研究[D]. 长春:吉林大学,2018.
ZHENG Jinyu. Research on parking space prediction methods based on wavelet neural network and cuckoo search al-

gorithm[D]. Changchun: Jilin University, 2018.

- [19] 吴云, 雷建文, 鲍丽山, 等. 基于改进灰色关联分析与蝙蝠优化神经网络的短期负荷预测[J]. 电力系统自动化, 2018, 42(20): 67-72.

WU Yun, LEI Jianwen, BAO Lishan, et al. Short-term load forecasting based on improved grey relational analysis and neural network optimized by bat algorithm[J]. Automation of Electric Power Systems, 2018, 42(20): 67-72.

- [20] 姚珊. 基于蝙蝠算法的神经网络预测模型的研究及应用[D]. 鞍山: 辽宁科技大学, 2019.

YAO Shan. Neural network prediction model based on bat algorithm research and application[D]. Anshan: University of Science and Technology Liaoning, 2019.

作者简介:



葛磊蛟

葛磊蛟(1983—), 男, 湖北咸宁人, 讲师, 博士, 通信作者, 研究方向为智能配用电、云计算和大数据(**E-mail**: legendgj99@tju.edu.cn);

秦羽飞(1996—), 男, 山东烟台人, 硕士研究生, 主要研究方向为分布式光伏智能运维技术(**E-mail**: 1366769402@qq.com);

刘嘉恒(1997—), 男, 辽宁铁岭人, 硕士研究生, 主要研究方向为分布式光伏智能运维技术(**E-mail**: 1277330496@qq.com)。

(编辑 王锦秀)

Virtual acquisition method of distributed photovoltaic data based on similarity day and BA-WNN

GE Leijiao¹, QIN Yufei², LIU Jiaheng¹, BAI Xingzhen²

(1. Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin 300072, China;

2. School of Electrical and Automation Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract: The distributed photovoltaic power stations in China have the characteristics of widely and disorderly distributed, etc, which makes photovoltaic operation and maintenance data acquisition face problems of huge investment cost and heavy task of real-time acquisition. A virtual acquisition method of distributed photovoltaic data is proposed based on the combination of similar day and BA-WNN (Bat Algorithm and Wavelet Neural Network), which realizes virtual acquisition of operation and maintenance data of all power stations in regional range under the scenario that only one distributed photovoltaic power station is equipped with complete data acquisition device while other distributed photovoltaic power stations are equipped with low-priced current acquisition terminals in a region. The gray correlation degree and cosine similarity are used to form a similarity composite index for selecting similar days and a sample set of similar days is established. The bat algorithm is used to continuously adjust and optimize the weights, expansion factors and translation factors of wavelet neural network, and the trained BA-WNN model is used for “real time + virtual” acquisition of the output data of distributed photovoltaic power stations in regional range. Nine distributed photovoltaic power stations in regional range are taken as examples to verify the feasibility and effectiveness of the proposed method.

Key words: distributed photovoltaic; similar day; similarity composite index; bat algorithm; wavelet neural network; virtual acquisition

附录：

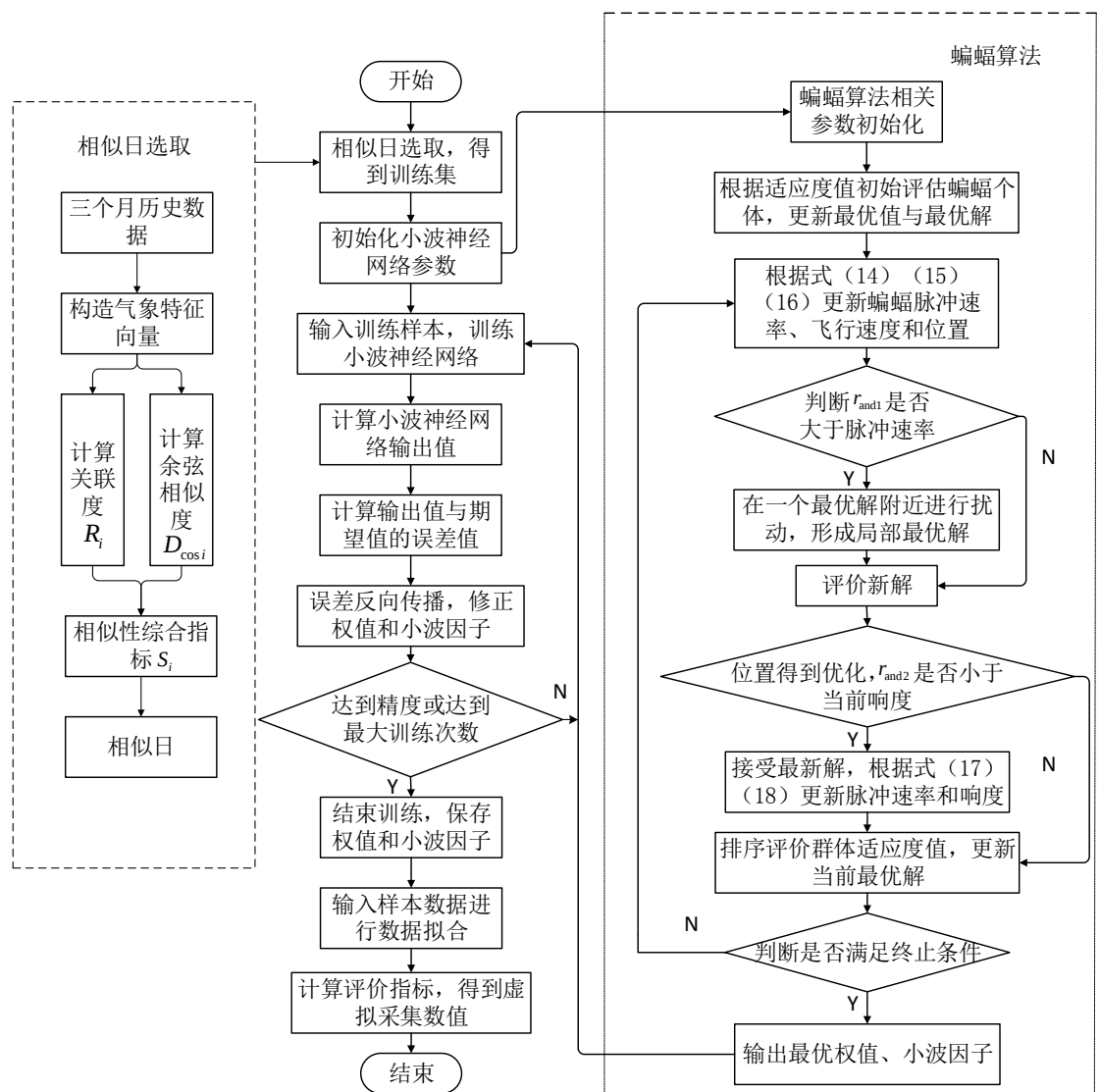


图 A1 虚拟采集模型示意图

Fig.A1 Schematic diagram of virtual acquisition model

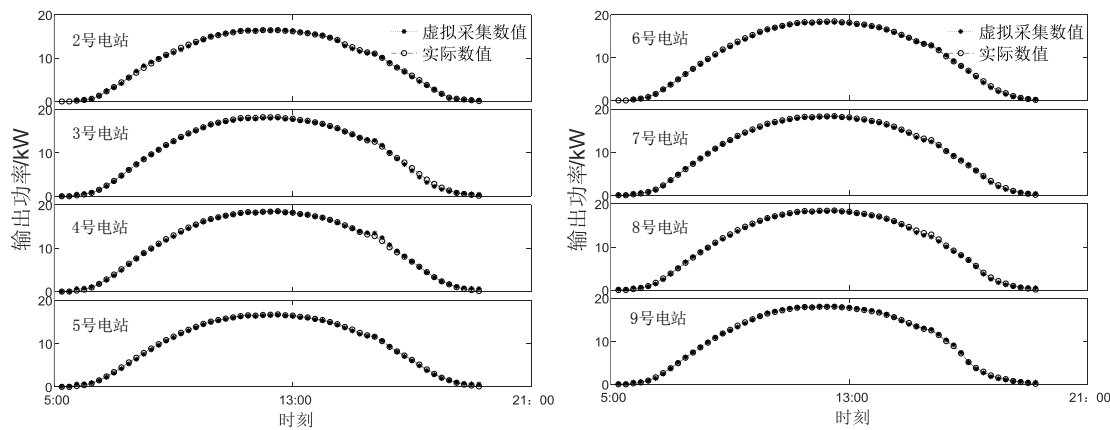


图 A2 平稳天气各电站虚拟采集数据

Fig.A2 Collected virtual data of each power station in stable weather

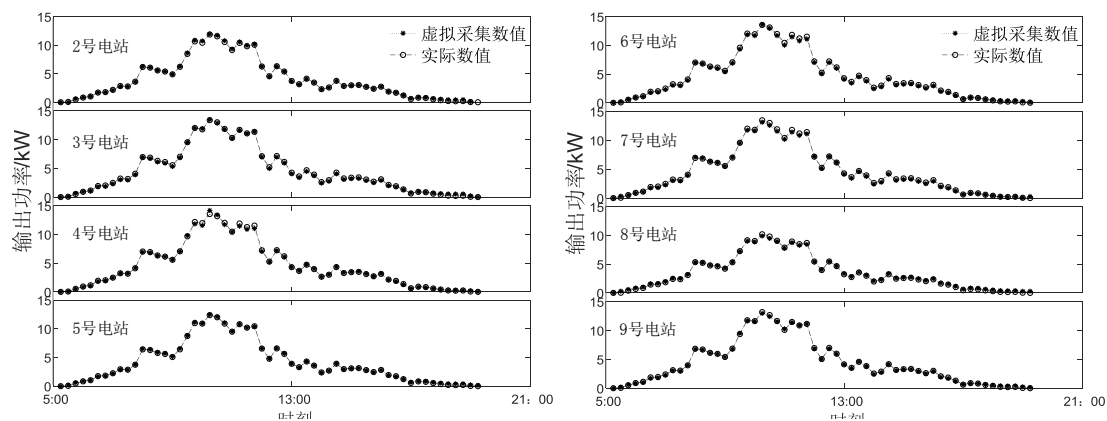


图 A3 突变天气各电站虚拟采集数据

Fig.A3 Collected virtual data of each power station in abrupt weather