

基于余弦相关性分析的光伏集群出力预测

吉畅¹, 唐辰旭²

(1. 六盘水师范学院, 贵州 六盘水 553000; 2. 国网四川省电力公司德阳供电公司, 四川 德阳 618000)

摘要:光伏发电具有波动性和间歇性, 分布式光伏发电的大规模并网, 对电网的调度、电能质量及电网的稳定性运行造成很大影响。光伏出力预测在有效解决上述问题的同时为电网调控提供了有力的支撑。本文建立基于相关性分析的单场电站组合预测模型, 利用选取典型场站、更新相关场站的相关系数和权重因子对光伏集群进行处理预测, 实践证明本文模型能得到比较准确的光伏集群预测结果。

关键词:光伏集群; 典型场站; 相关性分析; 预测

中图分类号: TK511; TM359.5; TM615

文献标识码: A

文章编号: 1001-2184(2022)05-0008-05

Photovoltaic Cluster Output Forecasting Based on Cosine Correlation Analysis

JI Chang¹, TANG Chenxu²

(1. Liupanshui Normal University, Liupanshui Guizhou 553000;

2. State Grid Sichuan Province Power Co. Ltd., Deyang Power Supply Branch, Deyang Sichuan 618000)

Abstract: Photovoltaic (PV) power generation is volatile and intermittent. With the large-scale grid-connected distributed PV power, power grid scheduling, power quality and stability of the operation will suffer a great impact. PV output forecasting can effectively solve these problems and provide a strong support for power grid regulation. In this paper, a single-field power plant combination forecasting model based on correlation analysis is established, and the correlation coefficient and weight factor of selected typical field stations and updated relevant field stations are utilized to deal with and forecast PV clusters. The practice shows that this model can be used to get more accurate PV cluster prediction results.

Key words: PV cluster; typical field station; correlation analysis; prediction

0 引言

随着经济发展, 环境问题越发受到重视, 早日实现“碳中和、碳达峰”成为全民共同愿景, 清洁能源得到了快速发展。但光伏发电的随机性和波动性给大规模光伏集群的并网带来困难。通过预测光伏集群的出力、掌握出力规律对平抑出力的波动、减小并网难度具有一定意义。

随着光伏电站的大量投运, 多个中小型光伏电站组合形成了光伏集群。现有文献主要针对单个光伏电站进行出力预测, 对于集群出力预测主要采用单个场站直接叠加、外推、统计升尺度、动态分群预测等方式进行^[1]。但单场预测叠加忽略了场站之间的相关性, 外推法通过子集预测进行总预测叠加, 同样忽略了子集之间的相关性, 且计

算过程输入特征的冗余度高; 统计升尺度则通过先划分子集, 再通过代表电站/典型场站的选取进行典型场站预测, 最后进行集群预测^[2]。动态分群预测与统计升尺度一致, 但在选取典型场站时利用了动态分群策略, 在预测过程中考虑权重矩阵进行处理预测, 其结果更加精确, 但动态分群在选取典型场站时数据量大, 选取时间长^[3]。

本文在考虑各场站出力的影响因素相关性以及与场站相关性的基础上, 通过典型场站的选取和权重优化进行光伏集群出力预测, 利用实际数据进行验证, 结果说明本文所用方法预测精确度较高, 预测结果良好。

1 光伏集群出力相关性分析及典型场站的选取

光伏集群由多个场站级联而成, 具有间歇性、波动性等特征, 其中场站之间出力存在明显相关性, 采用相关性分析不仅能够反应集群范围内电

收稿日期: 2022-06-17
基金项目: 六盘水师范学院科学研究计划项目
(编号: LPSSYZK202006)

站的出力情况还能有效避免部分电站数据缺失造成的预测准确率低的问题^[4-6]。

1.1 光伏集群出力相关性

一般而言,光伏出力包含了空间相关性和时间相关性。空间相关性是指各个场站之间地理位置的选择、场站大小等空间角度之间的相关性,空间相关性一般与时序无关。时间相关性是时刻对出力的影响,通常用自相关函数和偏相关函数来表示同一时刻的前后关联度,通常选址完成后,时间相关性与光伏集群出力最为密切。

余弦相似度不仅可以描述非线性变量间的相关性,还可以比较精准地反应变化趋势,是光伏集群出力研究的热门,表示方式如下:

$$\gamma = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (1)$$

式中 γ 为余弦相似度系数,其值大小为 $[-1, 1]$; x_i 、 y_i 为同一集群内两电站 i 时刻出力值; n 为总测量值个数。

1.1.1 辐照相关性

光伏电站能接受的辐照度是决定光伏电站总发电量的关键因素,太阳辐照主要分成直接辐照强度和散射辐照强度两方面,光伏电站的出力往往和这两方面都相关,因此,本文将考虑总辐照量。图1是2020年1月~2020年7月国家电投集团贵州某电站总辐照量和出力功率散点图。

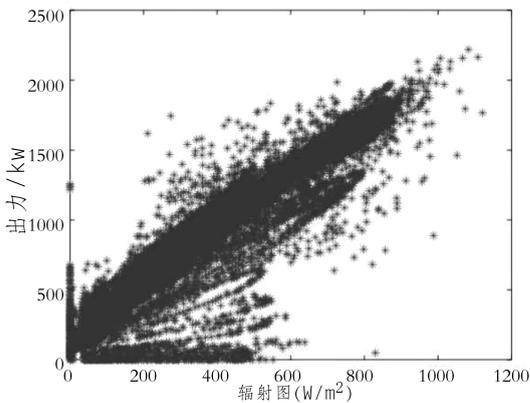


图1 某电站总辐照量和出力功率散点图

整理并计算得到余弦相似度为0.719 2,可以看出总辐照量和光伏出力之间存在较大的自相关性,且并不是严格的线性关系。记此时各场站所得相似度为 α_i 。

1.1.2 环境温度相关性

环境温度也是影响光伏出力的关键因素之一,在辐射量一定时,将温度维持在合适的温度值时,光伏组件的出力能达到峰值。图2是2020年1月~2020年7月同一电站件的环境温度与光伏出力散点图。

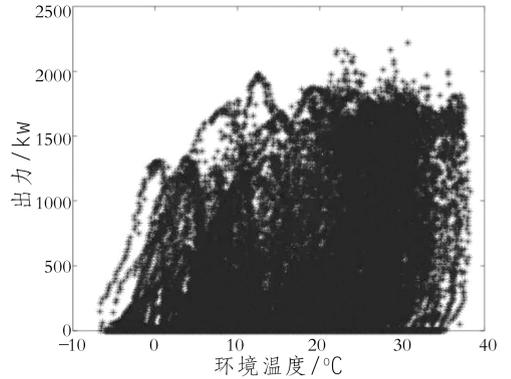


图2 环境温度与光伏出力散点图

整理并计算得到此时余弦相似度为0.764 1,可以看出温度与电站出力也有较为明显的相关性。

综上所述,光伏电站出力受太阳辐射和温度共同影响,其中太阳辐射影响更大。光伏组件的出力达到最大值是太阳辐射和环境温度的共同作用,因此,在进行光伏出力预测时需引入余弦相似度以提高模型的预测精度。记此时各场站所得温度引起的余弦相似度为 β_i 。

1.1.3 相邻电站出力相关性

光伏集群是多个光伏电站级联而成的大型光伏发电站,场站间的空间位置决定了不同场站的辐射量和温度变化影响。但相邻电站间往往具有相似的气象条件,其出力往往具有相同的趋势,相邻电站同一时刻光伏出力见图3:

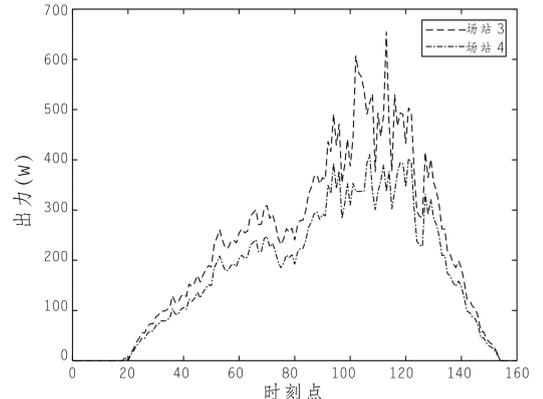


图3 相邻电站同一时刻光伏出力图

对于山区电站,相邻电站间的出力存在很小的时延,时延的存在使出力变化趋势不一致。为了简化预测量,提高预测精度,我们根据历史时刻对相邻电站出力的时延进行补偿,如下图所示,此时计算求得电站间最大余弦相似度为 0.795 1,可以看出此时的出力变化量一致,说明此时两个电站的出力存在相关性,在已知某一电站的出力情况下,可通过这个相关性预测另一电站出力。记此时各场站所得相似度为 δ_i 。

1.1.4 非相邻电站出力相关性

对于较远的同一集群电站而言,选取集群两侧电站观测同一时刻的出力情况,非相邻电站同一时刻光伏出力见图 4:

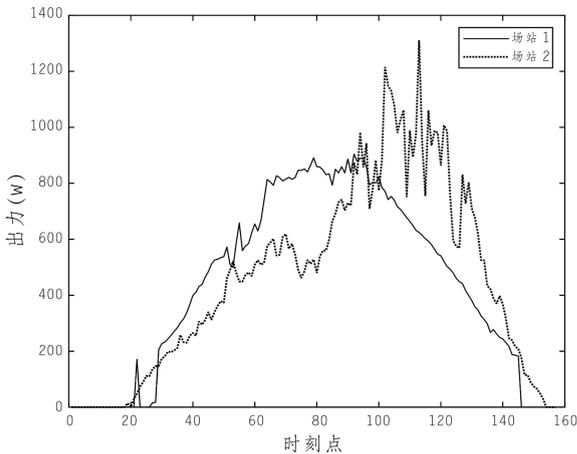


图 4 非相邻电站同一时刻光伏出力图

可以看出两个场站之间出力变化趋势有一定相似性,也存在一定的时延,但与相邻电站不同的是该站之间变化趋势不一致,计算可得最大余弦相似度为 0.911 3,可以看出其相关性更弱。记此时各场站所得相似度为 ϵ_i 。

1.1.5 单一场站对集群出力的相关系数

由于各个场站自身位置、气候影响和额定功率不尽相同,对集群出力预测时需要归一化处理才能避免超功率等问题,本文选择同一段时间内各场站的出力与集群总出力的比值作为此时的相关系数。

$$\phi = \frac{\sum_{i=1}^n P_{Ri}}{\sum_{i=1}^n P_{Ci}} \quad (2)$$

1.1.6 余弦相似度能够有效评价电站的出力相关性

在进行光伏集群出力预测时输入量至少有辐射、温度和不同场站位置等,因此,在预测时引入

余弦相似度可以极大提高预测精确度。

根据 1.1.1~1.1.5 可构建出力相关系数矩阵如下:

$$H = [\alpha, \beta, \delta, \epsilon, \phi] \quad (3)$$

1.2 典型场站的选取

对于光伏集群来说,直接对总体进行预测精度不高,主要体现在:

(1) 集群范围广,气候条件不均,按集群进行预测时无法得到准确的气象输入数据。而采用单个电站作为预测模型则气候条件更为精准。

(2) 将其作为单个模型的预测是,其灵活性、精度都不如多个场站的组合模型。

因此,本文采用选取典型场站,结合相关性系数和场站间的出力权重值对光伏集群的总出力区间进行预测。基本思路为典型场站→典型场站出力预测→利用相关性系数更新组合模型→计算集群出力→误差分析。

典型场站的选取原则主要包括以下两点:

(1) 数据不存在缺失;

(2) 典型场站是所求光伏集群中的相关性最高的场站。

本文通过滤波式特征选择,由于需要典型场站前期预测结果好,所以,需要最大化出力与输入变量之间的相关性,即所求电站预测精度和数据健康程度高;又由于典型场站需要与各场站之间都有相关性,因此,最小化各场站及场站之间出力的相关性,可以得到最小的场站之间相关性,减少了数据的冗余。

本文利用最大化特征集和最小化特征集作为约束条件,通过搜索形成最优集。具体为:假设一光伏集群共有 m 种特征,其中某一场站拥有 n 种特征,则一定有 $n < m$;进而假设 P 为某一场站特征集,则 P 为集群 F 的子集;令 d 为目标变量, v_i 为子集 P 的第 i 特征与目标变量之间的相关信息,则得:

$$\begin{cases} \max A(P, d) \\ A = \frac{1}{n} \sum_{i=1}^n f(v_i, d) \end{cases} \quad (4)$$

此时按照最大相关性原则选取得到的子集 P ,但需剔除其中的内部冗余信息。故有:

$$\begin{cases} \min B(P) \\ B = \frac{1}{C_n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n f(v_i, v_j) \end{cases} \quad (5)$$

式中 $f(v_i, v_i)$ 是子集 P 的特征向量之间的相关信息。因此,可以综合得到典型场站的选取为:

$$\begin{cases} \max \Psi(A, B) \\ \Psi(A, B) = A - B \end{cases} \quad (6)$$

值得注意的是按本文选择得到的典型场站可能不止一个。

2 基于贝叶斯网络的光伏集群出力预测模型

贝叶斯预测是一种以贝叶斯统计为基础的以动态模型为研究对象的时间序列预测方法。贝叶斯网络训练的结果可以得到权值的概率分布,因而贝叶斯网络常作为概率性网络:在给定的输入向量下,网络输出采用分布函数,也就是输出分布的概率区间。在需要单值输出时可以采用均值、期望值等,此时概率区间可以提供分布的可靠程度。

2.1 基于贝叶斯网络的单场电站模型

假设输入为光伏电站的历史出力值和对应时间段的气象因子,出力采集的时间间隔为 15 min。为了更好地处理输入输出关系,我们将以上三个值进行归一化处理。

此外需要设置贝叶斯网络的适应性函数:

$$\min \begin{cases} F1 = \text{MSE} = \frac{1}{L} \sum_{i=1}^L (P_{y(i)} - P_{act(i)}) \\ F2 = \text{std}(P_{y(i)} - P_{act(i)}) \end{cases} \quad (7)$$

式中 L 为测试数据集的个数, $P_{y(i)}$ 和 $P_{act(i)}$ 分别为预测值和实际值。据此可得到典型场站的光伏出力概率区间。

2.2 光伏集群预测算法

光伏集群的出力值是基于典型场站的出力预测值和各参数之间的相关性系数计算得到的。即在 2.1 中求出典型场站的出力值为基准,首先对代表场站与其余场站之间的相关系数进行计算并作为优化因子和判断值,求出集群内其余场站的出力值。再通过不同场站与光伏集群的特征比值作为场站在集群中的权重;则可根据所选场站预测值、相关系数和权重完成整体集群出力预测。其中预测值和相关系数在 1.1 中已经详细讲述,权重因子可表达为:

$$\omega = \frac{E(v_i, A)}{\sum_{i=1}^k E(v_i, A)} \quad (8)$$

式中 k 为选取得到典型场站的总数, $E(v_i, A)$ 为第 i 个电站所含有的特征量互信息。可得到权

重矩阵 $W = [\omega_1, \omega_2, \omega_3 \cdots, \omega_k]$ 。

则集群总出力值为 $P_c H W$ 。

3 实验验证及误差分析

3.1 实验验证

本文采用贵州省威宁某集群的数据进行验证。采集 2020 年 12 月~2021 年 11 月的出力数据,采集时间间隔为 15 分钟,由于光伏电池发电的特殊性,本文每日数据仅取早上 5:30 至晚上 10:00 的数据进行计算。

由于该电站群全部属于同一公司,且内部数据采集装置一致,地理位置接近,建设时间一致,因此,可作为同一类集群处理。按 2.1 中典型场站的选取方案,可得到该集群中的典型场站 4 号电站的位置(图 5)。

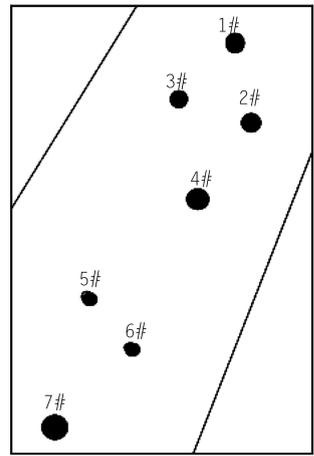


图 5 集群中的典型场站 4 号电站位置示意图

对选择的典型场站进行处理预测,按 75% : 25% 进行训练和测试,学习步长为 0.05,学习误差 0.01,训练函数采用 trainbr 函数,利用本文方法进行集群预测,此时可得集群总出力预测曲线(图 6)。

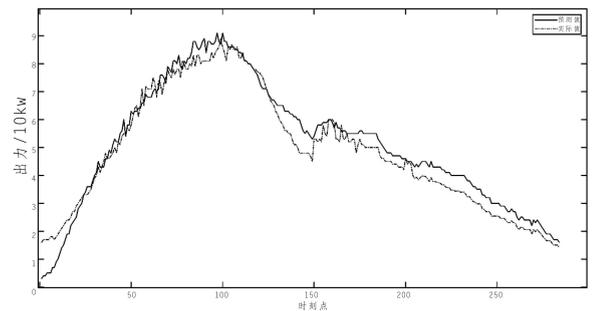


图 6 集群总出力预测曲线

3.2 误差分析

本文所用方法平均误差为 2.1%，误差来源主要有以下几方面：

(1) 输入量各物理量之间的相关性误差。即当输入量足够多时，空间相关性与时间相关性的输入量之间存在着非线性相关的关系，而本文采用算法中输入量与输入量之间认为是线性相关。

(2) 算法本身误差包含所选用网络误差和输入数据处理误差等多方面因素。

4 结 语

光伏集群出力预测是电网调度、稳定性分析等的基础支撑，常常采用直接叠加、外推和统计升尺度 and 动态分群预测等方式进行。由于光伏出力的随机性和波动性，预测难度较大，算法复杂，因此，本文在集群出力相关性分析的基础上，利用余弦相似度对影响场站出力的数据进行相关性计算，再利用集群内各场站的互信息比值可得权重，最后利用贝叶斯网络进行预测，得到集群总出力值。通过实际数据进行验证，采用本文算法进行集群出力预测结果的平均误差极小，后续可通过改进算法和优化相关系数进行优化。

参考文献：

(上接第 7 页)

极化隔离度 ≥ 32 dB；

上星性能测试方面，专网测试通过亚洲 7 号卫星进行了 4 h 的业务传输试验，通过对测试数据进行分析，完成专网业务 2M 带宽的传输要求，实现华能四川水电有限公司和各子公司的话音视频数据业务的传输^[5]，公网方面通过亚太 6D 完成同时 4 路音视频传输测试。

3 结 语

通过测试结果表明，双模卫星便携终端达到预期的设计要求，对于具有 VSAT 专用通信网的应急通信部门或者企业，同时又具有高速数据传输的业务需求，系统建成后具备公网通信功能，两种网络相互补充提高其应急支撑能力有显著的指导作用。终端性能突出，具有广覆盖、速度快、低成本、短时间、高效率等优势，该轻量化终端设备敏捷高效、韧性抗毁、随遇接入的优良特点，破解了无网断电等极端条件下的应急通信保障问题，一体化设计使该系统可以在几分钟内完全组

- [1] 孟安波, 陈嘉铭, 黎湛联, 等. 基于相似日理论和 CSO-WG-PR 的短期光伏发电功率预测[J]. 高电压技术, 2021(4): 1176-1184.
- [2] 李刚, 刘佳林, 王腾飞, 等. 基于相似日理论和 IPSO-Elman 模型的短期光伏发电功率预测[J]. 测控技术, 2020(2): 91-97+131.
- [3] 杨子民, 彭小圣. 基于集群动态划分与 BLSTM 深度学习的风电集群短期功率预测[J]. 高电压技术, 2021(4): 1195-1203.
- [4] 丁明, 方慧, 毕锐, 等. 基于集群划分的配电网分布式光伏与储能选址定容规划[J]. 中国电机工程学报, 2019(8): 2187-2201+2.
- [5] 石建平, 李培生, 刘国平, 等. 基于改进粒子群优化算法的混沌系统参数估计[J]. 华中科技大学学报(自然科学版), 2018(9): 70-76.
- [6] 侯松宝, 王侃宏, 石凯波, 等. 基于相似日和主成分分析的光伏发电系统短期出力预测[J]. 可再生能源, 2018(1): 55-59.

作者简介：

吉 畅(1991-), 女, 贵州毕节人, 工程师, 工学硕士, 主要从事电力系统参数辨识、故障诊断等；

唐辰旭(1988-), 男, 四川成都人, 工程师, 工学硕士, 主要从事电力系统保护与控制、配网调控、配网运营管控等。

(责任编辑: 卓政昌)

装, 单设备具备多星、多主站、多系统、公网结合的通信保障能力。根据现场情况灵活选公网或者专网有效支撑了线上远程调度与线下应急行动相结合的方式, 逐步完善了跨领域、跨层级、跨地域的应急联动机制, 推动了我国通信模式的创新。

参考文献：

- [1] 周晓斌. 应急通信 VSAT 便携站升级改造[J]. 信息通信期刊, 2009(6): 60-64.
- [2] 颜万生, 方芳, 张京, 等. 碳纤维复合材料天线反射面研制[J]. 现代雷达, 1999(6): 104-106.
- [3] 房海军, 涂彬. 卫星天线研制中反射面型面精度和尺寸稳定性的研究[J]. 航天返回与遥感, 2007(3): 68-71.
- [4] 国内卫星通信小型地球站(VSAT)通信系统工程设计规范[S]. YD/T 5028-2018. 0.
- [5] 吕龙, 曹伟. SKYWAN VSAT 卫星通信网络在水电行业的应用[J]. 数字通信世界刊, 2018(11): 159-163.

作者简介：

廖 川(1982-), 男, 四川成都人, 主管, 学士, 从事水电管理工作；

叶家帅(1996-), 男, 四川成都人, 通信系统工程师, 学士, 从事卫星通信研发工作。

(责任编辑: 吴永红)