

# 光伏超短期发电量预测算法研究及应用

熊敏, 吴文宝

(中国电建江西省电力建设有限公司, 江西 南昌 330001)

**摘要:**针对光伏场站运营中须对超短期即未来15分钟的发电功率与发电量提供预测的要求,结合国内当前光伏场站的实际数据条件,建立新型的自主算法并设计实施规程。首先,依据光伏发电机理模型中对关键变量以及变量间基本关系的阐释,针对三十余家并网发电光伏场站的实际运行数据,就关键变量的可获得性、可用性和使用方式进行系统的梳理;其次,提出了一种由随机相点滚动向量算法及自主学习规程两个部分构成的创新性预测算法;最后,使用场站运行数据从不同的角度对模型的预测表现进行了验证,展示了验证结果,并进而详述了实施策略。

**关键词:**超短期发电量预测;机理模型;实证分析;随机相点滚动向量模型;机器学习规程

**中图分类号:** TM 615 **文献标志码:** A **文章编号:** 1006-348X(2024)06-0052-06

## 0 引言

为新能源场站赋予发电量预测能力,并不断地提高预测精度,是促使新能源场站实现数字化经营与管理的重要环节。世界行业规范要求,从事并网发电的新能源场站必须向电力调度企业提供不同期长的发电量预测,包括未来15分钟的超短期、次日、短期以及中长期预测。这些预测是电网调度系统实现在不同期长内的产需平衡、从事跨地区电量调配、进行削峰填谷、完成发电额度分配的重要依据<sup>[1-2]</sup>。

在上述不同期长的发电量预测中,超短预测具有着独特的挑战性。首先,天气预报服务商无法对辐照度提供15分钟级时间精度的预测;即便假设能够获得高时间精度的局域平均辐照度预测,预测云块扫过特定光伏阵列的阴影也完全不现实。基于卫星动态云图以及AI图像识别技术的方法获得了一定的理论探讨与实验尝试,但结果证明这一方式距离实际应用尚遥遥无期,究其原因一是成本过高,二是日光对不同云层的穿透力无法预测<sup>[3-4]</sup>。

为此,文中的核心内容是阐释一个完全基于场站数据的超短期发电量预测算法,该方法由随机相点滚动向量算法及自主学习规程两个部分构成,就方法论

和算法的角度而言,这是一种创新性的预测方法。

## 1 基于场站数据的机理模型验证以及发电量预测建模综述

机理模型是通过光电转换的物理学原理而构建的发电功率与生电因素之间的关系,因而是指一切数据建模的理论基础。使用场站数据,验证机理模型所揭示的主要关系,将电学理论、场站运行数据和应用建模三者贯穿起来,从而在最大程度上实践建模过程的完整性。

### 1.1 基于场站数据的机理模型验证

发电量预测模型通常被分为三个大类,即机理模型、统计类模型和机器学习类模型。其中机理模型基于原理,探究的是光电转换的机制,即什么是决定发电功率高低的主导因素,又有哪些是影响因素。使用场站数据对机理模型导出的变量关系进行实证分析,是为了用场站运行数据来对应和印证理论,从而为建模设计,包括本文所聚焦的超短期发电量预测算法和自主学习规程,提供实证依据。

#### 1.1.1 发电功率与辐照度之间的实证关系

根据光伏电池的生电原理,发电功率与辐照强度之间为准线性关系,即对于既定的光伏电池(既定的生产工艺和服役时长),在给定的辐射倾角、背板温

收稿日期:2024-08-23

基金项目:中国电建集团科技项目(DJ-ZDZX-2021-1)。

作者简介:熊敏(1976),男,大学本科,高级工程师,研究方向为新能源电站运维技术开发及电厂自动化。

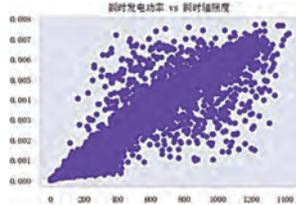
度、环境湿度以及光源的频谱特征下,光电池的最大可输出功率随着辐照度的增强以近似常数的比例增加。这一关系在实验室环境下容易得到验证,因为无论是辐照度,还是温度与湿度等各个因素,在实验室中都易于控制和测量。

使用场站运行数据进行实证分析的意义在于为建模设计中的具体选项提供判据。在实证分析中得到的结论,对于超短期预测算法中的变量选择和使用方式,均有着重要的指导意义。

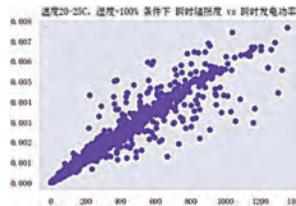
本研究使用的全样本是并网发电场站集中式逆变器在两年内积累的SCADA数据。以下是对分析中的主要观察与结论的总结,图1提供了相应的实证依据。

对于超短期预测,如果单从分析的角度出发,来假设可以获得未来15分钟辐照度的精准预测值,并假设没有发生限量发电、故障停机等外生干扰的情形,那么使用辐照度的预测值,并使用瞬时背板温度和湿度进行样本分箱,就可以得到十分精准的发电功率甚至是发电量的预测值。

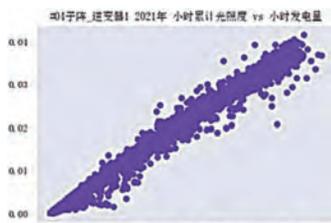
然而现实中很难从技术上准确预测未来15分钟内的辐照度。正因如此,文中所提供的算法方案,是使用建立在大数据上的算法,找到通过历史数据而不是预测数据,进行未来推断的途径。



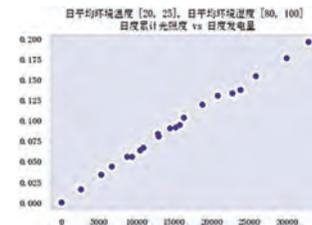
(a) 瞬时发电功率与瞬时辐照度的关系(瞬时值-全部样本)



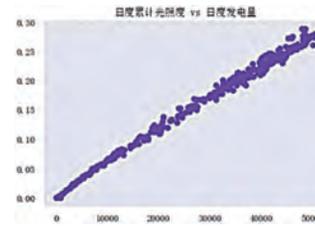
(b) 瞬时发电功率与瞬时辐照度的关系(瞬时值-样本分箱后)



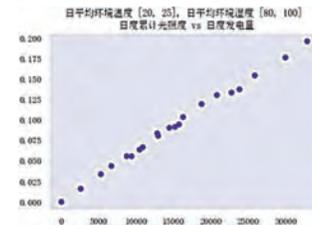
(c) 小时发电量和小时累计光照度的关系  
(使用年度历史数据生成)



(d) 小时发电量和小时累计光照度的关系  
(使用环境温度和湿度平均值区段进行分箱)



(e) 日度发电量和日度累计光照度的关系(年度历史数据)



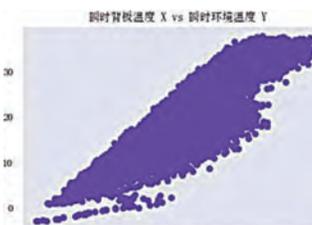
(f) 日度发电量和日度累计光照度的关系  
(使用环境温度和湿度平均值区段进行分箱)

图1 发电功率与辐照度之间的实证关系

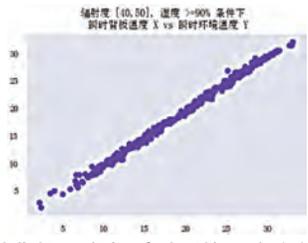
### 1.1.2 发电功率与背板温度之间的实证关系

依据光电效应的原理,对于固定的辐照度,光电池的光电转换率随着背板温度的升高而降低。光伏阵列的背板温度是场站发电过程中的一个实测变量,用来衡量阵列的实际工作温度,其主要影响因素是辐照度、环境温度以及风速(降温作用)。对于无法提供背板温度或数据质量可疑的场站而言,以环境温度替代是一个自然的选项。

通过图2可以观察到:由于辐照度、环境温度以及风速会同时影响背板温度,因此当使用全部样本数据,考察背板温度与环境温度之间的正向相关关系时,呈现的准线性关系并不精准;如果同样采用分箱的方式,用辐照度区间和环境湿度区间,对样本进行分箱处理,则背板温度和环境温度之间能呈现精准的线性关系。



(a) 瞬时背板温度与瞬时环境温度之间的关系(2021全年数据)



(b) 瞬时背板温度与瞬时环境温度之间的关系  
(使用辐照度和环境湿度区段进行分箱)

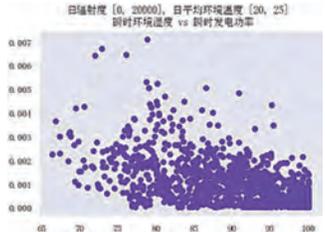
图2 瞬时背板温度与环境温度之间的实证关系

鉴于环境温度在样本分箱的条件下,能够精确地映射背板温度,因此,必要时可以做为背板温度的良好替代变量。

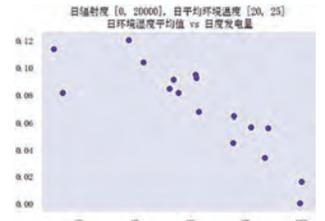
1.1.3 发电功率与环境湿度之间的实证关系

同样是基于光电转换的机理,在固定的辐照度下,光电池的光电转换率随着环境湿度的增加而降低。通过对图3中实证数据的观察,发现发电功率对环境湿度有着明显的响应惰性,即瞬时发电功率同瞬时湿度之间虽然没有呈现出统计上的关联关系,但日度发电量与每日平均环境湿度之间的线性关系则十分明显。究其原因,是环境湿度对发电功率的响应具有时间惰性,因为湿度的变化要“渗入”到阵列和导线内并导致漏电流以及相应发电损失的变化,需要一定的时间。

因此,对于在发电功率预测建模中如何使用环境湿度要取决于建模场景:对于期长在日度以上的建模,引入湿度作为解释变量的意义要远高于超短期预测这一场景,虽然对于超短期预测而言,使用环境湿度对样本进行分箱,有助于预测精度的提高。



(a) 瞬时发电功率和瞬时环境湿度的关系  
(使用辐照度和环境湿度区段进行分箱)



(b) 日度发电量和日环境湿度平均值的关系  
(使用辐照度和环境湿度区段进行分箱)

图3 发电功率与湿度的实证关系

1.2 对发电预测统计模型和机器学习模型的应用总结

目前,国内外针对超短期预测的研究焦点是寻求新的技术手段,以提高对超短期辐照度的预测精度,而基本思路是使用卫星云图或对天空区域连续拍照,从而计算云块的续行轨迹,或探寻改善辐照度预测的算法模型;而预测云块扫过地面上既定光伏阵列的阴影的措施,成本昂贵且缺乏实用精度,因而很难进入生产过程。

因此,文中的重点是针对超短期预测,来探讨和验证一种新型的算法,即随机相点滚动向量模型及配套的自主学习规程,以弥补现有解决方案中的欠缺,这一方法的本质是使用在预测时点上的发电功率,对预测期内发电功率变化方向和幅度推断,来预测15分钟后的发电功率,因此不以预测辐照度为先决条件。

2 用于超短期发电预测的随机相点滚动向量算法

建立基于场站生产数据的超短期发电量预测模型,首先需要考察场站数据字段的可获得性和可用性,并进而通过对可用数据的画像,建立数据现象与设备运行之间的对应关系。

2.1 场站发电数据字段的获取及其画像分析

该研究选择的是一个中等规模的综合型水上光伏发电场,装备有20台集中式逆变器和71台分布式逆变器,数据历史覆盖了两年零7个月的SCADA数据。

2.1.1 对实测数据的画像分析

与实验室数据不同的是,生产数据是发电设备在运行中产生的,因而各个变量的数值在物理关联中生成而不能单独可控。同时数据中也夹杂着如故障停机、限量发电、极端天气影响、设备计划维护停产等实际因素。因此,建立数据与设备实体运行之间的对应关系是生产场景建模中的一项基础工作。

2.1.2 实测数据与设备实体运行情形之间的对应关系

作为案例,在图4中显示了为期8天的具有代表性的时间序列数据,变量为单台逆变器交流输出端的发电功率与辐照度,数据采样间隔为10分钟/次。例子中连续8天的数据反映了整日晴朗无云、阴晴相

间、整日低辐照等多种情形,从而代表了超短期预测的实际场景。由图4可见,瞬时发电功率随着瞬时辐照度的变化同步起落,再次印证了两者之间稳定的正向相关关系。

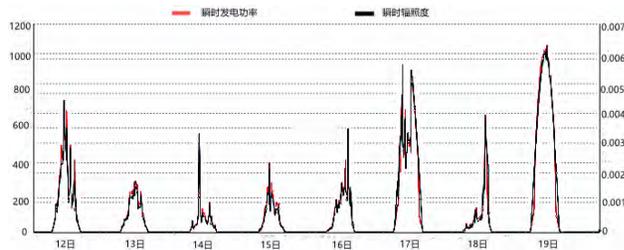


图4 单台逆变器交流输出端瞬时发电功率与瞬时辐照度的关系

### 2.1.3 对数据的甄选、截取与清洗

对原始数据进行甄选是为了避免使用失真的数据。传感器失灵和通讯故障是数据失真的主要原因,也是必然要发生的情况。因此任何负责任的数据应用方案都须建立对失真数据进行判别的逻辑,在生产过程中及时告警,并停止在失真出现之前的操作。

数据截取的目的是排除数据中与应用目的无关的部分。例如,对于发电量预测建模,在生产发电时段以外的数据,以及在生产出现异常时的数据,都应被排除在建模样本之外;对于预测应用,计划停机维护和应急维修时段也应排除在外。

数据清洗是对已知数据缺陷,通过编程做出的指定修复动作。例如,如果在一个连续的运行过程中突然出现了数据的异样跳跃和空挡,那么使用周边正常数据的均值进行替代或填充就是十分常用的清洗方式。对历史运行数据进行细致的画像分析是建立甄选、截取和清洗逻辑的基础。

## 2.2 随机相点滚动向量算法的建立

### 2.2.1 基本假设与算法逻辑

#### 1) 随机相点滚动向量模型中的基本假设

在预测时点 $t_0$ 上已知的历史发电功率值,尤其是在 $t_0$ 时点近旁的发电功率数值,可以用来作为未来超短期(如15分钟)发电功率预测值的锚点。

以上述锚点作为起点值,发电功率在超短期内变动的方向与幅度,在没有突发事件影响的情况下,可以通过对历史大数据的建模进行推断。

进行上述推断的途径是使用在 $t_0$ 时点近旁的发电功率、辐照度以及天气变量构建状态蜂房。状态蜂

房是一种分箱方式;在进行了这种分箱以后,随机相点滚动向量算法用来对每个蜂房(分箱)计算迁移向量,即发电功率在 $t_0$ 时点上的数值在随后的超短期将如何变化,包括变化的方向和大小。

发电功率在下一个超短期内的预测值,是发电功率在 $t_0$ 时点上的数值与迁移向量的代数和。在得到了发电功率预测以后,发电量预测即为功率曲线下的面积;对于未来15分钟的发电量预测,在数学上的计算为 $\int_{t_0}^{t_0+15} p(t) \times t_0$ 。

迁移向量的准确性和稳定性取决于样本规模的大小,以及在状态蜂房设计中选择的分类变量数目与时间颗粒度。

#### 2) 状态蜂房的概念与迁移向量的算法逻辑

如果用 $t_0$ 表示预测起始时点,那么对于期长为15分钟的超短期预测,其预测结束时点可表达为 $t_1 = t_0 + 15$ 。

状态蜂房由许多蜂房单元组成,因此要理解状态蜂房的概念,便捷的方式是先看一个蜂房单元的例子。在表1中显示的是对一个蜂房单元的定义。该单元是由五个维度界定的,分别为在 $t_0$ 时点上的发电功率、辐照度、背板温度、湿度和日内时间点;右侧一列是五个维度分别的区间范围。假设对某个逆变器有两年的历史数据,那么对于以15分钟为间隔的数据采集频度,就会得到约23 62360(2年 $\times$ 365天 $\times$ 8发电小时 $\times$ 4个15分钟)个数据点。任何数据点只要落入表1所定义五维空间内,它们就同属于表1所定义的蜂房单元。

一个完整的状态蜂房由许多蜂房单元构成。各个单元具有唯一性,即每个数据点只能落在一个单元中;所有单元作为一个整体又具有完整性,即每个数据点都必然能落在一个单元中。

由于各个变量之间相互关联,因此如果将所有单元聚集在一起画出立体图型,那么其形状将类似于蜂房而不是直边图形,这就是蜂房分箱这一称谓的来源。

继该例假设有233 60个数据点,每个数据点落在哪个蜂房单元中是随机的,对于落在如表1所定义的蜂房单元中的所有数据点,并具有相同的起始状态,而相点即为状态的同义语。至此,可做出一个关键的假设,即对于处在同样初始状态的数据点,其15分钟

以后的终止状态将具有相似性,这种相似性可以用一个迁徙向量来表示,其方向用来表达在 15 分钟内的 高低走势,其大小用来表达变动的距离。建模阶段的 任务是通过历史数据算出这些迁徙向量,从而能在预 测中使用。

在预测阶段,预期时点是不断先前滚动的。如果 某天的第一个预测时点是 7:00,那么第二个预测时 点就是 7:15。7:15 同时也是第一个瞬时预测的终结 时点;在这个终结时点上,发电功率的预测值,就是初 始时点(7:00)的发电功率测量值(已知)与迁徙向量 的代数和。

表 1 状态蜂房中的一个单元(举例)

划分维度	蜂房区间
功率( $t_0$ )	(10,20)kW
辐照度( $t_0$ )	(5,10)J
背板温度( $t_0$ )	(9.5,13)°
湿度( $t_0$ )	(50,60)%
日内时间( $t_0$ )	(10,11)点

蜂房的颗粒度主要由样本规模、分类维度数量、 变量变动范围和迁徙向量的统计稳定性等因素决定。 蜂房设计中的最主要依据是使各个蜂房的迁徙向量 具有统计上的一致性和可解释性。表 2 归纳了在本 项目中使用的维度和区间长短,包括功率、背板温度、 湿度和日内小时。

上述蜂房设计下,得到的某个特定蜂房的迁移向 量,如表 2 所示。

表 2 迁徙向量计算结果举例

蜂房序号	1	2	3	4	...	N-2	N-1	N
迁徙向量值	1.873 187	1.407 375	31.249 455	1.242 453	...	0.905 472	0.903 309	0.889 532

### 2.2.2 有监督机器学习规程的建立

随机相点滚动向量算法的目的是在无法得到超 短期辐照度预测值的情况下,通过大数据建模的方式 进行超短期发电量预测,这一算法只是解决方案的一 个组成部分;完整的解决方案还包括了一个有监督的 机器学习过程,这个机器学习过程包括两个基本步骤 和一个循环流程,两个基本步骤包括对瞬时发电功率 趋势的统计外延和随机相点滚动向量算法的嵌套;循 环流程是一个优化过程。

#### 1) 建立统计外延算法池

预测的本质是选择最佳的适用方式对历史轨迹 进行外延。对时间序列数据而言,基本算法主要包括 滚动平均、回归、ARIMAX、LSTM 以及基于因素分解

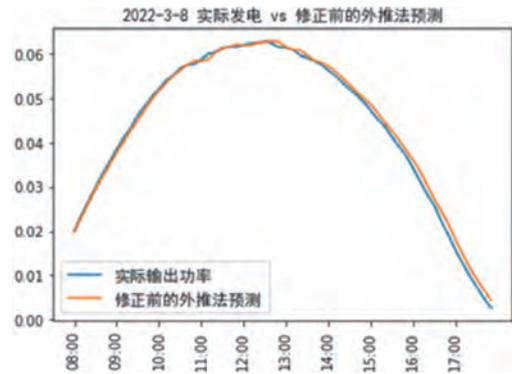
的动态模拟技术等。对于本项目而言,均以算法包的 形式存入了预测平台,以便算法程序的调用和根据数 据场景的自动优选。

具体到超短期发电功率预测,辐照度的场景可依 照云图的种类被归结为三种典型情况:情形 1:晴朗 或单一而持续的云况;情形 2:不同云况的渐变或边 界明显的切换;情形 3:无规则的云团、云块分布。

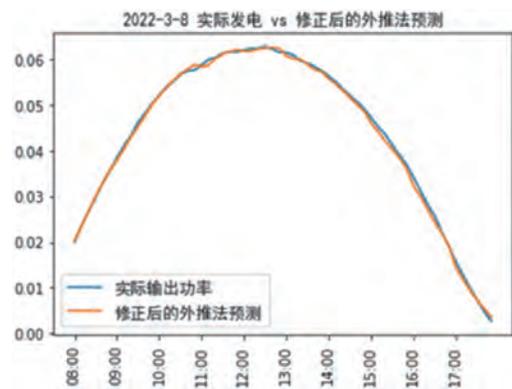
通过优选过程,对这三类云图进行适配,得到的 结论是两点滚动平均对于情形 1 和 2,而 ARIMA(1, 1, 1)对于情形 3 是最佳选择。

#### 2) 随机相点滚动向量算法的嵌套

嵌套的作用是使用随机相点滚动向量算法,对第 一步得到的外延预测值进行修正,消除系统性误差, 最终得到无偏差的预测结果。如图 5 所示,随机相点 滚动向量算法能够根据历史数据,建立进行系统纠偏 的向量矩阵,对外延法得到的预测数据进行调整,形 成最终的预测数值。



(a) 单纯通过外延法得到的预测结果



(b) 嵌套随机相点滚动向量算法后得到的结果

图 5 两步套嵌算法原理示例

### 2.2.3 实操应用策略

超短期发电功率及发电量预测在实操应用中的 一个必要步骤是实现算法同天气场景的匹配。以下

是在应用中的实操方案。

**自动匹配:**将当日凌晨获得的天气预报及日内更新信息导入编程,实现自动切换。

**人工干预:**由操作人员通过天气预报与实况目测获得数据,通过屏幕按钮进行算法匹配。

平台的用户界面将同时显示实时辐照度、发电功率、不同算法平行给出的预测值曲线,以及当日的天气预报更新信息,为操作人员的干预提供依据;平台亦将输出实际发电与预测值,形成当日的记录。

针对实操应用所需回答的另一个问题是,在没有保留历史数据的情形下,算法体系需要多长时间来完成机器学习过程。对于天气状况较为单一的场站,一个月的时间是充分的;而对于天气场景多样复杂的场站,则需要三个月的时间。

### 3 结语

文中首先依据光伏发电的机理模型中对关键变量以及变量间基本关系的阐释,针对三十余家并网发电光伏场站的实际运行数据,就关键变量的可获得性、可用性和使用方式进行了系统的梳理,得出当前的光伏场站数据可以为各种期长的发电量预测提供数据基础的结论。

迄今,在世界范围内针对超短期光伏发电功率与发电量预测的文献研讨和实用建模算法为数甚少,同时现有的探讨都是围绕着通过卫星云图和AI技术来

完善超短期辐照度预测这一思路展开的。本文的主要篇幅集中于对一个创新思路的展开,核心概念和内容包括状态蜂房分箱处理、随机滚送向量、通过趋势外延和随机滚动向量法嵌套,进行超短期预测,以及通过自学习算法对天气场景进行自动识别与分类等。本文使用场站运行数据,对模型的预测表现进行了系统的验证,并对落地实施策略进行了说明。

通过算法创新,来解决传统方式难于实现的超短期发电功率和发电量预测问题,并以更为精准和可控的方式落地到实际电力生产实践中,是该项研究所要达到的目的。为期数月的场站应用证明了文中所述方法的实用性,从而为普及性应用奠定了基础。

### 参考文献:

- [1] R.H.Inman, H.T.C.Pedro, C.F.M.Coimbra.Solar forecasting methods for renewable energy integration[J].Progress in Energy and Combustion Science,2013(6):535-576.
- [2] R.Perez, S.Kivalov, J.Schlemmer, et al.Validation of short and medium term operational solar radiation forecasts in the US[J].Solar Energy,2010(84):2161-2172.
- [3] L.Nonnenmacher, C.F.M.Coimbra, Streamline-based method for intra-day solar forecasting through remote sensing[J].Solar Energy,2014(108):447-459.
- [4] R.Marquez, H.T.C.Pedro, C.F.M.Coimbra.Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs[J].Solar Energy,2013(92):176-188.