

基于马尔可夫链的光伏电站发电功率超短期预测方法

文/范文华 鲍海铭 刘英健 侯杰栋 王 超
国家电投集团河北电力有限公司张家口分公司

提出基于马尔可夫链的光伏电站发电功率超短期预测方法，对预测数据进行预处理，为云层状态划分提供参考依据。根据云层状态划分结果，计算预测节点云层应用覆盖范围，并根据特定格式进行修改，完成结合预测模型中设计的预测方案，达到超短期的预测效果。

光伏发电是一种可再生能源，不消耗有限资源，且不会产生CO₂等温室气体排放，对环境的污染较小，有利于改善气候变化和空气质量，广泛应用于世界各地的能源转型和可持续发展中。近年来，我国可再生能源产业特别是光伏发电产业发展迅速。截至2019年末，我国可再生能源发电设备总产量达到 7.94×10^8 kW，占有所有电气设备的39.5%。其中光伏发电设备产量为 2.04×10^8 kW，年增长17.3%。在国家政策支持下，分布式光伏设备的产量增长尤为迅速，达到 6.234×10^7 kW，2019年同比增长24.2%，年产能增长12%。光伏发电对环境友好，在减少温室气体排放和改善空气质量方面起着重要作用。然而，由于光照、天气等因素的影响，光伏电站发电功率存在波动性。光伏电站发电功率的预测可帮助电网运营商和电力系统管理者更好地安排电力的调度，制定更加精确的调度策略，确保电力市场的平衡，提高电力系统的可靠性和稳定性。光伏发电功率的预测可分为超短期、短期和中长期等。其中，超短期预测主要关注未来几个小时的光伏发电产能，对调整日常电力计划和电力市场交易具有重要作用。有研究学者基于动态时间弯曲（DTW）和变分模态分解（VMD）的粒子群（PSO）优化的BP神经网络，设计了一种光伏

发电预测方法，根据DTW值确定对光伏发电功率影响较大的因素，利用变分模态分解降低数据的波动性，运用粒子群优化的BP神经网络实现光伏发电功率的短期预测。但在实际应用中，该方法易受外部因素的影响，预测精度控制能力较弱，可能增加预测误差，影响预测结果的准确性。还有研究学者提出基于时序动态回归的光伏发电功率超短期预测方法，建立回归模型，将历史数据和天气预报数据作为输入数据，再根据ARIMA模型预测回归残差序列，结合傅里叶谐波序列共同完成光伏发电功率短期预测。但在实际应用中，当产生异常电力时，预测误差的频率较高，自动定位程序相对独特，预测的结果缺少可靠性。

为了解决上述问题，本文基于马尔科夫链提出一种光伏电站发电功率超短期预测方法，并对设计方法进行研究和验证，以期降低预期事故的可能性。

基于马尔科夫链的光伏电站发电功率超短期预测方法设计

1. 预测数据预处理

通常，为了确保最终数据处理结果的可靠性，减少预测时间，提高整体预测能力，应使用相关设备和装置获取预测数据和信息。在这一过程中，可

以使用配置的数据收集单元开发自动收集程序，获得实时光伏数据，为预测工作提供参考依据。使用节点收集的数据是原始数据源，无法直接使用，需要修改数据集的格式或设计，减少数据处理错误。自动程序根据外部披露对数据进行分类，预期平均值为

$$H = (m+n)^2 \frac{m_i n^2}{v + \sum n_i - 1} \quad (1)$$

式中， H 为平均预测值； m 为异常预测的范围； n 为异常预测一致的范围； v 为数据预测点之间的距离；1为预测次数。

应根据所获得的实际估计平均值，确定相关预测范围，并在两个方向上建立所获得的异常数据，完成预测数据预处理。

2. 云层状态划分

云层状态划分是马尔科夫链的重要组成部分，在建立光伏发电状态转移矩阵时需要对云层状态进行划分。一般来说，光伏发电受云层直接影响，而太阳辐射强度影响云层。辐射遵循 β 概率的传播，但事实上，如果云层10点第一次挡住太阳，显然10点01分云层无法明显移动。因此， β 的分布函数没有暂时的特征，也不能描述云中时间变化对光伏特特的影响。为了探索云的变化过程，本文引入了云覆盖水平（CCL）指标对云的位置进行分类，计算式为

$$CCL(t) = 1 - \frac{HI(t)}{I_B(t)} \quad (2)$$

式中， $CCL(t)$ 为 t 时刻的云层覆盖水平； $I(t)$ 为 t 时刻的太阳辐射； $I_B(t)$ 为 t 时刻太阳辐射的值。参考值是全天变化的当天无云层覆盖的理想太阳辐射，最大值为1 000 W/m²。

初始马尔可夫链中 P 的固定转移矩阵对解决临时间题施加了限制。在CCL指标的基础上进行云层状态划分，应使用统计方法计算每15 min云位置 P 的转换矩阵。 $P_{ij}(t)$ 矩阵中的跃迁势 P 为

$$P_{ij}(t)E = \frac{CCLq_{ij}(t)}{\sum q_{ij}(t)} \quad (3)$$

式中， $\sum j=L1L4ij(t)$ 为 t 时刻云层状态发生变化的总次数； $P_{ij}(t)$ 为 t 时刻云层状态由 i 状态转移到 j 状态的次数， $ij \in [L1, L2, L3, L4]$ 。

云层状态分类见表1。CCL值越多，云就越多。

表1 云层状态分类

CLL等级	CLL范围
L1	0.0~0.25
L2	0.25~0.55
L3	0.55~0.85
L4	0.85~1.00

3. 基于马尔科夫链实现光伏电站发电功率超短期预测

传统的光伏发电生产预测模式和预测单元是单向的，预测范围仅限于外部因素和动态因素，预测结果不会被消除。建立周期性预测结构，并根据预测位置配置双向监测和预测单元。本研究结合马尔可夫链技术和超短期预测技术开发了一个自适应预测序列。首先，根据上述云层状态划分结果计算预测节点云层应用覆盖范围，表达式为

$$E = \left(0.5\delta^2 \frac{u}{hu - \gamma} \right) - U \sum h + e \quad (4)$$

式中， E 为源域应用覆盖范围的预测范围； δ 为检查预测距离； h 为预测源数据集； u 为预测序列的短预测标志偏差； e 为节点预测的数量。

根据预测节点的范围 γ ，自动控制程序将侧重于光伏设备的识别。需要注意的是，识别的位置与源域重复，应该建立一致的预测标准。此时，短期预测的序列用于评估周期性标准预测的能源效率，同时调整节点的位置，扩大实际预测的范围，并通过逐步发展周期性和动态的预测制度，及时发现异常情况。连同预测的目标，使最终得到的预测结果是准确和可靠的。具体为

$$Y = \omega - \int 0.5\eta \left(\rho\omega + \frac{1}{hQ} \right) \quad (5)$$

式中， Y 为周期性标准预测的效率； ω 为隐藏偏移预测区域； η 为发展的特定预期实际值； ρ 为重复预测的频率； Q 为预测的自动范围。

事实上，通过将模型划分为多层次和多目标的预测结构，可以使用预测序列分割控制场，提高模型预测的准确性，并评估预测的最短时间，计算式为

$$q = \psi - \frac{th(\psi^2 - 1)}{Lz \cdot 0.5t} v^2 \quad (6)$$

式中， q 为最短预报时间； ψ 为展开总体预报范围； l 为两极预报值； z 为预报单值频率； t 为以前的速率； ψ^2 为未来预报区域； v 为平均频率分量。

得到的最短预报时刻确定为预报标准时间，并加入预测模型。在实际操作过程中，应使用自动协议和程序进行控制。根据上述条件，建立一个预测矩阵，处理标准的预测用作自动预测的指南。首先提取预测数据，创建稳定的预测结构，并在预测目标的设计模型中识别提取特征，以识别特定的预测区域。本段中的可预见矩阵对应于可预见模型的处理阶段，计算欧几里得的相应距离，计算式为

$$Z_c = \arg \min \|qX - Z_j\|^2 \quad (7)$$

式中， Z_c 为欧几里得距离； X 为边界标签； Z_j 为更新后的权重。

合并以前获得欧几里得距离，并定义预测的某些点和矩阵中的适用位置。马尔可夫链包含两层随机过程，即马尔可夫链和一般的随机过程。马尔可夫链用于描述隐藏过程概率的变化、隐藏过程之间的关系及可能的观测值，使用整体统计过程进行描述。本文包括马尔可夫链中的经典训练问题和解码问题（也称为预测问题）。马尔可夫链用于分析不同阶段感知的超短期预测的影响，预测感知效果分析见表2。

表2 马尔可夫链预测感知效果分析

预测阶段	感知预测出完成率（%）	预测偏差
小幅功率预测	44.25	1.02
多维度功率预测	71.03	1.14
大功率预测	87.22	1.27

完成对马尔可夫链预测感知效果的分析研究，并以此为基础，结合预测模型中设计的预测方案，达到超短期的预测效果。与以往的标准进行比较，确保最终预测的效果，为未来的预测工作提供数据和信息支撑，并能继续改进预测方法。

实验论证

本实验主要应用于基于马尔可夫链的光伏电站发电功率超短期预测方法的效果分析和研究。为最终测试结果的真实性和可靠性，建立了传统方法1和传统方法2。以对比实验的方式进行测试，并对测试结果进行比较。需要创建一个实际的预测环境测试该方法的预测效果。

1. 实验准备

选取某地区2011年1月—2012年2月的负荷数据为研究对象。该地区的工业负荷约占总负荷的

72%，规律性强，预测容易。生活能耗约为14%，预测难度较大。在分析历史负荷曲线时，可使用数据库预测拐点。根据收集的基础数据和形式数据确定预测频率，设定预测周期，共4 h，即6个周期，以周期的形式进行预测分析，使用回归模型和转移预测模型评估超短期电力预测的基本二次偏差为

$$M = \varphi - \varepsilon^2 (B\varphi \times \int g) \quad (8)$$

式中， M 为预期平方误差； φ 为总功率； ε 为存在预先确定的能量消耗； B 为预测时间； g 为电量的差值。

与预期的均方误差一起，模型的允许极限值由误差准则设定。每个测量区域的电力生产应具有上限和下限。定义一个具体的预测范围，调整每个单元的预测，创建实验环境，然后进行具体的实验分析。

2. 对比实验

在建立测试环境的基础上，对基于马尔可夫链的光伏电站发电功率超短期预测方法进行检验和分析。在完成所有四个区域的预测后，每天记录并比较未来的预测。然后估计发电量为50 000 kW，并使用监测和预测单元重新获得实际时间数据和信息。根据自动预测，测量每种方法的超短期电力预测结果，并将其与标准预测结果进行比较。最终的预测误差结果见表3。

表3 三种方法结果对比

测试指标	预测标准	传统方法1	传统方法2	本文方法
预测步长变动/m	1.30	2.83	2.19	1.33
迁移预测距离/m	9.20	11.62	13.82	9.16

由表3可知，与传统方法相比，本文开发的超短期预测方法最终的预测误差最小。传统方法1和传统方法2的预测误差显著。实验结果表明，本文研究方法预测精度更高，具有更好的预测效果，同时具有充分的实用价值和创新价值。

结束语

目前，在马尔可夫链技术支持下开发的光伏电站发电功率超短期预测方法具有较强的顺应性和变化性。与传统的电力预测方法相比，自动化程度高，进一步缩短了实际预测时间，提高了实际预测的准确性，实现了动态扰动的处理，提高了光伏的超短期预测效果。EA