基于 QPSO-LSTM 混合模型的光伏发电异常检测算法

何勤联1,张杨2,胥倪1,朱毅1,金霞1

(1. 国网江西省电力有限公司南昌供电分公司,江西南昌 330006;2. 国网江西省电力有限公司信息通信分公司,江西南昌 330096)

摘 要:为有效检测光伏系统发电异常,文中提出了一种融合量子粒子群优化算法(quantum particle swarm optimization,QPSO)与长短期记忆神经网络(long short-term memory,LSTM)的光伏功率异常检测方法。首先,进行光伏系 统数据的预处理,剔除异常值以确保数据质量,运用相关性分析,确定对光伏输出功率影响显著的气象参数,并将 其作为模型的输入特征;接着,采用QPSO算法,优化LSTM网络的超参数,构建一个高效且精准的预测模型;最后, 通过实际值与模型预测值对比,检测光伏系统异常情况。实验结果表明,构建的QPSO-LSTM模型可以增强预测的 准确性,通过分析实际功率输出与模型预测值之间的差异,可以有效识别光伏系统中的发电异常情况,为系统的稳 定运行提供了坚实的技术支持。

关键词:光伏预测;异常检测;量子粒子群优化算法;长短期记忆神经网络 中图分类号:TM 744 文献标志码:B 文章编号:1006-348X(2025)02-0042-06

0 引言

随着传统能源资源的渐趋枯竭,全球范围内推动 能源转型的需求日益迫切。为了应对这一挑战并促 进可持续发展,许多国家大幅增加了对新能源基础设 施的投资,其中光伏发电装机容量显著增长^[1-2],成为 可再生能源领域的重要组成部分。光伏系统不仅有 助于减少温室气体排放和缓解气候变化的影响,还对 提高能源安全、促进经济发展和改善生活质量具有重 要作用。然而,光伏系统长期暴露于自然环境中,易 受温度变化、湿度、沙尘及自然灾害等因素影响,导致 性能下降或故障发生;组件老化、电气连接问题和逆 变器故障也会干扰其正常运行,影响发电效率和经济 效益。因此光伏系统异常检测对于确保光伏系统的 可靠性、效率和使用寿命至关重要^[3]。

目前国内外学者对光伏异常检测已有大量研究, 这些研究主要集中在两类技术路径上,一是在光伏组 件上通过热成像、图像识别和红外线等技术进行异常 检测^[4-6];二是通过模型分析法对光伏设备数据和环境 因素数据进行分析检测^[7],如直接通过数据训练和异常 特征设定的检测模型^[8-9],由于这类方法受故障类型的 影响较大,所以一般采用物理模型、统计模型或机器学 习的方法研究[10]。光伏异常检测物理模型是用于识别 和诊断光伏系统中潜在故障的一类数学和物理建模工 具,这些模型通常依据光伏组件的电气性能和物理特 性,通过分析输出电流、电压、功率等参数的变化来判 断是否存在故障或异常情况[11]。基于统计学的模型方 法通常需要根据历史数据和统计分析,建立故障模式 与正常模式之间的差异来实现异常检测[12-13]。但由于 实际中通常存在环境干扰因素对光伏系统的测量影 响,上述模型可能会存在精度不够的局限性,对此神经 网络的方法广泛应用于相关研究;文献[14]采用基于机 器学习技术的支持向量机异常诊断方法,能够提高太 阳能发电系统异常诊断的准确性和效率;文献[15-16] 提出了用改进算法对神经网络优化进行光伏阵列异常 诊断;文献[17-18]提出长短期记忆神经网络和不同算 法相结合,通过数据集训练,最后使用误差指标进行精 度判断,然而这两种模型在算法选择和结果精度上均 有提升空间,因此可以考虑用其他模型进行优化。

根据上述研究现状和问题,文中旨在采用量子粒 子群优化算法(QPSO)增强的长短期记忆神经网络 (LSTM)预测模型,以提高光伏发电预测的准确性, 并将此预测模型应用于光伏发电异常检测。首先,通

收稿日期:2024-12-12

作者简介:何勤联(1982), 女, 硕士, 高级工程师, 研究方向为电力技术。

过聚类分析方法对历史光伏数据进行细致的清洗,以确保数据的准确性和可靠性;接着,运用相关性分析确定对光伏输出功率影响显著的气象参数,结合经过清洗的历史数据,构建出用于训练和测试 QPSO-LSTM 组合模型的数据集,该模型的预测结果将通过误差评价指标进行严格评估,并与其他模型的预测结果进行比较,以凸显其性能优势;最终,本研究将通过设定合理阈值的方法,利用预测结果,对光伏发电中的异常状况进行有效检测。

1 数据预处理

光伏发电预测模型需重视数据质量,避免数据采 集过程中产生的异常值影响模型泛化能力。

1.1 数据清洗

在本研究中,采用 K-means 聚类算法,对光伏发 电数据进行清洗,通过动态阈值设定与异常值二次校 验机制,识别并修正异常数据,具体流程如下:

1) 聚类数确定与阈值计算

首先,通过肘部法则与轮廓系数来确定最佳聚类 簇数,以确保数据划分的合理性。计算各簇内样本到 聚类中心的欧氏距离,采用"均值+3倍标准差"(3σ原 则)作为初始阈值。

2) 动态阈值调整与周期性验证

考虑到光伏数据的周期性(如昼夜、季节波动), 单一静态阈值可能导致对正常波动的误判。因此,针 对不同时段的聚类结果分别计算阈值,并通过交叉验 证对比不同σ倍数(如2σ、3σ、4σ)的清洗效果,最终 选择使预测模型RMSE最小的阈值参数。

3) 异常值二次校验机制

为避免对周期性波动的误清洗,对标记为异常的 数据点进行二次校验:若其相邻时段数据连续偏离阈 值,则判定为真实异常;若为孤立点,则保留原始 数据,这一策略兼顾了清洗的严格性与对数据自然波 动的容错性。

1.2 相关性分析

光伏发电量的变化主要受不稳定气候因素的影响。在光伏系统收集的数据中,涵盖了丰富的气象参数,其中部分参数对光伏发电量的预测具有决定性的作用,而其他一些则影响甚微。鉴于此,本研究采用

皮尔逊相关系数作为筛选工具,挑选出对光伏功率预 测模型贡献显著的气象变量,并将其作为模型的关键 输入,旨在提高预测的精确度和效率。

皮尔逊相关系数^[19]由统计学家卡尔·皮尔逊 (Carl Pearson)首次提出,用于量化两个数据集之间 的线性关系,相关系数的值域介于-1到1,其中-1表 示完全负相关,1表示完全正相关,而0则意味着没有 线性相关,公式如下:

$$\rho_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$
(1)

式中:n为样本总数; x_i 为第i次气象参数;x为气象参数 的平均值; y_i 为第i次PV输出功率;y为PV输出功率 的平均值。如表1所示,本研究经过皮尔逊相关分析, 选择辐照度和气温数据作为预测模型输入。

序号	气象特征	相关系数
1	总辐射辐照度(W/m ²)	0.9884
2	直接辐照度(W/m ²)	0.9797
3	散射辐照度(W/m ²)	0.5732
4	气温(℃)	0.4146
5	相对湿度(%)	-0.4016
6	降水量(mm)	-0.0615
7	风向角度(°)	-0.1453

表1 各气象特征与光伏发电相关性

1.3 数据归一化

为了处理收集到的数据的大小范围和测量单位的 变化,数据归一化至关重要,以确保这些特征数据共享 一致的度量尺度并防止不同尺度的影响。本研究采用 极值归一化方法,将原始数据范围转换至介于0与1之 间的标准化区间。该归一化方法的计算公式如下:

$$X_i^* = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$$
(2)

式中: X_i^* 是归一化后的值; X_i 是归一化的初始值; X_{max} 是最大值; X_{min} 是最小值。

2 光伏出力预测与故障检测

2.1 QPSO-LSTM 预测模型

2.1.1 量子粒子群算法

在标准粒子群优化(PSO)算法中,粒子的随机性 及其全局收敛性能,通常会因速度更新公式确定的粒 子速度的大小和方向而降低。与粒子群优化相比,量 子粒子群优化(QPSO)表现出更强的位置随机性,并 能有效避免陷入局部最优解,因为它不考虑粒子运动

研究探讨 🦻 PERSPECTIVE

方向,而专注于粒子位置的更新^[20]。在量子粒子群优 化算法中,核心的迭代流程包含以下步骤:

1) 初始化阶段

$$\begin{cases} x_{i}(t) = [x_{i1}(t), x_{i2}(t), \cdots, x_{iD}(t)] \\ P_{i}(t) = [P_{i1}(t), P_{i2}(t), \cdots, P_{iD}(t)] \\ P_{g}(t) = [P_{g1}(t), P_{g2}(t), \cdots, P_{gD}(t)] \end{cases}$$
(3)

在算法的起始阶段,会在一个D维的空间中创建 一个群体,该群体由M个粒子组成,每个粒子代表着 一个可能的解决方案,这里i从1到M变化,表示粒子 的索引,M为粒子总数,在所有粒子中,选出的最佳位 置用g来表示, $g \in [1,2,\cdots,M]$ 。 $x_i(t)$ 为第i个粒子在 第t次迭代时的当前位置, $P_i(t)$ 则表示第i个粒子在 其历史中的最优位置,而 $P_s(t)$ 则表示整个种群在历 史中的最优位置。

2) 搜索更新机制

随着迭代进行,粒子会参考自身的历史最优位置 P_i(t)和群体的全局最优位置P_s(t)来调整自己的位 置。首先,粒子会计算这两者之间的一个加权平均 值,并将这个值作为局部目标点P_{id}(t),然后,粒子会 向这个局部目标点移动。位置更新公式为:

$$P_{id}(t) = \varphi P_{id}(t) + (1 - \varphi) P_{gd}(t)$$
(4)

式中:d取值为1至D,而 φ是一个在0到1之间的随 机数,用于调整两个最优位置的影响程度。

在 t+1 次迭代中, 粒子 i 在 D 维空间的新位置 x_{id} (t+1)可以通过蒙特卡洛模拟法确定^[21], 具体公式为:

$$x_{id}(t+1)P_{id}(t) \pm \beta \times \left| m_{\text{best}} - x_{id}(t) \right| \times \ln\left(\frac{1}{u}\right)$$
(5)

式中:β是一个调节算法步长的参数;u表示0到1之间的随机数;m_{best}是群体的平均最优位置,其计算公式为:

$$m_{\rm best} = \frac{1}{M} \sum_{i=1}^{M} P_{id}(t)$$
 (6)

2.1.2 LSTM 模型

长短期记忆网络是一种递归神经网络的优化形式, 其核心是由多个相互连接的子网络构成。每个子网络 即为一个LSTM存储单元,如图1所示。每个LSTM单 元由三个主要的门控单元组成:输入门、遗忘门和输出 门。这些门单元为存储单元提供连续的读、写和复位操 作,使其能够在不同的门之间动态调整。此功能的目的 是维护输入数据序列的长期和短期记忆。



遗忘门管理前一时刻的单元状态*c*(*t*-1)对当前 时刻的单元状态*c*(*t*)的影响,决定保留哪些信息,忘 记哪些信息。相比之下,输入门在将当前时刻的输入 *x*(*t*)同化到单元状态*c*(*t*)中发挥作用,从而确定新信 息的纳入。输出门监视单元状态*c*(*t*)对LSTM 网络 中当前输出*h*(*t*)的影响,调节单元状态传输到输出的 程度。计算过程^[22]如下:

$$i_{t} = \sigma (W_{xi}x_{t} + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_{i})$$

$$f_{t} = \sigma (W_{fi}x_{t} + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_{f})$$

$$c_{t} = f_{hi}c_{t-1} + i_{t} \tanh (W_{xc}x_{t} + W_{hc}h_{t-1} + b_{c}) \quad (7)$$

$$o_{t} = \sigma (W_{xo}x_{t} + W_{ho}h_{t-1} + W_{co}c_{t} + b_{o})$$

$$h_{t} = o_{t} \tanh (c_{t})$$

式中:x,表示输入,h,表示输出。更具体地说,h,对应 于输入门的输出,f,表示遗忘门的输出,c,表示单元在 当前时间点(t)的状态,o,表示输出门的输出。与这些 过程相关的参数矩阵和偏差项分别用W和b表示。

2.1.3 QPSO-LSTM

本研究使用一种结合 QPSO 算法优化 LSTM 网络的模型,以提高光伏功率预测的准确度。QPSO 算法通过模拟量子行为的粒子在量子空间中的运动来优化 LSTM 网络的超参数,从而提升模型性能。如图 2 所示,QPSO-LSTM 模型的算法流程如下:



44

步骤1:导入经预处理的历史光伏功率、辐照度 及气温数据集。

步骤2:划分数据集为训练、验证与测试三部分。

步骤 3: 初始化 QPSO 算法及 LSTM 网络,设定 LSTM 模型的超参数作为待优化变量。

步骤4:进行QPSO迭代,将粒子的位置向量输入 LSTM模型进行训练,使用均方误差衡量模型性能, 并据此更新粒子的个体最优位置与全局最优位置。

步骤5:重复步骤4的迭代流程,直至找到最优粒子 位置或达到QPSO算法预设的最大迭代次数,结束优化。

步骤6:将最优粒子参数配置给LSTM模型,并使 用该模型对验证集进行预测。若预测结果的根均方 误差(root mean squared error, RMSE)超过设定的误 差阈值,则需返回步骤3,调整模型训练,直到预测误 差降低到可接受的范围内^[23]。

2.2 异常检测

文中针对突发性能量衰减的检测目标,选用晴天 场景进行异常分析。这是因为阴雨天气因出力波动 剧烈,难以通过发电数据判断设备状态;而晴天无云 层干扰,光伏出力曲线呈平滑拱形,更易识别异常波 动^[24]。具体来说,在晴朗无云的天气条件下,阈值依 据历史数据分析和动态调整需求设定为预测光伏发 电量的80%至120%,结合连续两个采样点才判定异 常以排除偶发干扰^[25]。基于QPSO-LSTM预测模型 的光伏发电异常检测方法完整流程如图3所示。



文中方法通过预测与实发功率偏差实现光伏异常 检测,其核心优势在于:实时监测,无需红外/热成像等专 用设备,显著降低成本;采用双采样点超阈值策略,对组 件遮挡、逆变器效率骤降等突发故障的早期预警优于传 统均值阈值法;适用于无人值守电站的自动化监测场 景。但局限性在于:模型在极端天气下预测波动较大; 仅能判定发电异常而无法定位热斑、线路故障等具体问 题,需结合红外二次诊断。本方法更适合作为分布式电 站异常初筛系统,与图像识别、电气检测技术互补。

3 算例分析

3.1 实验数据和设置

本研究数据来源于某楼宇光伏系统,数据采集频 率为每15 min一次。这些数据涵盖了电力参数与环 境条件两方面,电力参数包括了光伏发电额定功率和 实际功率;环境条件则包含了辐照度和气温信息。文 中使用该系统2018年1月1日至2019年5月31日的 晴天数据,其中前70%的数据进行模型训练,使用中 间15%的数据用于验证预测模型性能,用最后15% 的数据测试故障检测效果。为提高故障检测的准确 性,研究中只选取发电量超过0.04 kWh的数据进行 分析,以忽略日落时的自然低发电量。

通过上述方法,可以实时检测光伏系统的发电异常 情况,及时发现并处理潜在的故障,确保系统的高效运 行。本研究建立了LSTM、PSO-LSTM和QPSO-LSTM 三种预测模型对比体系,所有模型均采用Xavier正态分 布初始化权重偏置,通过trainNetwork函数结合Adam 优化器及反向传播算法进行参数优化。其中,基准 LSTM模型通过网格搜索法在预设离散空间内确定最 优的固定超参数配置(各隐藏层神经元数、训练次数及 学习率);而PSO-LSTM与QPSO-LSTM模型则在统一 超参数空间内进行智能优化(各隐藏层神经元数[1, 100]、训练次数[10,100]、学习率[0.001,0.1]),并设置相 同的粒子种群规模与最大迭代次数。所有实验均在相 同计算环境及数据集划分下进行,确保模型性能对比的 公平性。表2列出了本次研究的具体参数配置。

表2 实验参数设置

模型	超参数类型	超参数值
LSTM	-	未进行优化
	隐藏层神经元数L1	114
DSO I STM	隐藏层神经元数L2	122
FSO-LSTM	训练次数	91
	学习率	0.0085
	隐藏层神经元数L1	160
ODSO I STM	隐藏层神经元数L2	6
QF3O-L31M	训练次数	83
	学习率	0.0079

3.2 预测结果分析

为评估所提出方法的有效性,文中对LSTM、 PSO-LSTM和QPSO-LSTM三种预测模型在光伏发 电量预测任务中的表现进行了比较。图4展示了这 些模型的预测曲线与实际功率曲线的对照。

分析图4的结果,可以发现LSTM模型虽然能够把 握功率变化的总体趋势,但在细节捕捉上略显不足。 PSO-LSTM模型在这方面有所进步,能够更准确地反映 实际功率曲线的波动,尽管预测误差仍然存在。而文中 提出的QPSO-LSTM模型在拟合实际功率曲线方面表现 卓越,其预测曲线与实际数据的吻合度显著高于其他两 个模型,从而验证了其在光伏发电量预测中的优越性能。



图4 各模型预测功率与实际功率对比

为了深入评价各模型的表现,本研究选用了三大 核心评价指标:根均方误差(RMSE)、平均绝对误差 (mean absolute error, MAE)以及平均绝对百分比误 差(mean absolute percentage error, MAPE)。这些指 标提供了一个量化的视角,以比较LSTM、PSO-LSTM和QPSO-LSTM模型在光伏发电量预测任务 中的表现。各模型评价指标对比如表3所示。

表3 各预测模型评价指标对比

预测模型	RMSE	MAE	MAPE/%
LSTM	0.2914	0.2082	13.577
PSO-LSTM	0.1501	0.1020	5.656
QPSO-LSTM	0.1282	0.0899	3.604

从表3中的数据可以看出,与传统的LSTM模型 相比,PSO-LSTM和QPSO-LSTM模型在预测精度和 拟合度上均有显著提升。具体来说,PSO-LSTM模型 在RMSE、MAE和MAPE三个指标上分别降低了 48.2%、50.5%和58.4%。而QPSO-LSTM模型的表现 更是出色,其RMSE、MAE和MAPE相比LSTM模型 分别降低了56.0%、56.7%和73.4%,相比PSO-LSTM 模型也分别降低了14.6%、11.9%和36.9%。

3.2 异常检测结果

为了评估异常检测方法的有效性,文中选取了若 干异常时间段进行分析,将实际功率与模型预测的功 率进行对比。如图5所示,通过设定的阈值,可以成 功地识别出发电异常,证明了该方法在实际应用中的 准确性和可靠性。



4 结语

本研究提出的 QPSO-LSTM 模型不仅提升了光

伏发电量的预测精度,而且为实时监测和异常诊断提 供了有效的工具,有助于保障光伏系统的长期稳定运 行。未来的工作可以考虑增加不同地理位置和气候 条件的数据,并引入更多影响因素,提升预测模型的 鲁棒性。此外,尽管文中能识别光伏功率的异常情 况,但未深入研究具体异常类型。未来工作应专注于 开发异常分类方法,确定具体的故障类型(如逆变器 故障、遮挡问题或组件老化),以提供更精准的异常诊 断和更具针对性的维护建议,从而实现更高效的系统 维护。

参考文献:

- [1] 吉兴全,刘健,叶平峰,等.计及灵活性与可靠性的综合能源 系统优化调度[J].电力系统自动化,2023,47(8):132-144.
- [2] 周勤勇,李根兆,秦晓辉,等.能源革命下的电力系统范式 转换分析[J].中国电力,2024,57(3):1-11.
- [3] Pillai D S, Rajasekar N.A comprehensive review on protection challenges and fault diagnosis in PV systems[J].Renewable and Sustainable Energy Reviews, 2018, 91:18-40.
- [4] 梁健锋,胡振球,黄灿,等.基于红外热成像检测的光伏电 站异常分析[J].太阳能,2024(1):70-76.
- [5] 林维修,李峰,王海峰,等.基于图像处理的光伏组件热斑缺 陷检测方法[J].计算技术与自动化,2024,43(3):121-126.
- [6] Mellit A.An embedded solution for fault detection and diagnosis of photovoltaic modules using thermographic images and deep convolutional neural networks[J]. Engineering Applications of Artificial Intelligence, 2022, 116:105459.
- [7] Qu J, Qian Z, Pei Y, et al. An unsupervised hourly weather status pattern recognition and blending fitting model for PV system fault detection[J]. Applied Energy, 2022, 319:119271.
- [8] Li B, Delpha C, Diallo D, et al. Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: A review[J]. Renewable and Sustainable Energy Reviews, 2021,138:110512.
- [9] Ramadan E A, Moawad N M, Abouzalm B A, et al. An innovative transformer neural network for fault detection and classification for photovoltaic modules[J]. Energy Conversion and Management, 2024, 314: 118718.
- [10] 朱润泽, 王德军. 基于 LSTM 神经网络的光伏系统功率预 测[J].电力科技与环保,2023,39(3):201-206.

- [11] Satpathy P R, Aljafari B, Thanikanti S B, et al. Electrical fault tolerance of photovoltaic array configurations: Experimental investigation, performance analysis, monitoring and detection[J].Renewable Energy, 2023, 206:960-981.
- [12] 丛伟伦,张博,夏亚东,等.基于马尔可夫链的光伏电站遮 挡实时诊断算法[J].太阳能学报,2020,41(4):67-72.
- [13] Harrou F, Sun Y, Taghezouit B, et al. Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches[J]. Renewable Energy, 2018, 116: 22-37.
- [14] Lu S D, Liu H D, Wang M H, et al. A novel strategy for multitype fault diagnosis in photovoltaic systems using multiple regression analysis and support vector machines[J]. Energy Reports, 2024, 12:2824-2844.
- [15] 李斌,高鹏,郭自强.改进蜣螂算法优化LSTM的光伏阵列故 障诊断[J]. 电力系统及其自动化学报, 2024, 36(8): 70-78.
- [16] Zhou N, Shang B, Xu M, et al. Enhancing photovoltaic power prediction using a CNN-LSTM-attention hybrid model with Bayesian hyperparameter optimization[J].Global Energy Interconnection, 2024, 7(5):667-681.
- [17] 秦宇,许野,王鑫鹏,等.基于改进FCM-LSTM的光伏出 力短期预测研究[J].太阳能学报,2024,45(8):304-313.
- [18] 王涛, 王旭, 许野, 等. 计及相似日的 LSTM 光伏出力预测 模型研究[J].太阳能学报,2023,44(8):316-323.
- [19] 薛阳,李金星,杨江天,等.基于相似日分析和改进鲸鱼算 法优化LSTM网络模型的光伏功率短期预测[J].南方电 网技术,2024,18(11):97-105.
- [20] Xin-gang Z, Ze-qi Z, Yi-min X, et al. Economic-environmental dispatch of microgrid based on improved quantum particle swarm optimization[J]. Energy, 2020, 195:117014.
- [21] 周晓东,肖正江,杨坤鹏,等.基于改进量子粒子群算法的叶片延 长翼型厚度优化设计[J].节能技术,2024,42(4):345-352.
- [22] A wind power prediction framework for distributed power grids[J].Energy Engineering, 2024, 121(5):1291-1307.
- [23] 谭才兴,岳雨霏,汤赐.基于QPSO-LSTM的短期风电负荷 预测模型[J].中阿科技论坛(中英文),2023(12):88-91.
- [24] 王晓倩,周羽生,毛源军,等.基于神经网络分位数的分布 光伏发电功率异常识别方法[J/OL].上海交通大学学报, 2024:1-18 [2024-02-28].DOI:10.16183/j.cnki.jsjtu.2023.412.
- [25] 赵博超,马嘉骏,崔磊,等.基于改进 VMD-XGBoost-BiL STM 组合模型的光伏发电异常检测[J]. 计算机工程, 2024,50(3):306-316.