

浑源风电场风速预报模型优化

胡彦昭¹ 吴雪彤² 吴亮³ 李博宇⁴

1 浑源县气象局, 山西省大同市, 037400;

2 内蒙古自治区丰镇市气象局, 内蒙古自治区, 012100;

3 山西省浑源县气象局, 山西省浑源县, 037400;

4 吉林省磐石市气象局, 吉林省磐石市, 132300;

摘要: 在全球能源结构加速向清洁能源转型的背景下, 风力发电已成为可再生能源的重要组成部分。然而, 复杂地形区域的风速预报精度不足, 直接影响风电场的发电效率和经济收益。本研究以山西浑源风电场为研究对象, 针对其峡谷地形导致的局地风场加速效应、垂直风切变显著及极端天气频发等问题, 提出了一套多维度协同优化方法。首先, 基于 30 米分辨率 DEM 数据构建地形动力因子, 量化峡谷加速效应; 其次, 融合 LSTM 时序建模与 CNN 空间特征提取能力, 构建时空混合深度学习架构, 并引入动态注意力机制强化关键时段预报; 最后, 采用改进粒子群算法 (PSO) 进行超参数自适应优化。数据预处理阶段创新性地结合 Z-score 统计检验与物理阈值筛选, 有效提升异常数据清洗精度。经 2022-2023 年实测数据验证, 优化模型在峡谷区域的平均绝对误差 (MAE) 降至 0.89 m/s, 极端风速预警命中率提升至 82%, 计算效率满足业务实时性需求 (单次预报<15 秒)。该研究为山地风电场的高精度风速预报提供了可推广的技术方案, 对提升风电消纳能力和电网稳定性具有重要实践意义。

关键词: 浑源风电场; 风速预报模型; 模型优化; 预报准确性

DOI: 10.69979/3041-0673.25.12.016

风速预报的准确性直接影响风电场的发电调度、设备安全和经济效益。当前, 全球风电装机容量持续增长, 但复杂地形区域的风速预测仍面临严峻挑战。传统数值预报模型 (如 WRF) 在平坦地形表现良好, 但对山地风电场特有的局地环流、地形强迫效应等非线性特征捕捉不足, 导致预报误差显著偏高。此外, 现有机器学习模型多依赖单一测风塔数据, 难以反映轮毂高度风切变及空间异质性, 且对突发性极端天气 (如雷暴、沙尘暴) 的预警能力有限。

本研究以浑源风电场为例, 创新性地将物理机制与数据驱动方法相结合: 通过 DEM 数据量化地形粗糙度与峡谷加速效应, 构建专属特征工程; 设计 LSTM-CNN 混合模型, 增强时空特征提取能力; 并采用改进 PSO 算法实现超参数自适应调优。研究结果不仅提升了预报精度, 还为同类山地风电场的预测模型优化提供了方法论参考, 对推动清洁能源高效利用具有重要现实意义。

1 浑源风电场风速预报现状及问题分析

1.1 本地风速特征

浑源风电场位于恒山南麓 (北纬 39° 23' -39° 45', 海拔 1650-1950m), 覆盖大仁庄乡、官儿乡等 7 个乡镇。

截至当前, 全县已建成 12 家风电场 (总装机 86.6 万千瓦), 其中大仁庄乡 (26.85 万千瓦) 和官儿乡 (24.9 5 万千瓦) 为装机集中区, 分别布局 5 家和 3 家风电场。受复杂地形及局地气候影响, 风速呈现显著时空分异特征。地形强化效应突出, NE-SW 走向的峡谷使盛行 NW 风加速 1.5-2 倍, 2022 年实测数据显示峡谷中部风速较周边高 38%, 最大瞬时风速达 28m/s。季节差异明显, 春季 (3-5 月) 受蒙古高压影响, 平均风速 5.2m/s, 大风日数占比 40%; 夏季 (6-8 月) 平均风速降至 3.5m/s, 但局地热力环流引发的雷暴大风占比达 12%, 且多伴随短时强降水。日变化呈现“双峰”特征, 受山谷风环流驱动, 午后 14-16 时风速较夜间高 45%, 而日落前后 (17-19 时) 因热力环流减弱出现风速骤降, 小时变幅超过 3m/s。

1.2 现有模型缺陷

现行预报模型难以适应浑源特殊地理环境, 存在三方面短板: 地形适应性方面, 传统模型将面积达 28km² 的风场简化为均质平面, 导致对#15-#22 机组 (位于峡谷隘口处) 的风速预报误差持续高于平坦区域机组 41%, 2023 年运维记录显示该区域因误报导致的发电损失占

比达 27%。数据代表性方面,模型仅依赖 10m 高度测风塔数据,忽略轮毂高度(150m)处的风切变效应,实测表明风速垂直梯度达 $0.12\text{m/s}\cdot\text{m}$,致使功率曲线换算误差超过 15%。极端事件预报方面,模型对突发性天气响应滞后,如 2023 年“4·15”沙尘暴过程(瞬时风速 26m/s)提前 6 小时漏报,造成全场 23 台机组紧急停机,直接经济损失达 58 万元。此外,现有线性模型对夏季雷暴大风和冬季寒潮大风的预报准确率不足 50%,严重制约风电场的竞价上网和设备安全管理^[1]。

2 数据预处理优化

2.1 本地化数据清洗

针对浑源风电场特有的数据质量问题,研究采用多维度清洗策略。在异常值处理方面,针对春季频发的沙尘天气干扰(2023 年沙尘日数达 28 天),创新性地将传统 Z-score 算法与风速物理阈值相结合:设置动态阈值 (3.5σ),当数据同时满足统计异常 ($|Z|>3.5$) 和物理异常 (10m 高度风速 $>25\text{m/s}$ 且相对湿度 $<30\%$) 时进行修正,使 2023 年 4 月数据有效样本提升 23%。针对地形导致的观测盲区,融合 30m 分辨率 DEM 数据,基于 CGCS2000 坐标系建立空间插值模型,通过计算地形粗糙度指数 (z_0) 和坡度方位角,将单点观测扩展为空间场数据,使#15-#22 机组区域的测风数据覆盖率从 68%提升至 92%。

2.2 特征工程增强

结合浑源地理气候特征构建专属特征体系:1) 峡谷动力因子,计算 NW 风向 ($300^\circ\sim330^\circ$) 与主峡谷走向 (35°) 的夹角余弦值,量化地形加速效应,实测显示该因子与风速的 Spearman 相关系数达 0.73;2) 热力环流指标,利用浑源站(海拔 1720m)与广灵站(海拔 980m)的 2m 温度差构建局地环流强度指数,在夏季午后强对流时段(14-16 时)与风速的滞后相关系数达 0.61;3) 极端事件记忆特征,标注近 5 年 26 次大风预警事件的时间戳,在模型训练中通过样本加权(权重系数 1.5-3.0)强化特殊天气模式的学习。所有特征均经过 Min-Max 归一化处理,如将风速 ($0\sim28\text{m/s}$) 线性映射至 $[0,1]$ 区间,确保不同量纲特征的均衡贡献。

3 模型结构优化

3.1 本地化模型架构

针对浑源风电场的特殊需求,本研究创新性地构建了时空融合的深度神经网络架构。该模型采用双分支并行结

构:时序处理分支采用三层 LSTM 网络(隐层单元数 64),时间步长设为 12 小时以匹配浑源风速的自相关周期(实测显示滞后 12 小时相关系数达 0.82);空间特征分支采用二维 CNN(卷积核 3×3),输入 5×5 网格化地形粗糙度数据(30m 分辨率),捕捉峡谷地形的局地加速效应。特别引入注意力机制对午后强风时段(12-18 时)赋予动态权重(3-5 倍),在 2023 年夏季测试中,该机制使关键时段的预报准确率提升 19%^[2]。

3.2 模型性能验证

通过 2022-2023 年实测数据对比分析:传统 ARIMA 模型因无法处理非线性地形效应,在峡谷区域的 MAE 高达 1.8m/s ,且对“4·15”沙尘暴过程的预报滞后达 8 小时;基础 LSTM 模型通过记忆单元有效捕捉了季节周期特征,将整体 MAE 降至 1.2m/s ,但对突发性天气的响应仍存在 30-40 分钟延迟;优化后的 LSTM-Attention 混合模型在三个关键指标上显著提升:1) 极端风速预警准确率(命中率达 82%);2) 峡谷区域 MAE (0.89m/s);3) 计算效率(单次预报耗时 <15 秒)。该模型已成功预测 2023 年 11 月寒潮过程,提前 9 小时发布预警,为风电场争取了宝贵的应急准备时间。

4 参数优化与业务应用

4.1 本地化 PSO 参数设置

针对浑源风电场特有的参数优化需求,研究采用改进的粒子群优化算法(PSO)对 LSTM 模型进行调优。在算法参数设置方面:1) 种群规模设为 20,与风电场 16 个测风塔及 4 个气象站的空间分布相匹配,确保参数搜索的充分性;2) 惯性权重采用线性递减策略 ($0.9\rightarrow0.4$),以适应浑源地区春季风速波动大、冬季风速稳定的季节特征;3) 优化目标函数特别强化 14-16 时关键时段的误差权重(占比 40%),以提升对午后强风时段的预报精度。与遗传算法相比,该方案迭代收敛速度提升 35%,且避免了模拟退火算法对初始温度敏感的问题。

4.2 参数优化实施

优化过程分为三个阶段:首先确定 LSTM 核心参数的搜索空间,包括学习率 ($0.001\sim0.01$)、隐藏层神经元数 (32-128)、dropout 率 ($0.1\sim0.3$) 等关键参数;其次构建适应度函数,采用验证集 MAE 与关键时段命中率的加权组合(权重比 6:4);最后进行迭代优化,经过 200 代进化后,获得最优参数组合:学习率 0.0057、隐藏层神经元 96、注意力头数 4。优化后的模型在测试

集上 MAE 降至 0.86m/s, 较基线模型提升 28%。

4.3 业务应用成效

该模型自 2023 年 12 月投入业务运行以来表现突出:

1) 实现每日 4 次滚动预报 (00/06/12/18 时), 通过 API 接口实时接入风电场集控系统, 平均延迟 <3 秒; 2) 在 2024 年 1 月寒潮过程中, 提前 9 小时准确预测 NW 向 10 级阵风 (24.7m/s), 为风电场争取了充分的机组保护时间, 减少直接经济损失 85 万元; 3) 发电计划匹配度从 72% 提升至 89%, 2024 年第一季度减少弃风电量达 320 万度。目前该模型已扩展应用于周边 3 个风电场, 形成区域风速预报服务网络^[3]。

5 优化后模型的验证与评估

5.1 验证方法

为全面评估模型优化效果, 采用分层抽样策略构建数据集。基于浑源风电场 2018-2023 年完整气象观测数据 (时间分辨率 10 分钟), 按 7:1:2 比例划分训练集、验证集与测试集, 确保各子集覆盖不同季节、天气类型及地形特征^[4]。训练阶段采用滑动时间窗技术 (窗口宽度 24 小时, 步长 1 小时) 生成样本序列, 通过验证集实时监控模型在春季大风 (3-5 月) 和夏季雷暴 (6-8 月) 等典型场景下的表现。特别设置“极端事件测试子集”, 包含 2021-2023 年记录的 32 次沙尘暴与寒潮过程数据, 用于专项评估模型在突发性天气中的预警能力。测试阶段严格隔离数据流, 测试集数据不参与任何参数调整, 确保评估结果客观反映模型实际应用性能^[5]。

5.2 评估指标

建立三级评估体系: 基础精度指标采用 MAE 和均方根误差 (RMSE), 其中 MAE 侧重反映平均预报偏差 (单位 m/s), RMSE 通过平方运算突显大误差事件的惩罚; 稳定性指标选用误差标准差 (SDE), 量化模型输出波动程度, 2023 年测试数据显示优化模型 SDE 降低至 0.73m/s, 较原模型改善 31%; 业务适用性指标包括预警准确率 (POD) 与误报率 (FAR), 针对风速 $\geq 12\text{m/s}$ 的大风事件, 优化模型 POD 提升至 85.7% (原模型 62.3%), FAR 控制在 14.2%^[6]。通过误差贡献分解发现, 地形复杂区域的预报误差占比从 41% 降至 28%, 验证了空间特征分支的有效性。相关系数分析表明, 优化模型在 NW 风向 ($300^\circ - 330^\circ$) 下的 R 值达 0.91, 较其他风向高 0.

15-0.2, 准确捕捉了峡谷地形加速效应。所有指标均通过显著性检验 ($p < 0.01$), 证实改进具有统计学意义^[7]。

6 结论与展望

6.1 结论

本研究通过融合动态阈值清洗与空间插值技术, 将峡谷区域数据覆盖率提升至 92%; 构建 LSTM-Attention 混合模型, 结合地形特征增强时空建模能力, 极端风速预警命中率达 82%; 采用改进粒子群算法优化参数, 关键机组 MAE 降至 0.89m/s。测试表明, 模型整体误差降低 25%, NW 风向峡谷区 ($R=0.91$) 和夏季强对流时段 (误差降幅 19%) 改进显著。

6.2 展望

未来研究将聚焦多源数据融合 (如激光雷达与 WRF-LES 耦合)、LSTM-Transformer 混合架构开发 (集成 SCADA 数据补偿尾流效应), 以及基于强化学习的动态调优系统, 实现极端天气的分钟级响应, 构建高精度智能预报体系。

参考文献

- [1] 王一博. 风电场运行风速数据清洗与重构方法研究 [D]. 华北电力大学 (北京), 2022.
- [2] 陆嘉铭, 奚增辉, 瞿海妮, 等. 电力量测数据缺失补齐方法研究与实践 [J]. 电力大数据, 2023, 26 (07): 40-49.
- [3] 冯英, 冯鑫媛. 四川盆地冬季气溶胶污染与气象条件的关系 [J]. 高原山地气象研究, 2022, 42 (03): 117-125.
- [4] 陈钊华. 基于深度学习的短期电力负荷概率预测 [D]. 广东工业大学, 2022.
- [5] 毕琪. 神经网络及其混合模型在风速预报中的研究 [D]. 西北师范大学, 2024.
- [6] 周统汉. 复杂地理环境下风电场风能特性分析与评估 [D]. 电子科技大学, 2017.
- [7] 靳春旭, 董福贵. 长期风电负荷预测方法比较 [J]. 广东电力, 2018, 31 (09): 70-76.

作者简介: 胡彦昭, 出生年月: 1995 年 10 月 15 日, 性别: 男, 民族: 汉, 籍贯: 山西大同, 学历: 本科, 当前职称: 助理工程师, 研究方向: 气象。