

整县光伏发电的综合 LSTM 功率预测模型研究

仓 敏, 翟晓萌, 王静怡, 吴 霜, 程 曦

(国网江苏省电力有限公司 经济技术研究院, 南京 210028)

摘要: 为了更好地实现整县屋顶光伏集约化发展, 提高大规模分布式屋顶光伏并网经济性, 分析基于气象数据的预测方法弊端问题, 挖掘海量时序光伏发电功率数据规律, 提出基于改进长短期记忆递归神经网络(LSTM)的整县光伏综合功率预测模型。使用所提模型进行单一光伏电源预测实验, 结果表明, 改进 LSTM 模型较其他模型更精准, 且在整县光伏综合功率预测场景下具有更高整体预测精度, 可以用于整县光伏发电功率预测, 有助于实现高比例分布式光伏消纳。

关键词: 整县光伏; 分布式光伏; 长短期记忆网络; 功率预测

中图分类号: [TK-9] **文献标志码:** A **文章编号:** 1671-1807(2023)06-0156-09

随着“双碳”目标的提出, 中国加快了风电、光伏等清洁能源对传统化石能源的替代进程, 明确将风能光伏等清洁能源作为重点支持行业。然而当前光伏资源一类区大多远离城市负荷中心, 远距离输电成本较高, 需要开发临近负荷中心的光伏资源, 进一步提高光伏发电功率在总负荷中的占比。

国内负荷中心主要位于城市中心及工业园区, 建筑空置屋顶资源丰富, 开发建设屋顶分布式光伏潜力巨大。同时, 建筑屋顶主体分散、单体规模小、聚集度低、开发协调工作量大等问题, 也在一定程度上制约了屋顶分布式光伏发展。

2021 年 6 月, 国家能源局提出用整县屋顶分布式光伏集约化发展模式代替传统分散模式, 要求做到“宜建尽建”“宜接尽接”, 在光伏领域加快实现“碳达峰、碳中和”目标, 用清洁能源助力乡村振兴^[1]。

在此背景下, 江苏加速建设以新能源为主的新型电力系统, 预计到 2025 年省内分布式光伏装机容量将超 2 000 万 kW, 远景可开发容量 7 600 万 kW, 与此同时, 分布式电源与多元负荷规模化接入配电网, “双高、双峰”特征显著, 导致系统中不确定和不稳定的因素增加, 系统调控更加复杂^[2], 亟须创新光伏功率预测方法, 提前把握光伏发电规律趋势, 为

其他类型电源协同调度提供支撑^[3], 进一步提高大规模分布式光伏使用效益。

1 光伏发电功率影响因素

学界对光伏发电功率的规律进行了大量研究, 提出了针对单一光伏站点的基于时间序列分析、回归分析的诸多预测方法^[4], 以地面气象站数据、卫星图像数据、光伏系统运行数据和数值天气预报数据为主要研究对象^[5], 结合光伏发电系统特征、地理位置等信息, 构建模型拟合输入特征数据与光伏发电系统电力输出的关系, 从而实现光伏发电预测。综合已有研究可知, 影响光伏发电功率数据的因素主要包括太阳辐照强度、温度、风力、云量、水汽、气溶胶、气压因素^[6], 各种影响因素之间存在相关性, 对光伏发电功率产生复杂影响。

1.1 光伏发电功率影响因素相关性

为探究整县光伏背景下光伏发电功率影响因素、各影响因素与光伏功率间的相关关系、研究以光伏电站 Q 为例, 基于气象及发电功率历史数据, 对光伏发电功率的影响因素进行相关性分析。光伏电站 Q 为集中式光伏电站, 总装机容量 20 MWP, 占地面积超 40 万 m², 所处地区气象条件稳定, 临近气象观测点数据可靠。本文采用皮尔逊相关系数衡量多个变量间相关关系计算公式为

收稿日期: 2022-11-10

作者简介: 仓敏(1979—), 女, 江苏南京人, 国网江苏省电力有限公司经济技术研究院, 技术经济中心副主任, 硕士, 研究方向为电网项目诊断评价; 翟晓萌(1988—), 男, 河北邢台人, 国网江苏省电力有限公司经济技术研究院, 高级工程师, 硕士, 研究方向为电网项目诊断评价; 王静怡(1973—), 女, 天津人, 国网江苏省电力有限公司经济技术研究院, 高级经济师, 研究方向为技术经济分析评价; 吴霜(1989—), 女, 江苏泰州人, 国网江苏省电力有限公司经济技术研究院, 高级工程师, 硕士, 研究方向为电网项目诊断评价; 程曦(1993—), 男, 安徽淮南人, 国网江苏省电力有限公司经济技术研究院, 工程师, 硕士, 研究方向为电网环保和投资管理。

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

式中： r_{xy} 为皮尔逊相关系数； x_i, y_i 为序列 x 、序列 y 中的第 i 个分量； $\bar{x}, \bar{y}$ 为对应序列均值。

皮尔逊相关系数 r_{xy} 是 $[-1, 1]$ 间实数,当 r_{xy} 大于 0 时,两变量正相关,当 r_{xy} 小于 0 时,两变量负相关, r_{xy} 的绝对值越趋近于 1,则表示两个变量之间相关程度越高。

经过计算,得到光伏电站 Q 中,多种气象因素与光伏发电功率的相关系数。如表 1 所示,光伏发电功率与辐照度相关系数达到 0.798,具有最高关联性,与环境温度相关度达到 0.724,呈高度相关,与其他气象因素间也存在一定相关关系。如图 1 所示,各弧表示不同变量,若两条弧中存在连接线,则存在相关性,弧间连线的宽度反映了各变量间相关度数量关系,连线越粗则相关度越高。

受气象数据限制,研究未能体现气溶胶和云量影响因素对其他影响因素及光伏发电功率的影响。通过上述相关性分析可知,环境温度和辐照度与光

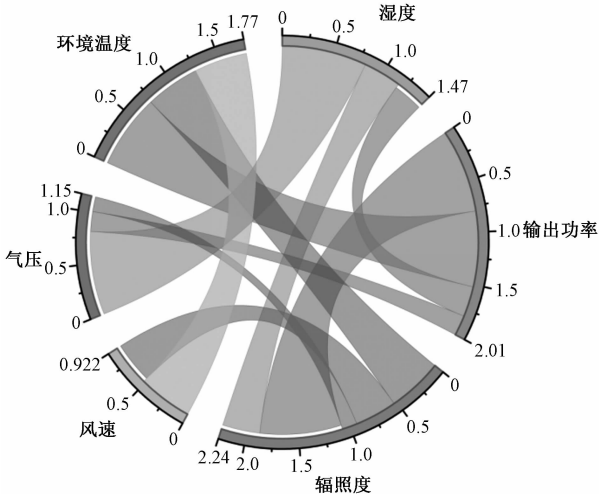


图 1 气象因素与光伏输出功率间影响关系

伏输出功率具有较强相关性,而组件温度是环境温度、辐照度、云量及风速的综合作用结果,基于精确实时的辐照度、环境温度数据,可进行气象范围内短期光伏功率预测。

1.2 光伏发电功率影响因素

实际场景下,县域内分布式屋顶光伏所在位置最大距离可达到 50 km 以上,且所在地理位置分散、地表形态各异,环境温度、辐照度等关键影响因素均具有较大差异。为探究县域内同一时刻气象条件差异,研究通过风云四号卫星数据表现江苏全域 2022 年 10 月 30 日早 9:00 气象条件示意图,图中标注了江苏省界和省内各县界。

如图 2 所示,图 2 中以不同灰度和数据点反映了江苏省内各地同一时刻地表大气温度,图中各条等温线曲折穿过不同县域,在部分县域内存在小范

表 1 光伏发电功率与气象因素之间的关联分析结果

变量	环境温度	辐照度	气压	湿度	风速	输出功率
环境温度	1	0.53	-0.706	-0.95	0.52	0.724
辐照度	0.53	1	0.144	-0.367	0.402	0.798
气压	-0.706	0.144	1	0.813	-0.371	0.194
湿度	-0.95	-0.367	0.813	1	-0.593	-0.29
风速	0.52	0.402	-0.371	-0.593	1	0.239
输出功率	0.724	0.798	0.194	-0.29	0.239	0

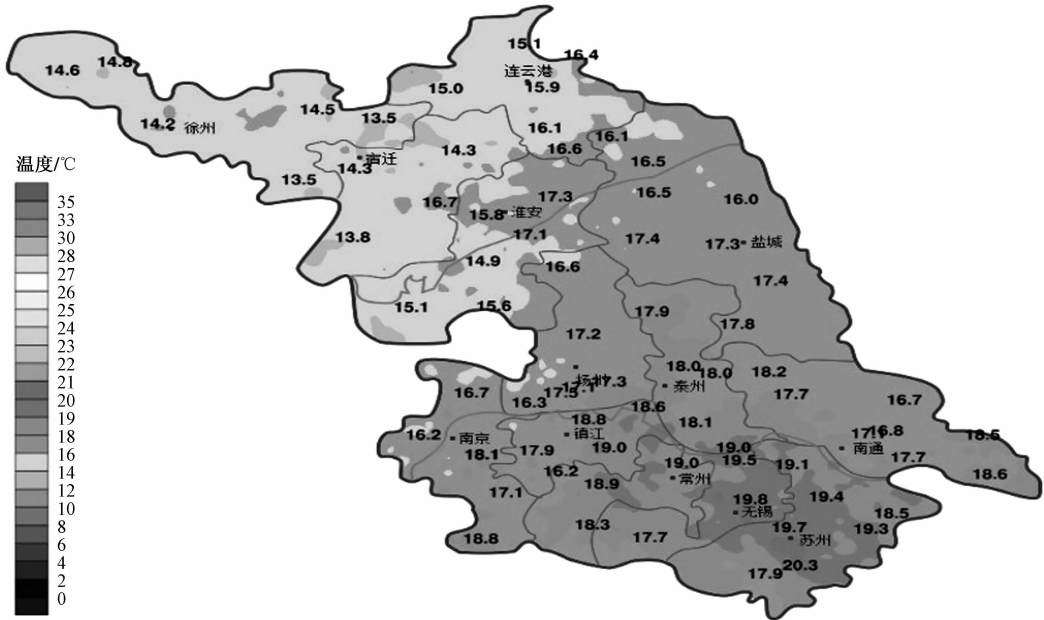


图 2 区域大气温度

围高温或低温区域。如图 3 所示,深色区域即太阳下行辐射强度低于光伏发电辐照强度阈值,图中颜色越亮的区域表示辐照强度越强,越暗的区域则辐照强度越弱,由图可知,淮安市金湖县等地部分县域内存在较大下行辐射差距,如果采用县内少量观测点数据,难以对各分布式光伏状况进行预测。如图 4、图 5 所示,图 4 中白色区域表示当地在此时刻天气条件为阴天,浅灰色地区当地为多云天气,深灰色地区当前天气为晴天。图 5 是红外波段在观测时间点中各地云量状态,可以直观看出各县域内云的厚度、形态存在显著差异。由图 2~图 5 可知,江

苏省各县域内在同一时刻可能存在多种不同天气形态,现有不同气象站点数据收集密度精度各不相同,且无法及时将温度及其他数据变化情况上传分析,难以满足负荷光伏运行主体需要,县域内分布式光伏电源无法通过现有气象数据进行功率预测。

2 光伏功率预测模型

2.1 已有预测方法研究

近年来,学界采用指数平滑模型法^[7]、卡尔曼滤波法^[8]、多元线性回归法^[9]自回归积分滑动平均模型^[10],对光伏功率预测进行了众多研究。董雷等提出一种条件概率预测方法,应用动态贝叶斯网络理

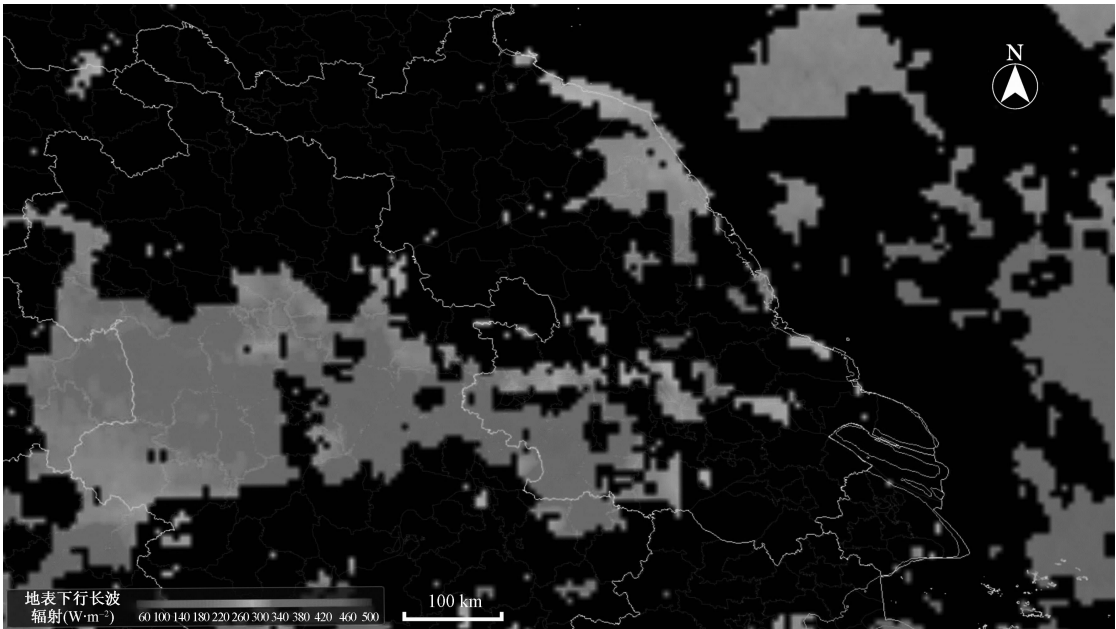


图 3 区域下行辐射

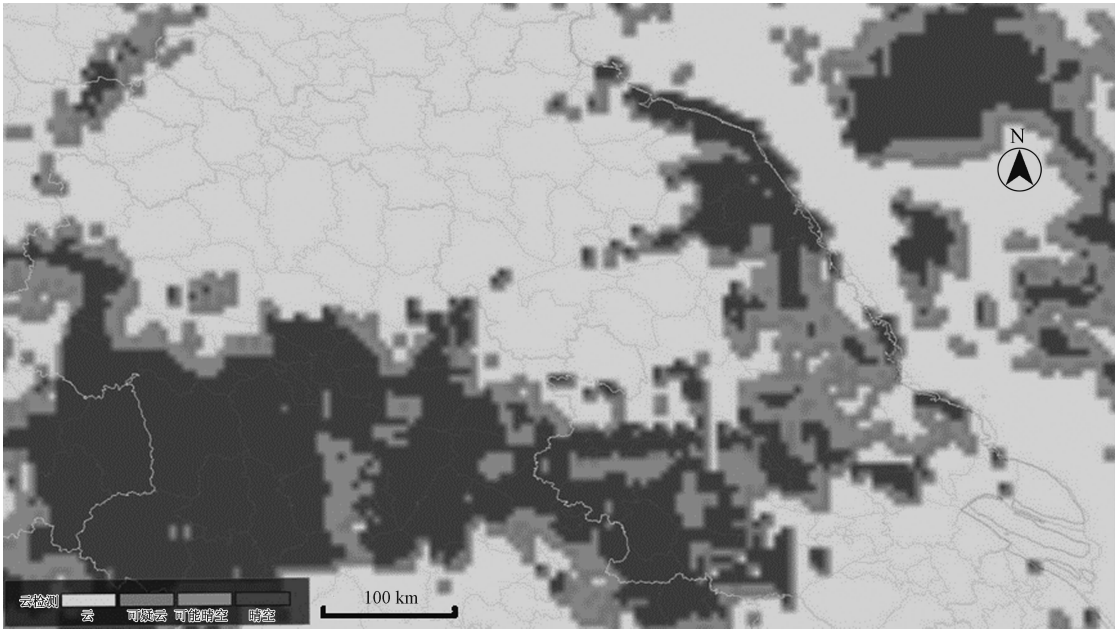


图 4 区域天气状况

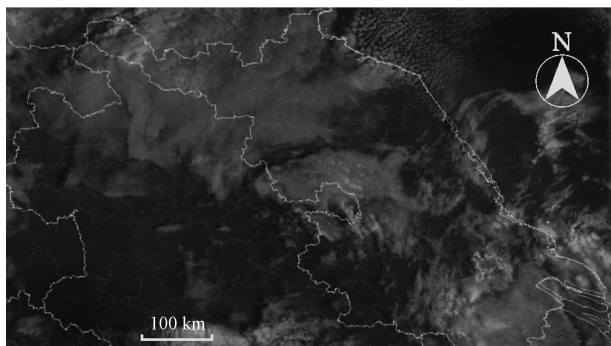


图 5 区域气象条件

论,建立光伏发电预测的 DBN 模型,预测未来短期光伏发电量的概率分布^[11]。牛东晓等提出基于数据挖掘预处理的支持向量机预测系统,进一步提升短期预测精度^[12]。高阳等提出基于发电功率历史数据的挖掘模型,利用特征分析及小波分析结果训练 SVM 分类器,结合发电功率曲线类型从自回归与滑动平均模型和神经网络模型中选取合适的方法对光伏发电功率进行预测,实现短期预测精度提升^[13]。

作为一种特殊的 RNN 模型,长短期记忆网络通过自身特殊的结构设计,有效规避了常规 RNN 训练过程中的梯度消失和梯度爆炸问题^[14],能够提高训练效率,更加有效地利用历史序列信息^[15]。目前,LSTM 已经在自然语言建模、神经网络翻译、图像分析等领域取得了较多应用,在电力负荷预测领域也有一定实践^[16]。基于上述分析,本文探索构建优化 LSTM 光伏功率预测模型,优化了长短期记忆网络学习算法,突出单纯时间序列预测学习功能,对模型学习率和隐藏层神经元个数进行参数优化,进一步提高了长短期记忆网络的预测性能,预测江苏地区典型县域内光伏发电功率,提出基于短期光伏功率预测结果的整县光伏运营模式。

2.2 LSTM 网络结构

LSTM 即长短期记忆递归神经网络,可实现对历史数据的动态传递,反映长期历史数据中具有传递价值的信息,体现短期历史数据中对当前状态具有影响的信息,在发电功率预测等长时间序列中有较好预测效果。

如图 6 所示,LSTM 网络由输入层、输出层和隐藏层构成。相对于传统 RNN 结构,LSTM 的隐藏层是具有特殊记忆模式的 LSTM 单元而非单一神经元。该模型中,记忆单元中的输入门用于控制当前时刻的信息输入,遗忘门用于更新单元的历史

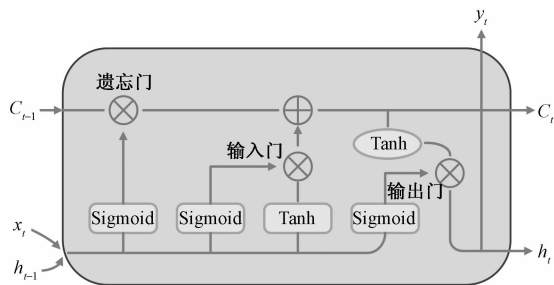


图 6 LSTM 模型基本结构

状态,输出门用于控制信息的输出。

在遗忘门中,LSTM 单元通过输入接口接收当前状态 x_t 和上一时刻模型的隐藏状态 h_{t-1} ,模型接收前一单元信息,即记忆单元的状态 C_{t-1} 。

在输入门中,每一个门将对不同来源的输入进行运算,并且由其逻辑函数决定其是否激活^[17],输入信息经过非线性函数的变换后,与遗忘门处理过的记忆单元状态进行叠加,形成新的记忆单元状态 C_t 。

在输出门中,通过 Sigmoid 层确定输出哪一部分信息,通过 tanh 对模型状态进行 $[-1,1]$ 归一化,并与 Sigmoid 层输出相乘,形成输出信息 h_t, y_t 。模型基本计算方法为

$$\begin{cases} f_t = \text{sigmoid}[W_f(h_{t-1}, x_t) + b_f] \\ i_t = \text{sigmoid}[W_i(h_{t-1}, x_t) + b_i] \\ \tilde{C}_t = \tanh[W_C(h_{t-1}, x_t) + b_C] \\ C_t = f_t C_{t-1} + i_t \tilde{C}_t \\ o_t = \text{sigmoid}[W_o(h_{t-1}, x_t) + b_o] \\ h_t = o_t \tanh(C_t) \\ y_t = \text{sigmoid}(W h_t) \end{cases} \quad (2)$$

2.3 模型预测流程

预测需要收集一定时间范围内屋顶光伏发电功率每日 96 点数据,为缩小数据绝对差异,避免模型神经元饱和,预测过程中应对光伏发电功率数据进行归一化,归一化极值采用历史数据内功率最高点及最低点,将已有功率数据映射到 $[0,1]$ 区间内。归一化过程公式为

$$x = x'(x_{\max} - x_{\min}) + x_{\min} \quad (3)$$

式中: x' 为归一化后某点功率值; x 为某点功率实际值; $x_{\max}$ 为历史数据中功率最大值; $x_{\min}$ 为历史数据中功率最小值。

预测过程中,需要确定输入时间步长、输入维数、隐藏层数量等参数。输入时间步长即用于预测的时间序列长度,该参数的设置主要考虑历史数据的准确性、代表性,不精确或反常历史数据输

人将降低模型预测精度,过长时间序列信息输入也将增加模型训练复杂程度,影响预测效果;输入维数即输入变量个数,在本研究中设定为 1;隐藏层数量即 LSTM 模型层的数量,适量增加隐藏层数量可以提高模型对训练集数据的拟合精度,过量隐藏层将延长模型运行速度,无法提高预测精度。

研究利用 LSTM 提取训练集中的输入时间序列与输出功率之间的非线性关系,建立基于长短期记忆递归神经网络光伏功率预测模型,并选取有更高运算效率的按时间展开的反向误差传播算法进行网络训练。对 LSTM 网络进行展开处理,运用 BP 算法即误差反向传播算法进行训练运算,在多次运用链式规则后形成训练结果,检查测试集预测效果,动态调整模型参数实现较高拟合水平,形成最优模型结构,输出预测值。

预测模型输出的预测值为[0,1]区间内数据,需要对输出值进行反归一化,得到最终预测结果,反归一化过程公式如式(3)所示。

为衡量预测模型性能,采用平均绝对百分比误差刻画不同预测方法与实际值的区别,该衡量方法可对各时间点下预测值与实际值的偏离情况进行分别刻画并汇总,综合反映预测区间内预测精度。

$$M = \frac{1}{n} \sum_{i=1}^n \left| \frac{x - x_i}{x_i} \right| \tag{4}$$

式中: M 为平均绝对百分比误差; n 为预测数据点数量; x_i 为实际值; x 为预测值。

3 算例分析

3.1 实验设计和说明

为验证上文提出预测模型可用性,采用江苏某园区建筑屋顶光伏约 2 年发电功率数据,该园区屋顶光伏装机容量 0.9 MWP,发电功率数据采样间隔为 15 min,共计 74 898 个数据点。为验证优化 LSTM 模型特性,设立灰色预测、BP 神经网络预测对照组,预测 5 d 内单个站点分布式光伏发电功率情况。光伏发电功率数据经过范围验证,进行了异常值处理,经过数据归一化后,设置 480 数据点作为测试集,其余数据作为人工智能模型训练集。

模型训练选取工作站平台,处理器为 Intel Xeon E5,内存为 32 G,硬盘容量 500 G,显卡为 NVIDIA RTX A2000,采用 MATLAB 2021b 软件进行模型计算。

3.2 预测过程及性能分析

分布式光伏功率预测过程中,试验通过控制变量,逐步进行模型调参寻优,通过增加 LSTM 网络层数提升预测效果,调参结果表明,当模型层数由 2 逐步增加至 7 时,模型预测能力上升,但超过 7 以后误差率上升,即表现为过度学习。

在确定模型结构参数后,使用历史光伏发电功率数据对模型进行训练测试,并采用传统 BP 神经网络模型及灰色预测模型进行对比测试。

如表 2、表 3 所示,在 1 d、7 d 预测中,改进 LSTM 模型预测精度明显优于 BP 神经网络方法和灰色预测方法,具有较高预测可行性。

如图 7 所示,3 种预测方法均对未来 7 d 光伏发电功率进行了较好趋势预测,其中改进 LSTM 算法预测精度最高,与真实值贴合紧密,LSTM 在尖峰预测、短时波动预测中表现仍有一定提升空间,需要增加训练数据量,改进模型对历史气象条件变化的学习。

表 2 不同预测方法对不同长度光伏发电功率的平均绝对百分比误差

模型	1 d	7 d
LSTM	0.028 3	0.104 0
BP 神经网络	0.043 1	0.131 0
灰色预测	0.180 7	0.564 1

与 BP 神经网络预测相比,LSTM 预测方法对于光伏发电功率变化趋势判断更精确,7 d 累计平均绝对百分比误差降低 3.3%,较灰色预测方法,LSTM 对光伏发电功率水平、多云阴天等气象条件下光伏减功率运行趋势有更好预测效果,7 d 累计平均绝对百分比误差降低 46.01%,说明运用改进 LSTM 预测模型对分布式光伏功率预测具有较高可行性。

3.3 整县光伏功率预测

基于上述光伏预测经验,本文提出基于优化 LSTM 的整县大规模分布式光伏综合功率预测模型,对分布在县域内不同区域的分布式光伏站点进行总体功率预测,规避以往预测中模型气象参数不适配分布式光伏运行状态的问题,通过分布式数据和分布式预测提升单一光伏站点预测精度,对逐组预测数据进行加总,减去线路损耗及其他损耗后,形成整县光伏功率 96 点预测值曲线,计算公式为

$$P_t = \sum_{i=1}^n p_{it} - \sum_{i=1}^n \Delta P(I,R) - \varphi \tag{5}$$

表 3 不同预测方法在发电功率阶段的预测误差对比

时间	真实值/kW	LSTM		BP		灰色预测	
		预测值/kW	绝对百分比误差/%	预测值/kW	绝对百分比误差/%	预测值/kW	绝对百分比误差/%
2022-09-30T19:15	0.372	0.335	9.946	0.363	2.419	0.301	19.086
2022-09-30T19:00	5.829	5.632	3.380	5.845	0.274	4.872	16.418
2022-09-30T18:45	19.325	18.962	1.878	19.637	1.614	15.380	20.414
2022-09-30T18:30	41.331	40.939	0.948	41.725	0.953	29.690	28.165
2022-09-30T18:15	69.538	69.348	0.273	70.062	0.754	47.855	31.182
2022-09-30T18:00	100.337	100.795	0.456	101.377	1.037	68.771	31.460
2022-09-30T17:45	133.361	134.568	0.905	134.871	1.132	90.633	32.039
2022-09-30T17:30	170.177	171.831	0.972	172.218	1.199	115.229	32.289
2022-09-30T17:15	208.829	209.963	0.543	210.008	0.565	142.109	31.950
2022-09-30T17:00	247.784	247.676	0.044	247.747	0.015	170.264	31.285
2022-09-30T16:45	286.712	285.288	0.497	285.548	0.406	200.198	30.175
2022-09-30T16:30	322.295	320.445	0.574	318.047	1.318	230.912	28.354
2022-09-30T16:15	353.832	352.786	0.296	344.880	2.530	255.409	27.816
2022-09-30T16:00	382.576	383.167	0.154	368.374	3.712	272.993	28.643
2022-09-30T15:45	407.479	410.567	0.758	388.300	4.707	290.036	28.822
2022-09-30T15:30	429.062	435.258	1.444	405.259	5.548	307.050	28.437
2022-09-30T15:15	452.831	461.284	1.867	430.980	4.825	322.822	28.710
2022-09-30T15:00	477.277	487.350	2.111	464.632	2.649	336.566	29.482
2022-09-30T14:45	497.041	509.197	2.446	495.962	0.217	349.401	29.704
2022-09-30T14:30	510.536	526.232	3.074	522.420	2.328	361.644	29.164
2022-09-30T14:15	513.122	531.125	3.509	533.377	3.947	355.922	30.636
2022-09-30T14:00	515.550	532.304	3.250	541.745	5.081	333.426	35.326
2022-09-30T13:45	514.069	529.253	2.954	544.730	5.964	310.633	39.574
2022-09-30T13:30	513.656	524.367	2.085	547.265	6.543	292.104	43.132
2022-09-30T13:15	514.725	515.280	0.108	550.798	7.008	279.656	45.669
2022-09-30T13:00	521.407	505.063	3.135	559.787	7.361	272.518	47.734
2022-09-30T12:45	530.254	494.044	6.829	568.950	7.298	266.127	49.811
2022-09-30T12:30	533.026	481.881	9.595	571.742	7.263	262.150	50.819
2022-09-30T12:15	524.504	464.985	11.348	562.407	7.226	247.069	52.895
2022-09-30T12:00	505.484	440.844	12.788	540.465	6.920	220.188	56.440
2022-09-30T11:45	479.940	411.415	14.278	512.330	6.749	195.962	59.169
2022-09-30T11:30	454.090	383.119	15.629	485.212	6.854	174.952	61.472
2022-09-30T11:15	417.806	350.340	16.148	451.199	7.992	162.927	61.004
2022-09-30T11:00	376.816	318.242	15.544	414.786	10.077	159.465	57.681
2022-09-30T10:45	335.518	286.353	14.653	377.768	12.592	157.503	53.057
2022-09-30T10:30	295.238	254.728	13.721	340.056	15.180	157.544	46.638
2022-09-30T10:15	254.619	221.581	12.975	300.957	18.199	146.821	42.337
2022-09-30T10:00	215.209	189.926	11.748	263.967	22.656	127.150	40.918
2022-09-30T09:45	177.842	160.892	9.531	227.046	27.667	109.053	38.680
2022-09-30T09:30	143.537	132.980	7.355	186.859	30.182	91.606	36.180
2022-09-30T09:15	116.147	108.318	6.741	149.313	28.555	75.957	34.603
2022-09-30T09:00	93.652	86.717	7.405	117.699	25.677	61.801	34.010
2022-09-30T08:45	71.238	65.476	8.088	88.459	24.174	47.841	32.843
2022-09-30T08:30	48.865	44.969	7.973	60.467	23.743	34.227	29.956
2022-09-30T08:15	28.739	26.940	6.260	35.169	22.374	21.735	24.371
2022-09-30T08:00	13.692	12.858	6.091	15.693	14.614	11.047	19.318
2022-09-30T07:45	3.922	3.621	7.675	4.286	9.281	3.192	18.613
2022-09-30T07:30	0.075	0.074	1.333	0.078	4.000	0.054	28.000

式中： P_t 为某时刻下整县光伏功率预测值，整县共有 n 个光伏电源； p_i 为第 i 个光伏电源在 t 时刻的功率预测值； $\sum \Delta P(I,R)$ 为线路中线损； φ 为其他损耗。

在整县光伏综合功率预测过程中，主要对 42 个光伏电源数据进行了以下工作：

1)数据筛选。分析各个光伏电源已有历史数据，提取具有共同采样时点的数据区间，根据预测

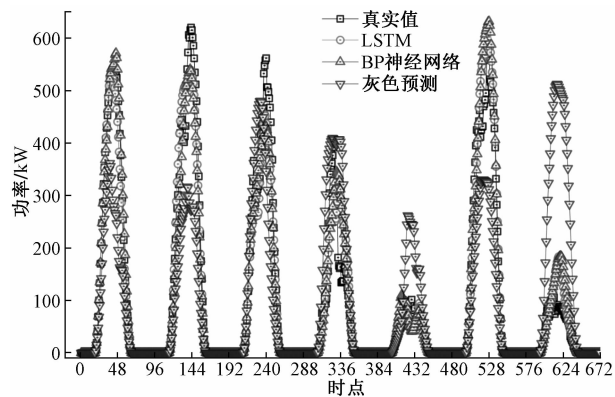


图 7 不同预测方法的 7 d 预测结果对比

时段需求确定输入序列长度。

2) 数据处理。对拟输入数据检查其一致性, 填补无效值和缺失值, 进行功率数据归一化。

3) 预测模型构建及优化。对 42 个光伏电源进行逐一预测建模, 根据不同模型预测特点调整参数, 提高预测效率精度, 得到各光伏电源未来功率预测值, 进行数据反归一化后检验数据可靠性。

4) 结果汇总。对经过可靠性检验的单个光伏电源功率预测数据进行加总, 形成整县光伏功率综合预测结果。

为验证基于优化 LSTM 的整县大规模分布式光伏综合功率预测模型的有效性, 本研究收集某县内 42 个分布式光伏电源历史发电数据, 光伏站点所处位置包括开阔地带屋顶、鱼塘水面、乔木林地、住宅楼外立面, 具有不同微气象条件, 反映了真实环境下整县分布式光伏运营场景。

在对 42 个光伏电源进行综合功率综合预测后, 各个光伏电源功率预测曲线如图 8 所示。在整县光伏综合功率预测过程中, 通过对 42 个单一光伏电源预测序列的加总, 可以规避以往其他预测方法中因

模型本身误差造成的整体预测问题, 本文通过分布式预测和预测结果加和, 分摊了各模型中的正向及负向误差, 预测效果全面优于单一光伏状态, 进一步提高了整体功率预测精度。

如表 4、图 9 所示, 基于改进长短期记忆递归神经网络的整县光伏综合功率预测模型取得较好预测

表 4 整县光伏综合功率预测部分时点结果对比分析

时点	整县光伏功率 预测值/kW	整县光伏功率 实际值/kW	绝对误差
8:00	1 163.32	1 167.98	0.004 0
9:00	7 302.67	7 346.49	0.006 0
10:00	17 265.15	17 455.07	0.010 9
11:00	24 751.09	25 765.89	0.039 4
12:00	28 318.49	28 799.91	0.016 7
13:00	28 730.79	27 725.21	0.036 3
14:00	20 555.57	21 377.80	0.038 5
15:00	18 875.69	18 403.80	0.025 6
16:00	14 913.41	14 406.36	0.035 2
17:00	11 575.52	11 147.23	0.038 4
18:00	5 270.32	5 296.67	0.005 0
19:00	370.70	357.72	0.036 3
20:00	0.00	0.00	0.000 0
8:00	807.87	810.30	0.003 0
9:00	6 692.32	6 812.78	0.017 7
10:00	13 351.99	13 859.36	0.036 6
11:00	22 888.25	22 682.26	0.009 1
12:00	28 291.67	28 037.04	0.009 1
13:00	27 212.05	26 041.93	0.044 9
14:00	25 889.76	24 854.17	0.041 7
15:00	21 349.40	21 135.91	0.010 1
16:00	16 856.67	16 081.26	0.048 2
17:00	12 192.54	12 168.16	0.002 0
18:00	4 863.66	5 097.11	0.045 8
19:00	300.56	295.15	0.018 3
20:00	0.00	0.00	0.000 0

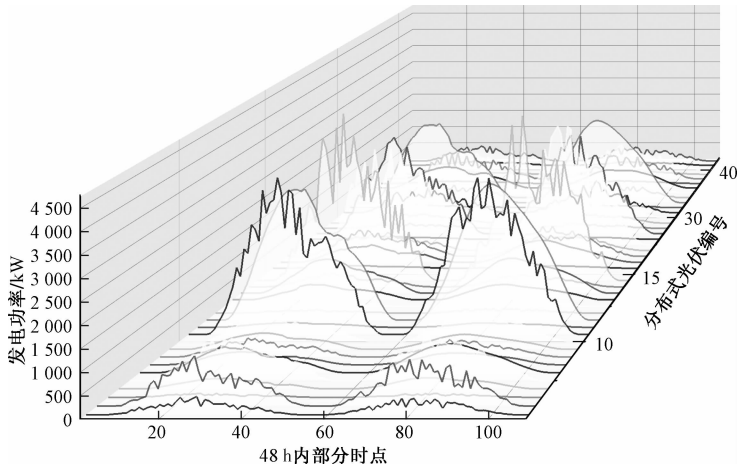


图 8 48 h 内部分时点各个光伏电源发电功率综合预测结果

效果,功率预测值与实际发电功率匹配度高,功率单调变化区间预测平均绝对百分比误差仅为 2.05%,预测模型对第 1 天、第 2 天尖峰功率预测误差仅为 1.46%和 4.27%,较好反映了未来光伏发电功率变化的增减规律、峰值区间、波动趋势,能够有效支撑区域电网提前调度。

如表 5、图 9 所示,整县光伏综合功率预测模型两天内预测平均绝对百分比误差仅为 2.32%,较单一电源预测精度提高 2.86%,明显提升了县域内大规模分布式光伏规模预测精准度。基于优化 LSTM 模型的整县光伏功率预测曲线与实际发电功率匹配度高,功率单调变化区间预测平均绝对百分比误差仅为 2.05%,对第 1 天、第 2 天尖峰功率预测误差仅为 1.46%和 4.27%,较好反映了未来光伏发电功率变化的增长规律、峰值区间、波动趋势、减少规律,能够有效支撑区域电网提前调度。

表 5 单一光伏预测与整县光伏综合预测结果误差对比

类别	误差/%			
	48 h	单调变化区间	第 1 天尖峰	第 2 天尖峰
单一光伏	5.18	4.39	1.61	4.58
整县光伏	2.32	2.05	1.46	4.27

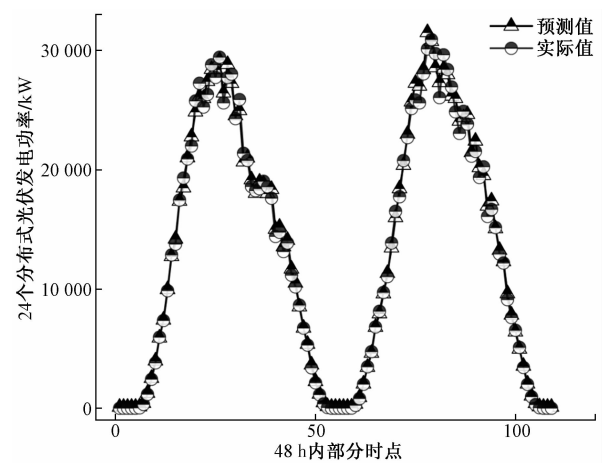


图 9 48 h 内部分时点整县光伏综合预测功率曲线与实际功率曲线

4 总结与展望

研究对大规模分布式光伏整县推进背景下江苏省典型光伏电站发电影响因素进行梳理,对省内不同县域同时刻气象数据差异进行分析,采用时间序列功率数据预测方法,提出改进 LSTM 光伏功率预测模型,减少模型训练时长及算力需求,改善了

模型对历史信息学习效果,提升了 LSTM 模型对单一光伏电源的功率预测精度,进一步提出了基于改进 LSTM 的整县光伏综合功率预测模型,结合整县光伏特点,对单一分布式光伏电源预测结果进行计算,形成整县光伏功率综合预测方法。在单一光伏电源和整县 42 个光伏电源中进行模型应用分析,模型均实现较高预测精度,在整县光伏预测环境下,改进 LSTM 模型实现 48 h 192 点实现 2.32%平均绝对误差,更好满足基于光伏功率预测的系统资源调度需求。

未来,基于精准大规模分布式光伏功率综合预测模型,现有零散小规模光伏电源可以更好融入现有县域综合能源系统,与电网、其他电源、储能、充电桩等设施相结合,形成以分布式光伏为骨干的区域电网,通过实时数据传输和功率预测,为需求响应等可调资源提供更精确综合发电功率曲线,实现域内的短时精确多能互补,解决县域内配电网全局协调控制问题,降低控制维度和复杂度,发挥虚拟电厂源网荷储高效调控的优势,进一步促进分布式光伏高效消纳。

参考文献

[1] 庞龙光,周靖仁,王树欢. 先行先试 当好样板:河北保定供电公司助力分布式光伏整县推进[J]. 国家电网,2021(10):50-51.

[2] 苏剑,周莉梅,李蕊. 分布式光伏发电并网的成本/效益分析[J]. 中国电机工程学报,2013,33(34):50-56,11.

[3] 崔红芬,汪春,叶季蕾,等. 多接入点分布式光伏发电系统与配电网交互影响研究[J]. 电力系统保护与控制,2015,43(10):91-97.

[4] 龚莺飞,鲁宗相,乔颖,等. 光伏功率预测技术[J]. 电力系统自动化,2016,40(4):140-151.

[5] 丁明,王磊,毕锐. 基于改进 BP 神经网络的光伏发电系统输出功率短期预测模型[J]. 电力系统保护与控制,2012,40(11):93-99,148.

[6] 栗然,李广敏. 基于支持向量机回归的光伏发电功率预测[J]. 中国电力,2008(2):74-78.

[7] OSTERAGOVA E, OSTERTAG O. Forecasting using simple exponential smoothing method[J]. Acta Electrotechnica et Informatica,2012,12:62-66.

[8] 潘志远,韩学山. 电网节点负荷的立体化预测方法[J]. 电力系统自动化,2012,36(21):47-52.

[9] 王勇,黄国兴,彭道刚. 带反馈的多元线性回归法在电力负荷预测中的应用[J]. 计算机应用与软件,2008(1):82-84.

[10] 李震,张思,任娴婷,等. 基于数据驱动的线性聚类 ARIMA 长期电力负荷预测[J]. 科学技术与工程,2020,20(16):6497-6504.

[11] 董雷,周文萍,张沛,等. 基于动态贝叶斯网络的光伏发电短期概率预测[J]. 中国电机工程学报,2013,33(S1):38-45.

[12] 牛东晓,谷志红,邢棉,等. 基于数据挖掘的 SVM 短期负荷预测方法研究[J]. 中国电机工程学报,2006(18):6-12.

[13] 高阳,张碧玲,毛京丽,等. 基于机器学习的自适应光伏超短期功率预测模型[J]. 电网技术,2015,39(2):307-311.

[14] 杨丽,吴雨茜,王俊丽,等. 循环神经网络研究综述[J]. 计算机应用,2018,38(S2):1-6,26.

[15] 陆继翔,张琪培,杨志宏,等. 基于 CNN-LSTM 混合神经网络模型的短期负荷预测方法[J]. 电力系统自动化,2019,43(8):131-137.

[16] 陈亮,王震,王刚. 深度学习框架下 LSTM 网络在短期电力负荷预测中的应用[J]. 电力信息与通信技术,2017,15(5):8-11.

[17] 程宇也. 基于人工神经网络的短期电力负荷预测研究[D]. 杭州:浙江大学,2017.

Study on Comprehensive LSTM Power Prediction Model of Photovoltaic Power Generation in Whole County

CANG Min, ZHAI Xiaomeng, WANG Jingyi, WU Shuang, CHENG Xi

(Economic and Technological Research Institute, State Grid Jiangsu Electric Power Co., LTD., Nanjing 210028, China)

Abstract: In order to better realize the whole county photovoltaic roof intensive development, improve the economics of large-scale distributed photovoltaic grid roof, analyze disadvantages of prediction method based on the meteorological data, explore the rule of the massive time-series photovoltaic power data, based on improved both short-term and long-term memory, recursive neural network (LSTM) of the whole county comprehensive photovoltaic power prediction model is proposed. The model is used to predict a single photovoltaic power supply. The results show that the improved LSTM model is more accurate than other models, and it has higher overall prediction accuracy under the scenario of the whole county photovoltaic comprehensive power prediction. It can be used to predict the whole county photovoltaic power generation power, which is helpful to achieve a high proportion of distributed photovoltaic consumption.

Keywords: county photovoltaic; distributed photovoltaic; long and short term memory network; power prediction