

改进神经网络模型在光伏发电预测中的应用^①



胡克用, 李云龙, 江霞, 李静, 胡则辉

(杭州师范大学钱江学院, 杭州 310018)

通讯作者: 胡克用, E-mail: hukeyong@yeah.net

摘要: 针对传统的 BP 神经网络模型在光伏发电功率预测中, 存在着的预测精度不高、收敛速度较慢的弊端, 本文提出一种改进型的 BP 神经网络模型, 运用增加动量项以及自适应选取最佳隐含层的方法来改进传统 BP 模型的缺陷. 文中首先进行了各类气象因素对于光伏发电功率输出影响的相关性分析, 提取出最能影响光伏发电功率的 6 个气象因素, 作为网络模型的输入, 然后建立了改进型的 BP 网络模型, 结合光伏功率输出的历史数据, 来进行发电数据的直接预测; 最后根据不同气候类型下的预测结果, 分析并验证了采用该模型进行功率预测的可行性和有效性.

关键词: 光伏发电; 相关性分析; 气象因素; 神经网络

引用格式: 胡克用, 李云龙, 江霞, 李静, 胡则辉. 改进神经网络模型在光伏发电预测中的应用. 计算机系统应用, 2019, 28(12): 37-46. <http://www.c-s-a.org.cn/1003-3254/6854.html>

Application of Improved Neural Network Model in Photovoltaic Power Generation Prediction

HU Ke-Yong, LI Yun-Long, JIANG Xia, LI Jing, HU Ze-Hui

(Hangzhou Normal University Qianjiang College, Hangzhou 310018, China)

Abstract: In view of the traditional BP neural network model used in Photo Voltaic (PV) power generation prediction, there are shortcomings of low prediction accuracy and slow convergence speed. In this study, an improved BP neural network model is proposed to improve the defects of traditional BP model by adding momentum term and adaptively selecting the best hidden layer. Firstly, the correlation of meteorological factors for PV power output is analyzed, and six meteorological factors that can affect the PV power are extracted as input of the network model. Then, an improved BP network model is established, combined with the historical data of PV power output, to directly predict the generation of data. Finally, according to the prediction results under different climate types, the feasibility and effectiveness of the model for power prediction are analyzed and verified.

Key words: Photo Voltaic (PV) power generation; correlation analysis; meteorological factors; neural network

现如今随着社会的建设和人口的增长, 环境恶化、资源减少的问题愈发严重, 世界各国努力寻求开发新型能源, 对太阳能的利用成了解决人类可持续发展问题的曙光. 光伏发电是将太阳能直接转变成电能的一种发电方式, 其具有转化率高且安装灵活、便于

维护的优势. 但研究表明, 光伏发电系统不同于常规能源发电方式具有连续可调控的特点, 会跟随太阳辐射强度和温度的变化而具有波动性和间歇性, 大规模光伏发电系统的并网运行会对传统电网的安全、稳定造成较大影响^[1-4]. 因此, 对光伏系统的输出功率进行准

① 基金项目: 浙江省自然科学基金 (LY17E070004), 浙江省教育科学规划研究课题 (2018SCG015), 浙江省高校实验室工作研究项目 (YB201904), 杭州师范大学钱江学院科学研究资助项目 (2019QJL05), 国家自然科学基金 (61702151)

Foundation item: Natural Science Foundation of Zhejiang Province (LY17E070004); Educational Science Planning Research Project of Zhejiang Province (2018SCG015); Laboratory Work Research Project of Higher Education of Zhejiang Province (YB201904); Scientific Research Program of Hangzhou Normal University Qianjiang College (2019QJL05); National Natural Science Foundation of China (61702151)

收稿时间: 2018-10-17; 修改时间: 2018-11-06; 采用时间: 2018-11-19; csa 在线出版时间: 2019-12-10

确预测就显得非常重要,首先可以降低发电系统接入电网的负面影响,保证电网的安全性和稳定性,同时可以充分利用太阳能资源,达到经济效益最大化的目标;其次,精准的预测能为电力系统调度部门实施合理的调度计划提供参考意见。

目前对于光伏系统发电的预测方法主要有直接预测和间接预测2种。直接预测法本质上都是统计方法,是将气象因素或历史数据作为输入量建立模型从而对数据算法进行训练,最终通过模型计算出预测结果^[5];间接预测是先根据天气状况对太阳辐射强度进行预测,然后根据发电站的电功率模型进一步预测输出功率。发电系统输出功率的预测是电力行业持续不断研究的重点,它是电网安全运行和技术进步的基础条件。国内外许多研究者都针对不同负荷类型的电力系统进行探究和应用。主要的预测方法有:线性回归法、人工神经网络法、卡尔曼滤波法、时间序列法等。

文献[6]通过对温度、风向、云量、湿度等气象因素对光伏发电系统影响的分析,利用BP神经网络建立预测模型,实现对其光伏输出功率的预测。缺点是算法容易陷入局部极小值,造成结果准确率低。

文献[7]是基于马尔科夫链的光伏发电功率预测模型,先收集统计光伏发电系统历史输出功率数据,然后采用其搭建的模型,直接输出预测光伏发电功率,然而不太适宜用于系统的中长期预测。

文献[8]是结合了近邻传播聚类和回声状态网络2种算法的优点,先采用近邻传播算法对光伏发电分类,然后基于预测日所属的类别建立回声网络状态方程来预测光伏发电功率,然而回声网络中随机生成的储备池与具体问题关联性低,并且参数较难以确定。

文献[9]提出一种自适应的光伏超短期出力预测模型,利用光伏发电历史数据的小波分析以及特征分析结果来训练支持向量机分类器,从而预测发电数据,然而小波基的选取要求高,加大了实验的难度。

文献[10]是利用支持向量机和神经网络,同时结合天气种类以及历史气象数据,基于各类天气状况分别建模或引入日类型指数来进行预测,然而支持向量机在解决多分类问题上存在困难。

文献[11]采用前向反馈神经网络(FFNN)、径向基神经网络(RBFNN)和回归神经网络(RNN)3种神经网络算法对辐射强度进行预测。经过数据分析对比,最后选择了RNN对光伏出力进行间接预测。

文献[12]先分析了影响光伏发电系统输出功率的各项因素,并加入天气信息利用人工神经网络法对其输出功率进行预测。结果能够一定程度上解决光伏发电系统随机性问题,表明了模型的可行性,然而在训练中容易陷入局部极小值而得不到最优解。

经过对比分析,间接预测的结果精度较高,但需要建立复杂的预测模型和太阳辐射强度测量装置,其中光伏阵列的安装角度以及转换效率也是必要的考虑因素。而直接预测的过程就较为简便,但若要达到较高精度的预测结果,需要对历史输出功率以及相关气象数据进行科学地选取,同时要对算法模型进行改进优化。经过探究和对比,本文提出一种改进型的BP神经网络短期预测模型,以光伏发电历史数据和相关气象因素作为模型输入量,并且对传统BP神经网络进行了优化,通过增加动量项和自动选取最优隐含层数目使模型避免陷入极小值并且加快了训练速率。经过算例对比验证,该模型具有较高的预测精度以及收敛速度,具有一定的参考意义。

1 光伏发电输出功率影响分析

1.1 相关性分析理论

相关性分析就是指对两个及以上具备相关性的变量元素进行分析,来衡量两个变量因素之间的相关密切程度^[13]。

用来衡量两个变量之间的相互关系,最常用的是皮尔逊积矩相关系数(pearson product-moment correlation coefficient),定义是两个变量协方差除以两个变量的标准差(方差的平方根)。

两个变量之间的皮尔逊相关系数计算如下:

$$\begin{aligned} \text{corr}(X, Y) &= \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E(X - \mu_X)E(Y - \mu_Y)}{\sigma_X \sigma_Y} \\ &= \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1) \end{aligned}$$

式中, $\text{cov}(X, Y)$ 是变量 X, Y 的协方差; σ_X 是变量 X 的标准差; σ_Y 是变量 Y 的标准差; 结果 $r(\text{corr})$ 范围在 $[-1, 1]$ 之间。

(1) 若 $r > 0$, 表明两个变量是正相关, 即当一个变量值增大的同时另一个变量值也会越大;

(2) 若 $r < 0$, 表明两个变量是负相关, 即当一个变量值减小的同时另一个变量值也会越小;

(3) 若 $r=0$, 表明两个变量间不是线性相关。

r 的绝对数值反映了相关性的 大小, 通常情况下: 当 r 的范围在 $0.00\sim\pm 0.30$ 之间时, 两变量微相关; 当 r 的范围在 $\pm 0.30\sim\pm 0.50$ 之间时, 两变量实相关; 当 r 的范围在 $\pm 0.50\sim\pm 0.80$ 之间时, 两变量显著相关; 当 r 范围在 $\pm 0.80\sim\pm 1.00$ 之间时, 两变量高度相关。

在本文的研究中我们基于以上相关理论使用 Matlab 中的 $corr(X, Y)$ 函数, 来计算皮尔逊相关系数。

1.2 气象因素与光伏系统输出功率的相关性分析

影响光伏系统输出功率的因素众多, 如逆变转换效率、太阳辐射强度、气压、太阳能电池板安装角度、温度、风速、湿度等^[14,15], 并且这些因素与光伏发电输出功率之间也有着复杂的关联。如果将所有因素都作为神经网络预测模型的输入, 光伏发电系统历史数据的收集整理步骤难度必定增加, 且一定程度上影响了预测的准确度。其中气象因素主要包括温度、湿度、风速、云量、气压等, 它们都具有明显的不连续性和不确定性, 而且与光伏发电系统输出功率都有着紧密的联系。为了选择相对合适的输入模型, 我们需要运用到相关性分析来甄别气象因素中对光伏发电系统输出功率的影响程度。

本文选取澳大利亚的 Desert Knowledge Australia Solar Center (DKASC, 纬度: -23.76 , 经度: 133.87) 的一处 10.5 KW 太阳能光伏发电系统作为研究对象, 采集其 2015 年 1 月、4 月、7 月、10 月的日发电量与数个气象因素历史数据分别进行相关性分析。DKASC 地理位置如图 1 所示。



图 1 DKASC 地理位置

1.2.1 太阳辐射强度和光伏发电功率的相关性分析

由于单位面积光伏面板的输出功率, 由式 (2)^[16] 表示:

$$P_s = \eta S I [1 - 0.005 \times (t_0 + 25)] \quad (2)$$

其中, P_s 是光伏面板的输出功率; η 是光伏面板的转换效率; S 是光伏阵列的面积; I 是光伏面板斜面上的太阳辐射强度; ϕ 是环境温度。

由上式可以看出, 光伏面板斜面上的太阳辐射强度与输出功率成正比, 也就是说每日光伏面板斜面上所吸收的辐射量越大, 发电量就越多。为了验证太阳日辐射强度总量与光伏发电功率两个变量有着较强的相关性, 本文对二者做了相关性分析, 如表 1 所示。

表 1 太阳日辐射强度总量与光伏发电功率的相关性分析

月份 (月)	相关系数
1	0.9932
4	0.7944
7	0.8901
10	0.9786

由表 1 中 4 个月的太阳日辐射强度总量与光伏发电输出功率的相关性分析数据可知, 2015 年 1 月与 10 月的相关性系数均在 0.95 以上, 由于 4 月与 7 月澳大利亚正值秋冬季, 太阳辐射强度略低于 1 月和 10 月, 但也都达到了显著相关的标准, 由此证明了二者之间有着很强的相关性。

1.2.2 湿度与光伏发电功率的相关性分析

湿度是衡量空气中水蒸汽含量的指标, 水蒸汽会对太阳辐射有着反射和吸收的作用, 间接影响了光伏系统所接收到的太阳辐射值, 最终关联着光伏系统的输出功率, 文中所收集到的数据来源于澳大利亚天气预报网站中的相对湿度。

表 2 反映了最大湿度、最小湿度、平均湿度与光伏系统输出功率的相关系数。

表 2 湿度与光伏发电功率的相关性分析

月份 (月)	最大湿度 H_{\max}	最小湿度 H_{\min}	平均湿度 H_{avg}
1	-0.5961	-0.8654	-0.8484
4	-0.6434	-0.8344	-0.8158
7	0.1928	-0.5235	-0.2942
10	-0.1790	-0.5690	-0.4831

由表 2 中的 4 个月的相对湿度与光伏发电输出功率的相关性分析数据可知, 两个变量之间成负相关, 说明湿度越大则输出功率越小; 同时最小湿度 H_{\min} 与平均湿度 H_{avg} 的相关系数值较高, 基本上都呈显著相关状态, 但与最大湿度 H_{\max} 的相关系数就较弱。

1.2.3 温度与光伏发电功率的相关性分析

通常温度的上升会使光伏发电系统的转换效率下降,从而导致其输出功率也呈现减小的趋势.

表3反映了最高温度、最低温度、平均温度与光伏系统输出功率的相关系数.

月份(月)	最高温度 H_{max}	最低温度 H_{min}	平均温度 H_{avg}
1	0.7517	0.0394	0.6978
4	0.4353	-0.0896	0.2142
7	-0.0779	-0.6486	-0.2815
10	0.1455	-0.3184	0.0268

由表3中4个月的温度与光伏发电输出功率的相关性分析数据可知,两个变量出现负相关的情况主要在冬季,这是因为太阳辐射量对发电功率的影响效果要远大于温度,阴雨天的太阳辐射量较晴天要少,一般情况下冬季太阳辐射总量偏低,温度对光伏发电功率的影响效果才得以显著.2015年1月对于澳大利亚来说,在当地处于夏季,太阳辐射量较大,才使得两变量成正相关.

1.2.4 平均风速与光伏发电功率的相关性分析

风速大小的产生原因相对复杂,与很多气象因素相关联,且一般情况下风速会影响到太阳能光伏发电系统散热.

表4反映了平均风速与光伏发电功率的相关系数.

月份(月)	相关系数
1	0.2768
4	0.1646
7	-0.0221
10	0.4937

由表4中4个月的平均风速与光伏发电输出功率的相关性分析数据可知,两变量在春、夏季相关度相对较高,秋、冬季则均呈微相关,且在7月呈负相关.由此可见,两者之间仍然是有相关性的,为了保证预测模型的准确性,我们仍需将其作为参考量输入预测模型.

1.2.5 气压与光伏发电功率的相关性分析

气压是指在单位面积上作用的大气压力.通常来说,气压基本不会对光伏发电输出功率造成影响.表5反映了气压与光伏系统输出功率的相关系数.

由表5中4个月的气压与光伏发电输出功率的相关性分析数据可知,两变量间的相关系数都在0.1以

下,全部处于微相关.由此可见,气压与光伏发电输出功率的密切程度很弱,因而没有将其作为参考量输入预测模型.

表5 气压与光伏发电功率的相关性分析

月份(月)	相关系数
1	-0.058
4	-0.023
7	0.071
10	-0.051

1.2.6 影响光伏发电输出功率的因素小结

经过以上多种主要气象因素与光伏日发电量的相关性分析,我们了解到影响光伏系统输出功率的因素众多且影响趋势较复杂.与之关系最密切的是温度、相对湿度、和太阳辐射强度,而风速则取决于季节因素.数据显示,我们的光伏发电系统数据采样地,澳大利亚的 Desert Knowledge Australia Solar Center 在每年12月、1月和2月风速影响较强,其它月份相对较低.因此,选取每日太阳日辐射总量,最大、最小以及平均湿度,最大、最小以及平均温度和平均风速,共8种气象因素作为光伏发电系统功率预测模型的输入量.

2 BP神经网络的预测模型设计

2.1 BP神经网络模型结构

神经网络是一种分布式并行信息处理的算法数学模型,现在已广泛应用与模式识别、自动控制、数据挖掘、人工智能等领域^[17].BP (Back Propagation)神经网络作为众多人工神经网络实现方法之一,是迄今为止研究最成熟、应用最广泛的人工神经网络模型之一^[18],它的拓扑结构包括输入层、隐含层及输出层,典型的三层神经网络结构如图2所示.

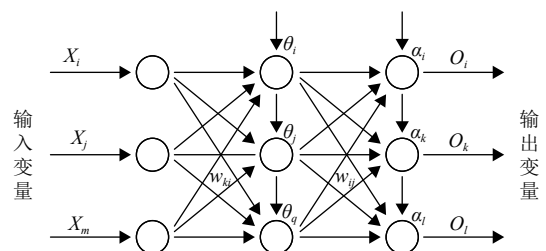


图2 BP神经网络结构图

工作信号在正向传递过程中,各层节点输入与输出的数学表达如下:

隐含层第 i 个节点输入值为:

$$net_i = \sum_{j=1}^M w_{ij}x_j + \theta_i \quad (3)$$

隐含层第 i 个节点输出值为:

$$y = \phi(net_i) = \varphi \left(\sum_{j=1}^M w_{ij}x_j + \theta_i \right) \quad (4)$$

输出层第 k 个节点输入值为:

$$\begin{aligned} net_k &= \sum_{i=1}^q w_{ki}y_i + \alpha_k \\ &= \sum_{i=1}^q w_{ki} \varphi \left(\sum_{j=1}^M w_{ij}x_j + \theta_i \right) + \alpha_k \end{aligned} \quad (5)$$

输出层第 k 个节点输出值为:

$$\begin{aligned} o_k &= \psi(net_k) = \psi \left(\sum_{i=1}^q w_{ki}y_i + \alpha_k \right) \\ &= \psi \left[\sum_{i=1}^q w_{ki} \varphi \left(\sum_{j=1}^M w_{ij}x_j + \theta_i \right) + \alpha_k \right] \end{aligned} \quad (6)$$

式中, w_{ij} 是隐含层中的第 i 个节点到输入层中的第 j 个节点之间的权值; x_j 是输入层第 j 个节点的输入值, 其中 $j=1, 2, 3, \dots, m$; θ_i 是隐含层第 i 个节点的阈值; $\varphi(x)$ 表示隐含层的激励函数; α_k 表示输出层第 k 个节点的阈值, $k=1, 2, \dots, l$; w_{ki} 表示输出层中的第 k 个节点到隐含层中的第 i 个节点之间的权值, $i=1, 2, \dots, q$; $\psi(x)$ 表示输出层的激励函数。

2.2 传统 BP 神经网络算法存在的问题

由于传统 BP 神经网络算法主要利用梯度下降的原理, 按负梯度方向调整权值。所以一般会存在以下问题:

(1) 网络模型训练中容易陷入局部极小值而得不到最优值, 原因在于 BP 神经网络计算非线性问题时, 权值是通过局部改善并逐渐调整的, 这样就容易导致算法陷入局部极值, 同时权值也会收敛至局部极小点。因此, 当初始网络权重不同时, 每次训练网络都会收敛于不同的局部极小值, 并且网络训练得到的结果也有较大差异。

(2) 训练次数较多, 收敛速度慢, 造成学习效率的降低。这是因为神经网络需要优化的目标函数较复杂, 训练过程中当神经元输出接近 0 或 1 的情况下, 权值误差的波动较小, 所以此时训练速度非常缓慢甚至几乎停顿。

(3) 隐含层节点数量选取缺乏一定的理论指导。

2.3 改进 BP 神经网络算法的方法

针对传统 BP 神经网络算法存在的问题, 本文主要从 3 个方面加以改进:

(1) 增加动量项

针对网络模型训练中容易陷入局部极小值而得不到最优值的问题, 提出一种增加动量项的方法, 采用动量因子来传递最后一次的权值变化量。当动量因子为 0 时, 由梯度下降法产生最后一次的权值; 当动量因子为 1 时, 采用动量因子方法而产生的新权值为最后一次的权值变化量。增加动量法实质上等同于在训练过程中加入阻尼项, 降低网络对于误差曲面局部细节的敏感性, 减小学习过程的振荡趋势, 从而避免网络陷于局部极小点, 改善了收敛性。

动量项的修正公式如下:

$$\Delta w_{ij}(k+1) = (1-m)\eta\delta_i p_j + m\Delta w_{ij}(k) \quad (7)$$

$$\Delta b_i(k+1) = (1-m)\eta\delta_i + m\Delta b_i(k) \quad (8)$$

其中, w_{ij} 是隐含层中的第 i 个节点到输入层中的第 j 个节点之间的权值; k 是训练次数; m 是动量因子, 一般 $m \in (0, 1)$; η 是学习步长; δ_i 为在输出层和隐含层中的局部误差信号; p_j 为输入信号在隐含层的传递; b_i 是隐含层学习速率。这样操作的优点在于可以有效防止 $\Delta w_{ij}(k)=0$, 帮助网络优化从误差曲面局部极小值中脱离出来。

(2) 自适应调节学习步长

学习步长 η 在 BP 神经网络算法中是一个常数, 但在算法模型中学习步长是常数不一定都合适。例如在误差曲面处于较为平坦的范围时, 学习步长太小会引起学习次数的增多, 这样就要求适当增大 η ; 然而在极小值附近时, 学习步长大了则会引起振荡, 这样又会要求减小 η 。

鉴于此, 文中设置自适应学习步长的规则为: 设 $\eta(n)$ 为第 n 次的学习步长, 若第 $n+1$ 次学习中得到的误差 $E(n+1) > E(n)$, 则忽略修正, 并使得 $\eta(n+1) = \lambda\eta(n)$ ($\lambda < 1$), 一般情况下 λ 设为 0.95; 若第 $n+1$ 次学习中得到的误差 $E(n+1) < E(n)$, 则此次修正有效, 并使得 $\eta(n+1) = \tau\eta(n)$ ($\tau > 1$), 一般情况下 τ 设为 1.05。总的调整思想为: 在学习收敛的情况下, 自适应增大 η , 以缩短学习时间; 当 η 偏大导致不能收敛时, 要马上减小, 直到收敛为止。

(3) 最佳隐含层的选取

针对隐含层节点的数量选取缺乏理论指导的问题,文中根据相关经验公式选取合适的节点个数,再由算法计算出范围内误差最小的节点个数最终参与神经网络的预测模型。

经验公式设置如下:

$$h = \sqrt{m+n} + a \quad (9)$$

式中, h 是隐含层节点数目; m 是输入层节点数目; n 是输出层节点数目; a 是 1-10 之间的常数进行调节. 在训练中利用以上公式反复试验, 根据输入和输出神经元个数确定在 5-10 个之间, 并编写相应代码在其范围内自动选择误差最小的节点数作为隐含层神经元个数。

改进后的 BP 神经网络算法流程, 如图 3 所示。

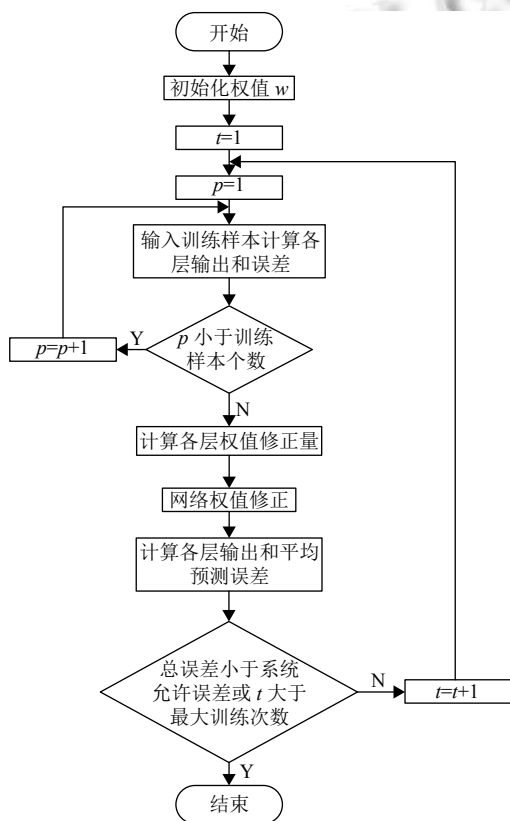


图 3 改进后的 BP 神经网络算法流程

2.4 预测模型的建立

2.4.1 训练样本的预处理

根据之前描述的基于皮尔逊积矩相关系数的相关性分析可知, 输入层选取太阳日辐射总量、最大湿度、最低湿度、平均湿度、最高温度、最低温度、平均温度、平均风速共 8 个相关程度较大的气象因素。

由于不同变量的数值差异非常大, 为了确保各因子之间处于同等地位而不会引起神经元的过饱和, 所以对原始数据进行归一化处理。

归一化是使所有输入样本在 [0, 1] 的范围内, 相关公式如下:

$$P_{t_s}^n = \frac{P_t^n - P_{\min}}{P_{\max} - P_{\min}} \quad (10)$$

$$X_{t_s}^n = \frac{X_i^n - X_{i,\min}}{X_{i,\max} - X_{i,\min}} \quad (11)$$

式中, $P_{t_s}^n$ 是第 n 个样本 t 时刻归一化后的输出功率, P_t^n 是第 n 个样本 t 时刻归一化前的输出功率, P_{\min} 是样本中最小的输出功率, P_{\max} 是样本中最大的输出功率, $X_{t_s}^n$ 是第 n 个样本中第 i 个气象参数归一化后的值, X_i^n 是第 n 个样本中第 i 个气象参数归一化前的值, $X_{i,\min}$ 是样本中第 i 个气象参数的最小值, $X_{i,\max}$ 是样本中第 i 个气象参数的最大值。

在 Matlab 中神经网络工具箱中, 归一化函数有 *premnmx*、*prestd*、*mapminmax* 3 种方法, 它们都具有将数据归一化并分布在 [0, 1] 之间的作用. 相比之下, *mapminmax* 函数可以使收敛速度更快, 结果更精确, 所以在预测模型中我们选取了 *mapminmax* 函数来对数据归一化, Matlab 语法设置如下:

$$[Y, PS] = \text{mapminmax}(X, FP) \quad (12)$$

2.4.2 网络结构的建立

由于一个三层结构的神经网络基本可以逼近任何非线性连续函数, 因此这里采用基于三层 BP 神经网络对光伏发电输出功率进行预测。

(1) 输入层作用是接受外部输入数据, 起到了缓冲存储器的作用, 所以神经元个数由输入矢量的维数决定. 研究中输入层到隐含层的传递函数使用 S 型曲线来完成, 利用的是正切函数 (*tansig*).

(2) 隐含层的作用是为了从训练样本中将映射规律提取出来并加以存储. 由于输入层和输出层都是可以先行确定, 然而隐含层节点的个数却不能预先设定. 如果隐含层的节点个数太多, 就会将训练样本中的非规律误差进行学习并且存储起来, 同时也延长了训练时间, 降低了神经网络的泛化能力和预测准确率; 但是如果隐含层节点的个数太少, 从样本中获取信息的能力就差, 不能更好的反应映射规律, 故而也降低了预测准确率. 至今隐含层的节点数量选取仍没有统一标准,

只能通过一些经验公式来大概确定数目。

(3) 输出层的值为预测当日7点到19点之间每小时的平均功率,因此输出层神经元数共13个。研究中采用了 purelin 型传递函数 (purelin transfer function),该函数特点在于输入等于输出,不用做任何附加处理。当隐含层数目较多时,网络可以采用任意精度来逼近一个有限间断点的非线性函数。训练函数使用的是带有动量项的梯度下降函数 (traingdm)。

模型的输入与输出变量结构,如图4所示。

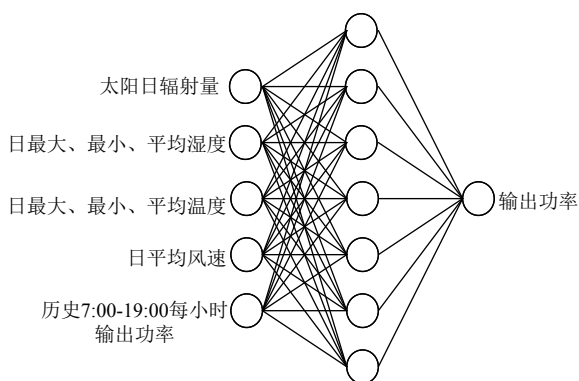


图4 网络结构

2.4.3 预测模型的评估

文中通过计算平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE) 的方法来对预测模型进行精度评估。

$$MAPE = \frac{100}{N} \sum_{i=1}^N \frac{|p_f^i - p_a^i|}{p_a^i} \times 100\% \quad (13)$$

式中, N 为数据总个数; p_f^i 为第 i 个样本中第 f 个输出神经元的预测值; p_a^i 为第 i 个样本中第 f 个输出神经元的目标值。

3 预测结果及分析

为了验证本文所提出的改进 BP 神经网络模型的可行性,选取了澳大利亚的 Desert Knowledge Australia Solar Center (DKASC, 纬度: -23.76, 经度: 133.87) 的一处 10.5 KW 光伏发电系统作为测试对象,采用 2015 年 1 月、4 月、7 月、10 月各的一天数据,分别来测试模型在春、夏、秋、冬 4 个季节的预测情况,同时并将数据分为两组,一组为训练样本,用于训练光伏发电系统的预测模型;另一组为测试样本,用于验证预测模型的有效性和准确性,并利用 Matlab 编写算法实现

BP 神经网络的学习训练和迭代过程。

3.1 夏季算例

传统 BP 神经网络模型与改进后模型对 2015 年 1 月 30 日 7:00-19:00 光伏发电系统输出功率预测结果,如图 5 与表 6 所示。

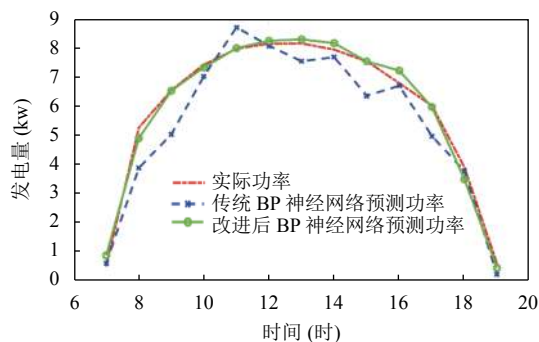


图5 1月30日光伏输出功率预测结果

表6 1月30日光伏输出功率预测误差

时间 tt:mm	实际值 (kw)	传统 BP 预 测值 (kw)	传统 BP 预测 平均误差 (%)	改进 BP 预 测值 (kw)	改进 BP 预测 平均误差 (%)
7:00	0.47	0.58	11	0.85	37
8:00	5.24	3.85	139	4.88	36
9:00	6.54	5.02	152	6.52	2
10:00	7.42	7.02	40	7.32	10
11:00	7.97	8.70	73	7.99	2
12:00	8.14	8.07	7	8.24	10
13:00	8.15	7.54	61	8.29	15
14:00	7.94	7.68	26	8.16	22
15:00	7.53	6.35	118	7.53	1
16:00	6.78	6.69	9	7.21	43
17:00	6.01	4.96	105	5.96	5
18:00	3.88	3.75	13	3.47	41
19:00	0.56	0.19	37	0.42	14

为了更直观显示模型的训练速度和精确度,分别比较了 2 种预测模型的迭代次数以及 MAPE 指标,如表 7 所示。

表7 预测模型的迭代次数和 MAPE

预测模型	迭代次数	MAPE (%)
传统 BP 模型	1162	27.82
改进后 BP 模型	359	2.41

由预测结果图 5 和误差表 6 以及表 7 来看,模型对于夏季的输出功率预测准确率较高,改进后 BP 模型预测值的 MAPE 为 2.41%,相对于传统 BP 预测值,精度得到了大幅度的提升。主要原因是夏季天气以晴朗为主,太阳辐射量大,温度较高,预测日的输出功率随

时间变化规律和训练样本十分相似, 复杂气候因素较少. 因此, 神经网络的训练误差小, 预测能力提升较大, 同时训练速度也提升了 68%.

3.2 秋季算例

传统 BP 神经网络模型与改进后模型对 2015 年 4 月 30 日 7:00-19:00 光伏发电系统输出功率预测结果, 如图 6 与表 8 所示.

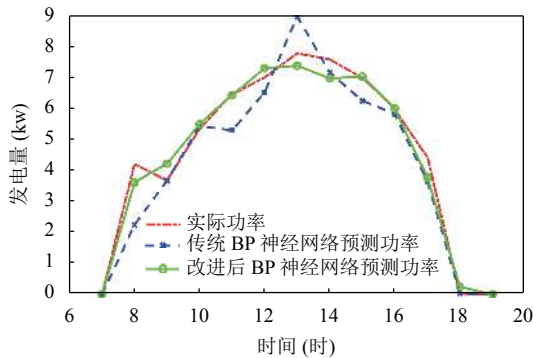


图 6 4 月 30 日光伏输出功率预测结果

表 8 4 月 30 日光伏输出功率预测误差

时间 tt:mm	实际值 (kw)	传统 BP 预 测值 (kw)	传统 BP 预测 平均误差 (%)	改进 BP 预 测值 (kw)	改进 BP 预测 平均误差 (%)
7:00	0	0	0	0	0
8:00	4.20	2.24	196	3.61	59
9:00	3.68	3.67	1	4.21	52
10:00	5.34	5.44	10	5.50	15
11:00	6.45	5.30	115	6.42	2
12:00	7.00	6.51	49	7.30	30
13:00	7.78	8.95	117	7.38	39
14:00	7.59	7.17	42	6.97	62
15:00	6.99	6.24	75	7.03	4
16:00	5.98	5.82	17	5.99	0
17:00	4.42	3.65	77	3.77	65
18:00	0.03	0.03	0	0.25	22
19:00	0	0	0	0	0

这 2 种预测模型的迭代次数以及 MAPE 指标, 如表 9 所示.

表 9 预测模型的迭代次数和 MAPE

预测模型	迭代次数	MAPE (%)
传统 BP 模型	865	97.96
改进后 BP 模型	266	21.09

由预测结果图 6、表 8 和表 9 数据来看, 秋季的光伏系统输出功率预测误差较夏季有所增大, 原因是秋季平均温度有所降低, 在 22°C 左右. 预测日天气类型为晴转多云, 湿度较小, 随机分量较大, 使得预测日的

输出功率与实际变化规律和训练样本有着一定的差距.

3.3 冬季算例

传统 BP 神经网络模型与改进后模型对 2015 年 7 月 12 日 7:00-19:00 光伏发电系统输出功率预测结果, 如图 7 与表 10 所示.

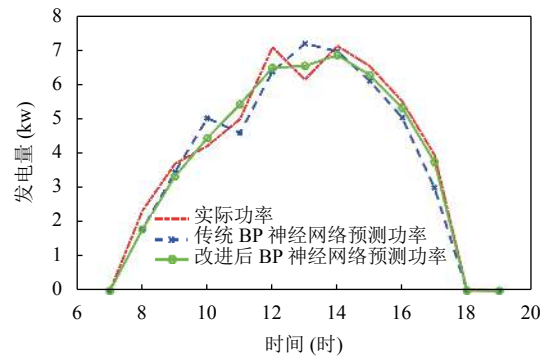


图 7 7 月 12 日光伏输出功率预测结果

表 10 7 月 12 日光伏输出功率预测误差

时间 tt:mm	实际值 (kw)	传统 BP 预 测值 (kw)	传统 BP 预测 平均误差 (%)	改进 BP 预 测值 (kw)	改进 BP 预测 平均误差 (%)
7:00	0	0	0	0	0
8:00	2.34	1.80	55	1.80	54
9:00	3.71	3.47	24	3.34	36
10:00	4.23	5.04	81	4.46	22
11:00	5.01	4.61	40	5.44	43
12:00	7.1	6.39	71	6.50	60
13:00	6.16	7.21	105	6.56	40
14:00	7.14	6.98	16	6.87	27
15:00	6.55	6.12	43	6.29	27
16:00	5.53	5.05	48	5.34	19
17:00	3.97	3.01	96	3.76	21
18:00	0.03	0.01	1	0.01	1
19:00	0	0	0	0	0

这 2 种预测模型的迭代次数以及 MAPE 指标, 如表 11 所示.

表 11 预测模型的迭代次数和 MAPE

预测模型	迭代次数	MAPE (%)
传统 BP 模型	865	55.95
改进后 BP 模型	227	19.07

由预测结果图 7、表 10 与表 11 数据来看, 冬季因为气温偏低, 太阳辐射量减少, 且多云天气占比大等原因造成预测误差相比夏季有所升高. 预测日天气类型为多云, 除了气温、太阳辐射量外, 湿度、风速也都偏低. 但总体上, 经过改进后的模型预测, 将误差从 55.95% 进一步缩小至 19.07%, 收敛速度提升了 73%,

可见改进之后的预测结果仍具有较高参考意义。

3.4 春季算例

传统 BP 神经网络模型与改进后模型对 2015 年 10 月 12 日 7:00–19:00 光伏发电系统输出功率预测结果,如图 8 和表 12 所示。

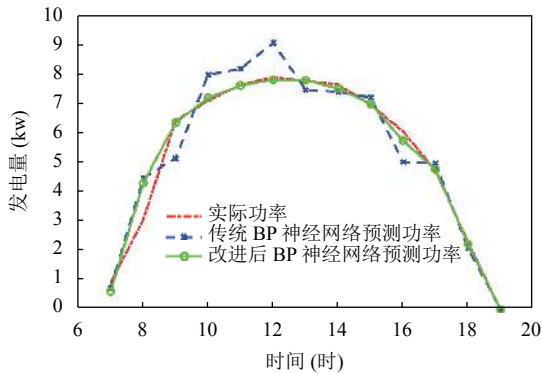


图 8 10 月 12 日光伏输出功率预测结果

表 12 10 月 12 日光伏输出功率预测误差

时间 tt:mm	实际值 (kw)	传统 BP 预 测值 (kw)	传统 BP 预测 平均误差 (%)	改进 BP 预 测值 (kw)	改进 BP 预测 平均误差 (%)
7:00	0.91	0.73	17	0.62	28
8:00	3.06	4.48	141	4.31	12
9:00	6.44	5.16	128	6.38	6
10:00	7.12	8.00	87	7.21	9
11:00	7.66	8.21	56	7.64	1
12:00	7.93	9.09	115	7.83	11
13:00	7.79	7.49	30	7.82	3
14:00	7.67	7.42	25	7.54	13
15:00	6.98	7.23	25	7.02	4
16:00	6.07	5.03	104	5.75	32
17:00	4.73	4.98	25	4.78	5
18:00	2.29	2.08	21	2.23	6
19:00	0	0	0	0	0

这 2 种预测模型的迭代次数以及 MAPE 指标,如表 13 所示。

表 13 预测模型的迭代次数和 MAPE

预测模型	迭代次数	MAPE(%)
传统 BP 模型	1148	44.34
改进后 BP 模型	652	10.55

由预测结果图 8、表 12 和表 13 数据来看,春季气温回暖,预测日天气类型为阵雨。根据当地气象数据报告显示当天 14:00 和 16:00 雨量偏大,对光伏发电输出功率预测的准确性有一定的影响。另外误差较大的时刻还出现在早晨,无论传统 BP 模型还是改进后模型的

相对误差都大于其他时刻,但此时输出功率较小,不会因此造成绝对误差增大从而影响电力系统。

综上所述,总体来看 BP 神经网络预测模型在夏季时的预测准确度最高,而在秋季时的预测准确度最低,经以上分析可见,气象环境的随机因素对预测模型影响仍然较大;同时通过 4 个季节的预测结果,验证了改进后的 BP 模型比传统 BP 模型在预测精度和收敛速度上均有大幅提升,充分证明了此方法在在工程应用上具有一定的可行性和参考价值。

4 结论

神经网络能够很好反映出输入与输出之间的非线性映射关系,相对于人工建模的预测方法,其方法更简便、准确。本文建立了一种基于改进 BP 神经网络的光伏发电预测模型,在模型构建时充分考虑了多种气象因素的影响,经过相关性分析,说明了这些影响因素对于光伏发电的紧密关系。针对传统 BP 学习训练中容易陷入局部极小值而得不到最优值、收敛速度慢等问题,文中使用了增加动量项、自动选取最佳隐含层的方法对其进行改进。最后,通过澳大利亚一处光伏电站的运行数据对预测模型进行比较验证,针对春、夏、秋、冬 4 个季节的预测数据进行分析,其结果在预测精度和收敛速度方面均有明显提升。

参考文献

- 沈瑶, 罗安, 陈燕东, 等. 基于功率前馈的光伏并网控制方法及稳定性分析. 电网技术, 2014, 38(9): 2449–2454.
- 索江镭, 胡志坚, 刘宇凯, 等. 大规模光伏发电并网对互联电力系统阻尼特性的影响及其阻尼控制策略. 西安交通大学学报, 2015, 49(2): 99–105.
- 简献忠, 魏凯, 郭强. 蜂群算法在光伏电池双二极管五参数模型中的应用. 光子学报, 2015, 44(1): 168–172.
- Rikos E, Tselepis S, Hoyer-Klick C, *et al.* Stability and power quality issues in microgrids under weather disturbances. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2008, 1(3): 170–179. [doi: 10.1109/JSTARS.2008.2010557]
- 李安寿, 陈琦, 王子才, 等. 光伏发电系统功率预测方法综述. 电气传动, 2016, 46(6): 93–96. [doi: 10.3969/j.issn.1001-2095.2016.06.022]
- 李军, 许继生, 王生渊, 等. 一种并网型光伏电站光功率及发电量预测的方法. 青海电力, 2010, 29(2): 18–20. [doi: 10.3969/j.issn.1006-8198.2010.02.007]

- 7 丁明, 徐宁舟. 基于马尔可夫链的光伏发电系统输出功率短期预测方法. 电网技术, 2011, 35(1): 152–157.
- 8 李乐, 刘天琪. 基于近邻传播聚类 and 回声状态网络的光伏预测. 电力自动化设备, 2016, 36(7): 41–46.
- 9 高阳, 张碧玲, 毛京丽, 等. 基于机器学习的自适应光伏超短期出力预测模型. 电网技术, 2015, 39(2): 307–311.
- 10 Shi J, Lee WJ, Liu YQ, *et al.* Forecasting power output of photovoltaic systems based on weather classification and support vector machines. IEEE Transactions on Industry Applications, 2012, 48(3): 1064–1069. [doi: [10.1109/TIA.2012.2190816](https://doi.org/10.1109/TIA.2012.2190816)]
- 11 Yona A, Senjyu T, Saber AY, *et al.* Application of neural network to 24-hour-ahead generating power forecasting for PV system. Proceedings of 2008 IEEE Power and Energy Society General Meeting—Conversion and Delivery of Electrical Energy in the 21st Century. Pittsburgh, PA, USA, 2008. 1–6.
- 12 陈昌松, 段善旭, 殷进军. 基于神经网络的光伏阵列发电预测模型的设计. 电工技术学报, 2009, 24(9): 153–158. [doi: [10.3321/j.issn:1000-6753.2009.09.023](https://doi.org/10.3321/j.issn:1000-6753.2009.09.023)]
- 13 严丽坤. 相关系数与偏相关系数在相关分析中的应用. 云南财贸学院学报, 2003, 19(3): 78–80. [doi: [10.3969/j.issn.1674-4543.2003.03.018](https://doi.org/10.3969/j.issn.1674-4543.2003.03.018)]
- 14 李冬辉, 王鹤雄, 朱晓丹, 等. 光伏并网发电系统几个关键问题的研究. 电力系统保护与控制, 2010, 38(21): 208–214. [doi: [10.7667/j.issn.1674-3415.2010.21.041](https://doi.org/10.7667/j.issn.1674-3415.2010.21.041)]
- 15 徐维. 并网光伏发电系统数学模型研究与分析. 电力系统保护与控制, 2010, 38(10): 17–21. [doi: [10.3969/j.issn.1674-3415.2010.10.004](https://doi.org/10.3969/j.issn.1674-3415.2010.10.004)]
- 16 Yona A, Senjyu T, Funabashi T. Application of recurrent neural network to short-term-ahead generating power forecasting for photovoltaic system. Proceedings of 2007 IEEE Power Engineering Society General Meeting. Tampa, FL, USA, 2007. 1–6.
- 17 高隽. 神经网络原理及仿真实例. 北京: 机械工业出版社, 2003.
- 18 Kalogirou SA. Artificial neural networks in renewable energy systems applications: A review. Renewable and Sustainable Energy Reviews, 2001, 5(4): 373–401. [doi: [10.1016/S1364-0321\(01\)00006-5](https://doi.org/10.1016/S1364-0321(01)00006-5)]