

基于 HPO-BP 光伏发电功率预测研究

高 昕¹ 陈 龙¹
GAO Xin CHEN Long

摘 要

光伏发电作为一种清洁、可再生的能源形式，在全球范围内得到了广泛的关注和应用。然而，由于气象条件的复杂性和不确定性，光伏发电的功率波动较大，给电网系统的稳定运行带来了一定的挑战。因此，提高光伏发电预测精度对于保障电网系统的安全和稳定具有重要意义。为了实现这一目标，提出了一种基于 HPO-BP 神经网络的预测模型。首先，利用灰色关联度和高斯混合模型相结合的方法，筛选出与光伏发电功率相关性较高的气象因素，从而构建出更加准确的预测模型。其次，针对 BP 神经网络的参数较多、易陷入局部最优等问题，采用了 HPO 技术对 BP 神经网络的参数进行优化，提高了预测精度和收敛速度。最后，通过实验验证了模型的优越性。经过测试，HPO-BP 神经网络预报模式显著优于传统光伏发电预报模式，其准确性、泛化性以及其它特性都远超过传统模型，从而使得供电系统的安全、可靠度获得了更大的保障。

关键词

光伏发电；猎人猎物算法；灰色关联度；预测精度；收敛速度

doi: 10.3969/j.issn.1672-9528.2024.06.038

0 引言

伴随着全球气候变化问题日益严重，我国于 2020 年提出 2030 年“碳达峰”与 2060 年“碳中和”目标。随着科技的进步，太阳能已经成为可持续的可再生能源，而且它在光伏发电方面也有着巨大的潜力。然而，由于其容易受到外界条件的变化，若不进行有效的调整，就会给电力市场带来不可估量的损失。所以，正确地预估光伏发电的效益，以及减少不必要的损失，对确保电力市场的安全性至关重要。为了提高和改善光伏发电的精度，国内外很多学者做了相关的工作。一些国内学者提出了粒子群算法来改善 BP 神经网络的预测性能，但由于其本质特性，仍有很大的改进空间。乔路丽等人的免疫遗传算法在预测精度上具有优势，但其算法自身复杂度较高，对资源具有一定要求。总体而言，这些方法在预测效果方面各有优缺点。

本文采取灰色关联度和高斯混合模型相结合的方法，以此来构建相似日模型。选取猎人猎物算法（HPO）和 GA 算法来对比分析 BP 神经网络预测模型结果。通过实验验证，本文的 HPO-BP 预测模型在光伏发电预测精度上准确性更高。

1 相似日的选取

1.1 气象因素相关性分析

本文选取某市光伏电站 2021 年 1 月 1 日—2021 年 12

月 31 日全年数据，其中每间隔 15 分钟对数据进行一次采集。共计 35 040 条数据。其中包含总辐射强度、环境温度、组件温度、气压、相对湿度等因素和实际的发电功率数据。具体列举该市某天采集的数据如图 1 所示。

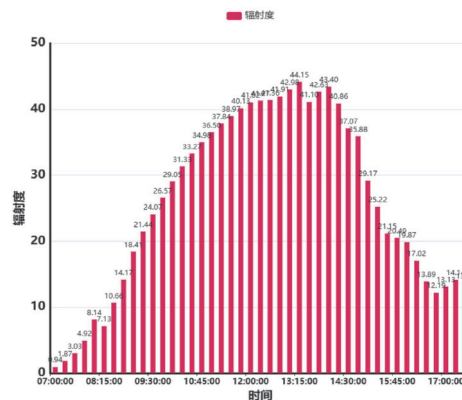


图 1 日辐射影响因素变化趋势图

考虑到一年中存在昼夜温差不同的情况，本文通过采取一天中 7:30—17:30 来进行实验，以此来让实验过程和结果更加有说服力。通过应用灰色关联度的方法来选取和实际光伏发电功率关联程度最密切的参数，具体公式为：

$$\varepsilon_i(k) = \frac{\min_k \min_i |Y(k) - X_i(k)| + \rho \max_k \max_i |Y(k) - X_i(k)|}{|Y(k) - X_i(k)| + \rho \max_k \max_i |Y(k) - X_i(k)|} \quad (1)$$

式中：分辨系数 ρ 是一个介于 0 和 1 之间的数值，它用于衡量两个变量之间的相似程度。在某些情况下，分辨系数越高，

表明两个变量越相似；反之，分辨系数越低，表示两个变量相似程度越低。在本文中，分辨系数设定为 0.5，这意味着在寻找相似日时，会关注那些与目标日相似程度较高的日子。关联系数反映的是每个时刻的关联程度，为了对数据进行整体比较，需将数据进行整合，公式如下：

$$r_i = \frac{1}{T} \sum_{k=1}^T \varepsilon_i(k) \quad (2)$$

式中： T 表示每个时刻。

在灰色关联度分析之前，通常需要对数据进行均值化的无量纲化处理，以消除数据量纲和数值差异对分析结果的影响。这有助于找出关键因素，发现数据之间的关联规律，从而为后续的预测和决策提供依据。具体参数与实际发电功率的关联度如表 1 所示。

表 1 灰色关联度

参数	辐射度	环境温度	组件温度	气压	相对湿度
关联度	0.973	0.723	0.692	0.626	0.614
排名	1	2	3	4	5

由表 1 可知，辐射度的关联度为 0.973，具有强相关性。紧随其后的是环境温度的关联度，接近 0.7。最后是气压和湿度湿度的相关性，分别是 0.626 和 0.614。

1.2 高斯混合模型选取相似日

在高斯混合模型中，通过将多个高斯分布组合在一起描述一组数据。相似日模型则是从这些高斯分布中选取与给定日期最接近的日期。这里选取相似日的原理主要基于以下几点。

(1) 距离度量。首先，需要计算给定日期与每个高斯分布代表的日期的距离。通常使用欧氏距离或余弦相似度等度量方法。距离越小，表示日期越相似。

(2) 概率融合。计算每个高斯分布对应的概率，并将它们融合在一起。融合方法可以是简单的加权平均，也可以是更复杂的融合策略，如投票或决策树等。概率越高，表示该日期与给定日期的相似度越高。

(3) 阈值设定。为了筛选出最相似的日期，需要设定一个阈值。只有当某个日期的概率超过这个阈值时，它才被认为是相似的。

(4) 迭代优化。上述过程可能需要多次迭代，以不断优化相似日的结果。可以通过调整高斯分布的参数（如均值、协方差矩阵等）来实现优化。

如果 GMM 的聚类数量为 K ，那么可以使用 0 作为参数，包括均值、协方差和权重。

EM 算法由期望步（E-step）和最大步组成，其中 E-step 用于计算样本点 z_i 属于第 k 个分布的可能性，具体公式为：

$$\gamma_k(z_i) = \frac{\omega_k N(z_i | \mu_k, \Sigma_k)}{\sum_{k=1}^K \omega_k N(z_i | \mu_k, \Sigma_k)} \quad (3)$$

式中： k 、 N 、 z 是高斯概率密度函数。第 k 个数据的平均数、平均方差以及平均权重都是由第 k 个数据计算出来的。通过 M-step，可以对每个数据的参数进行计算，然后进行更改，具体的形式如下：

$$\mu_k = \frac{\sum_{i=1}^N \gamma_k(z_i) x_i}{\sum_{i=1}^N \gamma_k(z_i)} \quad (4)$$

$$\Sigma_k = \frac{\sum_{i=1}^N \gamma_k(z_i) (z_i - \mu_k)(z_i - \mu_k)^T}{\sum_{i=1}^N \gamma_k(z_i)} \quad (5)$$

$$\omega_k = \frac{1}{N} \sum_{i=1}^N \gamma_k(z_i) \quad (6)$$

多次重复上述步骤，获得高斯混合模型，最终将样本点进行聚集。

在本文中，通过使用均差、标准值、具有高度相关的气候变化因子，以及过去的发电量，构成了一个具有高度模型化的日变化特征向量，其中每个变化都包含 $X_j = [X_{j,1}, X_{j,2}, \dots, X_{j,10}]$ ($j=1,2,\dots,N$)，并利用贝叶斯信息准则（BIC）的方式确定了高斯混合模型的最优聚类参数，从而获得了具有高度模型化的光伏相似的日变化样本。使用贝叶斯信息准则（BIC）的方法来确定高斯混合模型的最佳聚类数，由此建立光伏相似日样本数据。

2 预测模型构建

2.1 猎人猎物优化算法（HPO）

HPO 算法的主要思想是基于猎人追踪猎物的自然现象。在算法中，猎物代表问题空间中的解，它们会不断寻找全局最优解，以保证自身的安全性。而猎人则代表着搜索过程，通过不断迭代和更新策略，致力于找到最优解。HPO 算法的具体寻优过程主要包括以下四个环节。

(1) 种群位置初始化：HPO 的初始化个体位置采取随机的方式分布，表达式为：

$$x_i = \text{rand}(1, d) \cdot (u_b - l_b) + l_b \quad (7)$$

式中： x_i 为猎人和猎物位置； d 为搜索空间维度； u_b 为解向量空间的上限。

(2) 猎人位置更新：猎人通过猎物位置和平均距离，猎人可以调整自己的位置。猎物位置更新通常包括线性移动和随机扰动两部分，以实现搜索过程中既有方向性又保持一定的随机性。猎人的更新公式为：

$$x_{ij}(t+1) = x_{ij}(t) + 0.5(2CZP_{\text{pos}(j)} - x_{ij}(t)) + 0.5(2(1-C)Z\mu_{(j)} - x_{ij}(t)) \quad (8)$$

式中： $x_{ij}(t)$ 是指当前的猎手的位置； $x_{ij}(t+1)$ 是指他们下一秒的位置； P_{pos} 是指他们捕获的目标； μ 是指每种目标的平均距离； C 是用来调整平衡算法的参数，它的取值会根据迭代的次数由 1 逐渐降低至 0.02； Z 是用来调整的自适应参数。 C

和 Z 的值公式为:

$$C = 1 - I(\frac{0.98}{\max I_i}) \quad (9)$$

$$IDX = (P == 0) \quad (10)$$

$$Z = R_2 \otimes IDX + R_3 \otimes (\sim IDX) \quad (11)$$

式中: I_i 表示当前的迭代次数, $\max I_i$ 是指最大迭代次数, R_1 和 R_3 是两者之间的随机变量 $[0,1]$, R_2 是它们之间的值, IDX 表示当 R_3 满足条件 (即 $P=0$) 时的个数。

(3) 猎物位置更新: 当猎物被猎人发现并开始被捕捉时, 猎物会向安全位置进行移动, 此时安全位置被称为全局最优位置, 其位置更新的公式如下:

$$x_{ij}(t+1) = T'_{pos(j)} + CZ\cos(2\pi R_4) \times [T'_{pos(j)} - x_{ij}(t)] \quad (12)$$

式中: x_i 、 j 、 $t+1$ 分别表示当前和下一时刻的猎物位置, $T'_{pos(j)}$ 是当前获得的最佳位置, 而 R_4 则是两者之间的随机值 $[-1,1]$ 。

2.2 BP 神经网络

BP 神经网络是一种复杂的多层神经网络, 它通过模拟人脑神经元来实现信息的前向传播。它由输入、隐藏和输出三个部分组成, 其中输入部分负责接收外界信号, 而隐藏部分则负责处理这些信号, 最终将其转化为输出。当输出层得不到预期的结果时, BP 神经网络会进行反向传播。通过对数据误差的分析, BP 神经网络的阈值被精确地控制, 从而实现了一种高效的非线性映射。它的表现优于传统的函数拟合, 并且表现出良好的拟合效果, 见图 2。

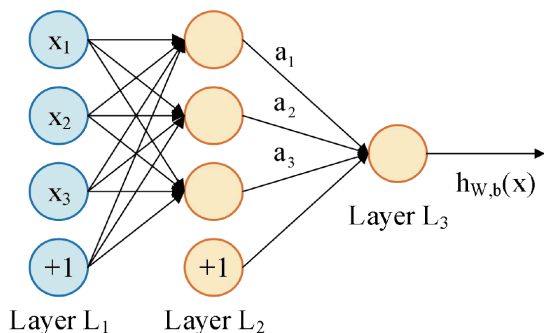


图 2 BP 神经网络结构图

根据本文中提出的优化步骤, 构建了一个 HPO-BP 神经网络算法, 其流程如图 3 所示。

3 模型验证

为了验证猎人猎物算法优化 BP 神经网络模型在发电功率预测上的准确性, 分别构建 GA-BP 神经网络预测模型和 HPO-BP 神经网络预测模型, 对比实验结果分析。

通过图 4 中神经网络参数的梯度、学习率以及迭代次数可以发现, 对比经过 HPO 优化的 BP 神经网络在性能上显著优于原始的 GA 优化的 BP 神经网络。

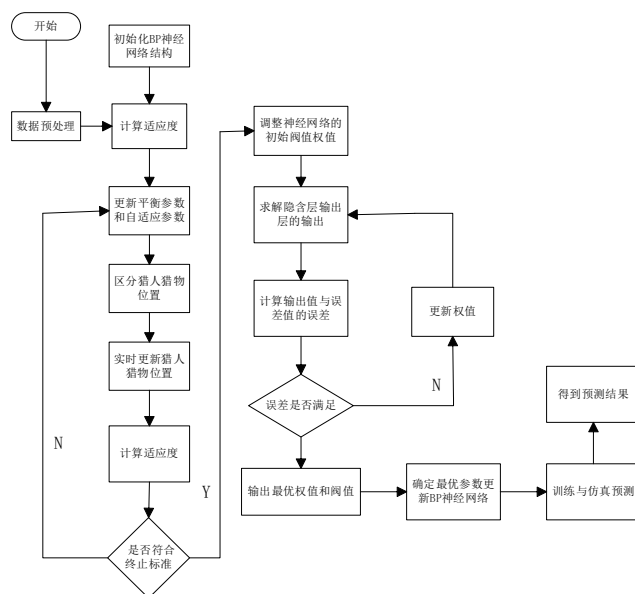


图 3 HPO-BP 神经网络结构

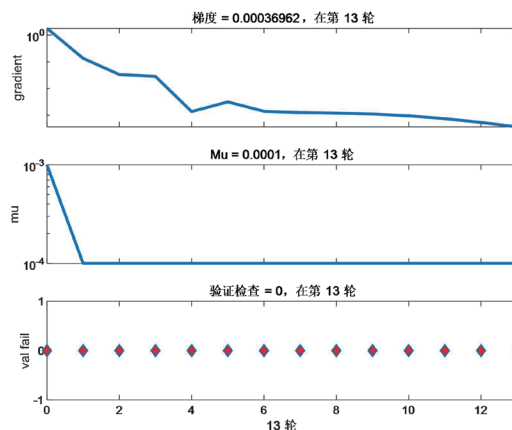


图 4 神经网络参数

选择均方误差 (mean squared error, MSE)、均方根误差 (mean squared error, RMS)、平均数绝对差错 (mean absolute error, MAE) 和平均数一定值比例错误 (mean absolute percentage error, MAPE) 用作估计成果的指标。将预测模型的预测结果和真实值进行对比, 分析结果如图 5~7 所示。

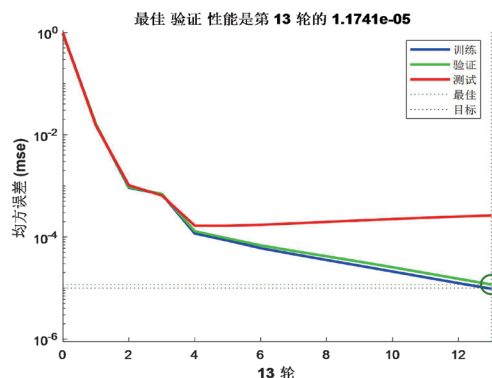


图 5 均方误差

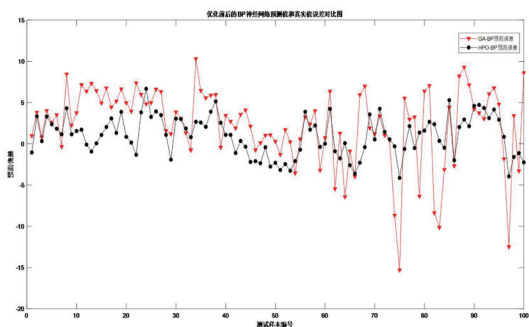


图6 BP神经网络预测值与真实值偏差

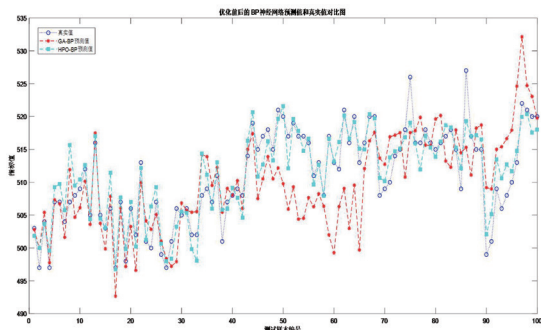


图7 预测值与实测值对比

对比上述图不难发现，HPO优化的BP神经网络表现出更低的误差。这表明HPO优化能够帮助神经网络更好地逼近真实函数，提高预测精度。其次，从梯度、适应度和迭代次数上来看，HPO优化的BP神经网络在更少的迭代次数内达到了较低的误差。这说明HPO优化能够有效地加速神经网络的收敛速度、寻优效果，同时还减少了一定的训练时间。

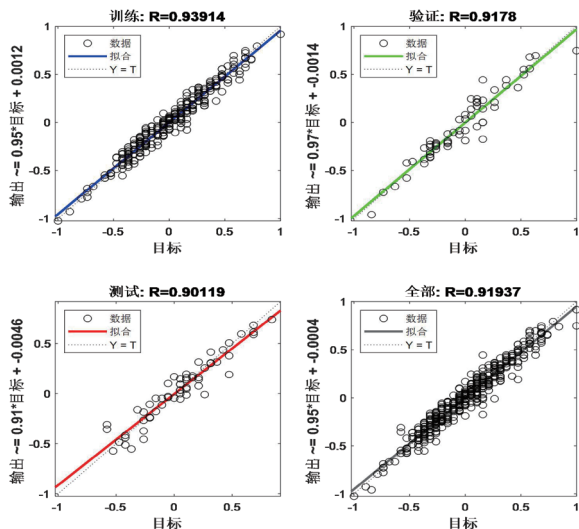


图8 神经网络预测样本拟合情况

此外，训练和测试的 R 值也是为了直观比较不同模型之间的区别，实验结果如图8所示。其中均方根误差（RMSE）也可以在数量级上直观反映预测偏差。

综上所述，通过HPO优化BP神经网络，能够显著提高模型的性能和效率。HPO技术的引入使得BP神经网络在训

练过程中能够自动调整超参数，从而达到更好的学习效果。

HPO优化对BP神经网络的性能提升效果较为明显，不仅提高了预测精度，还加速了网络的收敛速度，并增强了模型的泛化能力。这种优化方法不仅减少了人工干预的需要，而且提高了模型的稳定性和可靠性。

4 结论

本文提出了一种新的方法来评估光伏发电的效益。使用HPO-BP技术来处理这些信息，并利用这些信息来评估它们的准确性。本文研究表明，这种方法能够更准确地预测光伏发电的效益，从而保障电力系统的稳定性。

参考文献：

- [1] 王捷, 林余杰, 吴成坚. 碳中和背景下太阳能光伏产业现状及发展[J]. 储能科学与技术, 2022, 11(2): 731-732.
- [2] 张慧斌, 刘大贵, 朱婷婷, 等. 基于相似日和动量法优化BP神经网络的光伏短期功率预测研究[J]. 智慧电力, 2021, 49(6): 46-52.
- [3] 赖昌伟, 黎静华, 陈博, 等. 光伏发电出力预测技术研究综述[J]. 电工技术学报, 2019, 34(6): 1201-1217.
- [4] 吴鹏. 新能源并网对电力系统电能质量的影响[J]. 光源与照明, 2022(3): 246-248.
- [5] 朱二周, 孙悦, 张远翔, 等. 一种采用新型聚类方法的最佳类簇数确定算法[J]. 软件学报, 2021, 32(10): 3085-3103.
- [6] 王莹. 基于深度学习的光伏发电功率短期预测研究[D]. 青岛: 青岛科技大学, 2022.
- [7] 成珂, 孙琦琦, 马晓瑶. 基于主成分回归分析的气象因子对光伏发电量的影响[J]. 太阳能学报, 2021, 42(2): 403-409.
- [8] 耿博, 高贞彦, 白恒远, 等. 结合相似日GA-BP神经网络的光伏发电预测[J]. 电力系统及其自动化学报, 2017, 29(6): 118-123.
- [9] 王芬, 杨媛. 基于猎犬优化算法求解TSP问题[J]. 宁夏师范学院学报, 2022, 43(7): 59-63.
- [10] 孙德刚. BP神经网络遗传算法的图像识别技术分析与实现[J]. 信息技术与信息化, 2022(5): 190-193.
- [11] 辛宇, 童孟军, 华宇婷. 一种基于最优特征选择改进的遗传算法[J]. 传感技术学报, 2018, 31(11): 1747-1752.
- [12] MISHRA M, DASH P B, NAYAK J, et al. Deep learning and wavelet transform integrated approach for short-term solar PV power prediction[J]. Measurement, 2020, 166: 108250-1-108250-16.

【作者简介】

高昕（1965—），女，安徽淮南人，副教授，研究方向：电力电子及其自动化、电机与电气等。

陈龙（2000—），男，安徽合肥人，硕士研究生，研究方向：电力系统及其自动化。

（收稿日期：2024-04-29）