

采用特征优选和优化深层核极限学习机的 短期风电功率预测

商立群, 黄辰浩, 侯亚东, 李洪波, 惠泽, 张建涛
(西安科技大学电气与控制工程学院, 710054, 西安)

摘要: 针对风电出力非线性、不稳定且用传统方法难以准确预测的问题,提出了一种基于对深层混合核极限学习机(DHKELM)参数进行优化的短期风电功率预测。利用核主成分分析(KPCA)方法进行特征优选得到的最优特征集,既能表达风电功率的有效信息,也能避免冗余信息的出现,有利于 DHKELM 模型的学习与训练,同时也降低了模型的复杂度。针对 DHKELM 超参数难确定的问题,利用改进的野犬优化算法(IDOA)对 DHKELM 的 8 个超参数进行寻优,可以发掘原始序列特征信息,从而使模型能够充分掌握数值天气预报(NWP)与风电功率之间的非线性关系。以国外某风电场真实数据为算例,结果表明:提出的预测模型相较于野犬算法、差分进化算法和粒子群优化算法的平均绝对百分比误差(MAPE)分别降低了 0.979 3%、2.342 1%、3.383 2%,有效提高了风电功率的预测精度。

关键词: 短期风电功率预测;深层混合核极限学习机;改进的野犬优化算法;特征优选;核主成分分析

中图分类号: TM715 **文献标志码:** A

DOI: 10.7652/xjtuxb202301007 **文章编号:** 0253-987X(2023)01-0066-12

Short-Term Wind Power Prediction by Using the Deep Kernel Extreme Learning Machine with Well-Selected and Optimized Features

SHANG Liqun, HUANG Chenhao, HOU Yadong, LI Hongbo, HUI Ze, ZHANG Jiantao
(School of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China)

Abstract: Aiming at the problem that wind power output is nonlinear, unstable and difficult to be accurately predicted by traditional methods, this paper proposes a short-term wind power prediction based on the optimization of parameters of the deep hybrid kernel extreme learning machine (DHKELM). The kernel principal component analysis (KPCA) method is used to well select the features to form an optimal feature set, which can not only express the effective information of wind power, but also avoid the appearance of redundant information, and is thus conducive to facilitating the learning and training of the DHKELM model and reducing the complexity of the model. In view of the problem that it is difficult to determine the hyperparameters of DHKELM, the improved dingo optimization algorithm (IDOA) is used to find the eight optimal hyperparameters of DHKELM and explore the original sequence feature information, so that the model can

收稿日期: 2022-07-03。 作者简介: 商立群(1968—),男,教授,硕士生导师。 基金项目: 陕西省自然科学基金基础研究计划资助项目(2021JM-393)。

网络出版时间: 2022-10-06

网络出版地址: <https://kns.cnki.net/kcms/detail/61.1069.T.20220930.1157.003.html>

fully grasp the nonlinear relationship between numerical weather prediction (NWP) and wind power. Taking the real data of a foreign wind farm as an example, the results show that the proposed prediction model effectively improves the accuracy of wind power prediction, with the mean absolute percentage error (MAPE) 0.979 3%, 2.332 1% and 3.383 2% lower than that of the dingo optimization algorithm, the differential evolution optimization algorithm and the particle swarm optimization algorithm respectively.

Keywords: short-term wind power prediction; deep hybrid kernel extreme learning machine; improved dingo optimization algorithm; feature optimization; kernel principal component analysis

风电功率预测在应对全球气候变化问题上和在电网规划中都有着不可或缺的作用。风能强烈的间歇性和波动性影响电网的稳定运行^[1]、造成电能质量下降,因此风电功率的精准预测对于电力系统经济运行至关重要^[2]。

早期的风电功率预测方法大多是基于数值天气预报(numerical weather prediction, NWP)的物理预测。数值天气预报数据的准确性和可靠性将直接决定模型最终的预测精度,其中风速数据作为直接影响风电功率的关键因素,与风电输出功率有较强的正相关性。利用聚类分析和主成分分析对数值天气预报样本进行特征提取,能够准确反映风电场的物理特征^[3-4]。主成分分析虽然能实现数据降维,提取高维特征,但不适于处理非线性、波动性强的序列,且不能有效去除原始信号中的噪声,会影响预测结果的准确性^[5-6]。核主成分分析(kernel principal component analysis, KPCA)方法能够很好地处理非线性数据。文献[7]对制约光伏输出功率的5种气象因素进行特征提取,可有效提高功率预测精度。物理预测模型建模及求解过程复杂,需要依靠大量的NWP数据,并要考虑风电场的地理因素,故其在风电功率中短期预测上表现不佳。

目前,已有诸多学者在风电功率预测中将物理法与统计法结合使用,以提高风电功率的预测精度。随着人工智能的广泛应用,统计法已不限于灰色预测^[8]、ARIMA^[9]等时间序列法,还包括极限学习机^[10]、支持向量机^[11]、深度学习^[12]等。由于单一模型无法在处理大量数据的情况下建立变量之间良好的非线性关系,预测精度有待提高。文献[13]将详尽可能性(ELM)模型应用于时序预测,经过算例分析验证了其有效性,但ELM初始的输入权重和隐含层的偏置是随机的,无法充分体现模型性能。在ELM的基础上引入核函数可以很好地解决上述问题,将核极限学习机应用于功率预测^[14],有效解决

了随机映射对建模的影响,大大提高了功率预测的精度。深层混合核极限学习机(deep hybrid kernel extreme learning machine, DHKELM)是一种新型智能预测模型,它在极限学习机的基础上拥有更深层次的结构^[15],引入了自动编码器以及混合核函数思想,能有效改善因初始权值和阈值随机生成而导致输出稳定性低下的问题,但该模型的超参数颇多,预测性能受参数选择的影响,利用群智能算法优化其参数可有效提高模型性能。常见的群智能优化算法有粒子群优化(PSO)算法^[16]、乌燕鸥优化算法(STOA)^[17]、差分进化(DE)算法^[18]等。野犬优化算法(DOA)在考虑个体生存概率的情况下又包含4种捕食策略,能够有效求解各种未知搜索空间中的复杂问题^[19]。它在全局搜索能力和寻优精度等方面具有显著优势,但其易陷入局部最优的缺点仍然存在,利用差分进化算法、无限折叠混沌映射(ICMIC)和精英反向学习机制改进野犬算法,本文拟采用改进的野犬优化算法(IDOA)对深层混合核极限学习机的参数进行寻优。

基于以上问题,本文提出了基于IDOA优化DHKELM的风电功率短期预测方法。风电场高维NWP数据经过KPCA降维处理,能有效提取数据特征。针对DHKELM超参数难确定的问题利用IDOA对DHKELM的8个超参数进行寻优,可以发掘原始序列特征信息。以国外某风电场真实数据为算例,算例结果表明KPCA-IDOA-DHKELM预测模型能够映射出NWP与风电功率之间的非线性关系,有效提高了风电功率的预测精度。

1 深层混合核极限学习机

1.1 混合核极限学习机

传统ELM只需要设置网络中隐含层节点数量就可以生成局部最优解^[20],而不需要调整输入层的权值和隐含层的偏置,也不需要像传统神经网络那

样,通过设定大量参数来训练网络,而且 ELM 具有较快的收敛速度,学习效率更高。

对于风电数训练集 $T_0 = \{(\mathbf{x}_i, \mathbf{y}_i), i = 1, 2, \dots, N\}$, $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{im}] \in \mathbf{R}^n$ 是输入特征向量, $\mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}] \in \mathbf{R}^m$ 是相应的目标向量, $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{im}]$ 是通过 ELM 网络得到的输出向量。如果输出 $\mathbf{y}$ 与目标 $\mathbf{t}$ 之间的误差可以近似为 0, 即满足 $\sum_{i=1}^N \|\mathbf{t}_i - \mathbf{y}_i\| = 0$, ELM 模型表示为

$$\mathbf{y}_j = \sum_{i=1}^l \beta_i \mathbf{g}_i(x_j) = \sum_{i=1}^l \beta_i \mathbf{g}(a_i x_j + c_i) = \begin{bmatrix} \mathbf{g}(a_1 x_1 + c_1) & \cdots & \mathbf{g}(a_l x_1 + c_l) \\ \vdots & & \vdots \\ \mathbf{g}(a_1 x_N + c_1) & \cdots & \mathbf{g}(a_l x_N + c_l) \end{bmatrix}_{N \times l} \cdot \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_l^T \end{bmatrix}_{l \times m} = \underbrace{\begin{bmatrix} \mathbf{g}(a_1 x_1 + c_1) & \cdots & \mathbf{g}(a_l x_1 + c_l) \\ \vdots & & \vdots \\ \mathbf{g}(a_1 x_N + c_1) & \cdots & \mathbf{g}(a_l x_N + c_l) \end{bmatrix}}_{\mathbf{H}_{N \times l} = [\mathbf{h}(x_1), \dots, \mathbf{h}(x_N)]^T} \cdot \underbrace{\begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_l^T \end{bmatrix}}_{\beta_{l \times m}} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N \times m}^T, j = 1, \dots, N \quad (1)$$

式中: $\mathbf{a}_i$ 为输入权值向量; β_i 为输出权值向量; $\mathbf{b}_i$ 为偏置向量; $\mathbf{g}(\cdot)$ 为隐含层的激活函数; $\mathbf{T}$ 为目标输出矩阵; $\mathbf{H}$ 为隐含层的输出矩阵; β 为连接输出层和隐含层的输出权值矩阵。

通过求最小二乘解来表征 ELM 的学习过程, 其解如下式

$$\hat{\beta} = \mathbf{H}^+ \mathbf{T} \quad (2)$$

式中 $\mathbf{H}^+$ 为隐含层输出矩阵 $\mathbf{H}$ 的摩尔-彭罗斯广义逆矩阵。

核方法拥有强大的非线性映射能力, 为了使 ELM 获得较强的学习能力和泛化能力, 将不同类型的核函数融合。引入核参数 $\mathbf{H}\mathbf{H}^T$, 以提高 ELM 模型的泛化能力, KELM 的输出函数可以表示为^[21]

$$\mathbf{y}(x) = \mathbf{h}(x)\beta = \mathbf{h}(x)\mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{T} \quad (3)$$

式中: 正值常数 C 是惩罚参数; $\mathbf{I}$ 是单位矩阵。

KELM 的核函数定义如下

$$f_{\text{KELM}} = \mathbf{H}\mathbf{H}^T; f_{\text{KELM}, i, j} = \mathbf{h}(\mathbf{x}_i)\mathbf{h}(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j) \quad (4)$$

KELM 的模型函数可以写为

$$\mathbf{y}(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix} \left(\frac{\mathbf{I}}{C} + f_{\text{KELM}} \right)^{-1} \mathbf{T} \quad (5)$$

传统的核函数包括全局核函数和局部核函数。全局核函数强调数据作为一个整体对核函数的影

响, 泛化能力较强, 但学习能力较弱。局部核函数强调了关键点附近的数据对核函数的影响, 具有较强的学习能力, 泛化能力较弱。

高斯核函数是一个典型的局部核函数, 高斯核函数的学习能力优于多项式核函数, 高斯核函数的表达式如下式

$$K_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (6)$$

式中 σ 是用于控制径向范围的参数。

多项式核函数作为典型的全局核函数, 因其较强的泛化能力被广泛应用于 ELM 中^[22]。多项式核函数的表达式如下

$$K_{\text{poly}}(\mathbf{x}_i, \mathbf{x}_j) = ((\mathbf{x}_i \mathbf{x}_j) + q)^p \quad (7)$$

式中: q 和 p 分别为多项式核函数的常数参数和指数参数。

为了提高 KELM 的学习能力和泛化能力, 提出了两个核函数的组合作为混合函数。本文采用了线性权重法, 并将新的混合函数的方程表示如下

$$K(\mathbf{x}_i, \mathbf{x}_j) = vK_{\text{RBF}} + (1-v)K_{\text{poly}}, v \in [0, 1] \quad (8)$$

式中 v 是混合函数的权重系数, 可以给全局核函数和局部核函数分配不同的线性权值。此外, 所有参数都会影响 HKELM 模型的性能。因此, 为了实现 HKELM 模型良好的学习能力和泛化能力, 最优核参数的确定具有重要意义^[23]。

1.2 基于自动编码器的深层混合核极限学习机

基于自动编码器的深层混合核极限学习机能够捕获数据中的非线性特征。网络由输入层、 m 个堆叠的隐含层、核映射层、全连接层以及输出层构成。自编码器通过对数据的逐层抽取, 能够有效提取数据特征, 有利于区分容易混淆的类型, 提高分类精度。多输入、单输出的广义单隐含层前馈神经网络的输出函数可以表示为

$$f_i(x) = \sum_{i=1}^l \beta_i h_i(x) = \mathbf{H}(x)\beta \quad (9)$$

为了提高模型的稳定性, 当训练样本数大于隐含层节点数时, β 采用下式计算

$$\beta = \left(\frac{\mathbf{I}}{\lambda} + \mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{T} \quad (10)$$

基于自动编码器的 HKELM 模型必须满足目标输入等于输出这一条件, 其主要目的是学习输入特征中有意义的压缩、稀疏或相等的表示。在 DHKELM 模型中, 选择具有正交随机权值和偏差的加性节点作为隐含节点^[24]。

DHKELM 模型如图 1 所示, 该模型由编码器

和解码器组成的。编码器部分通过对隐含层和核映射层的压缩来初始化隐含层的权值 $\{\beta_i\}_{i=1}^m$ 。初始化过程可总结如下:

- (1) 随机生成输入层和隐含层的正交参数 a_i 和 b_i ;
- (2) 计算暂时隐含层状态矩阵 H_i , $H_i =$

$\left(\frac{I}{\lambda} + H^T H\right)^{-1} T$, 并将 $\sigma(x) = \frac{1}{1+e^{-x}}$ 作为非线性激活函数 $g(x)$;

- (3) 第 $i+1$ 个隐含层的权重 β^{i+1} 通过下式计算

$$\beta_{i+1} = \left(\frac{I}{\lambda} + H_i^T H_i\right)^{-1} H_i^T H_i \quad (11)$$

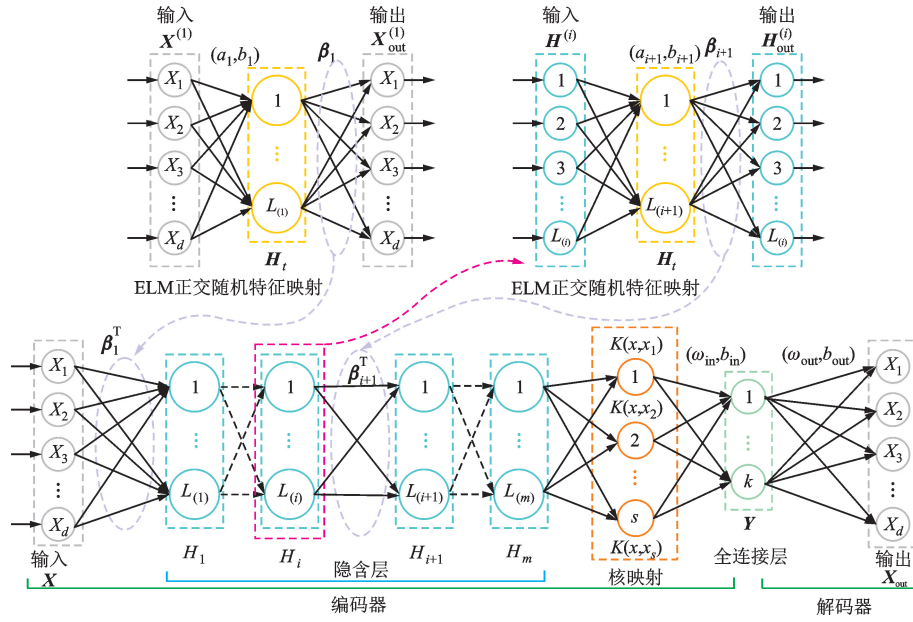


图 1 DHKELM 模型

Fig. 1 DHKELM model

然后,将这些隐含层堆叠起来。将前一层的输出 H_i 作为当前层的输入 H_{i+1} , 形成层次结构

$$\begin{cases} H_0 = X \\ H_{i+1} = g(\beta_{i+1}^T H_i), i = 0, 1, \dots, m-1 \end{cases} \quad (12)$$

完全连接层的输出 Y 表达式如下

$$Y = g(\omega_{in} H_m + b_{in}) \quad (13)$$

通过微调手段,可得到最后一层隐含层 H_m 与编码器的输出 Y 之间的参数。最后,将解码器部分设置为传统的神经全连接层。可以得到重建数据

$$\begin{aligned} X_{out} &= g(\omega_{out} Y + b_{out}) = \\ &g(\omega_{out} g(\omega_{in} H_m + b_{in}) + b_{out}) \end{aligned} \quad (14)$$

DHKELM-AE 的目标函数 F_L 是最小化重构数据和输入数据的误差,数学表达式如下

$$F_L = \arg \min_{\omega_{in}, b_{in}, \omega_{out}, b_{out}} \|X_{out} - X\|_2^2 \quad (15)$$

式中 ω_{in} 、 b_{in} 、 ω_{out} 、 b_{out} 分别为进入全连接层前后的权值和偏差。

DHKELM 中的隐含层权值、隐含层状态矩阵

可通过直接计算得到,模型的网络结构参数在初始化后就不需要重复训练计算,因此只需要使用反向传播算法对最后两层之间的权值参数进行训练。与传统的深度自动编码器相比,DHKELM-AE 中需要进行微调的参数数量大幅下降,进而实现效率的大幅提升。

2 野犬算法及其改进

2.1 传统野犬优化算法

受澳洲野犬的觅食和狩猎行为的启发,基于群启发式的野犬算法于 2021 年被提出。野犬的狩猎行为是可变的,它的狩猎策略包括群体攻击、迫害攻击、食腐行为,此外还考虑了野犬的生存概率。这些搜索策略有利于更好地探索目标空间,并且在求解优化问题时具有更高的成功率,与其他传统竞争方法相比,该算法适用性强,能快速地向着目标收敛。

野犬种群在搜索边界内随机初始化

$$\mathbf{X}_i = u_i + r_i(d_i - u_i) \quad (16)$$

式中: $\mathbf{X}_i$ 为初始个体; u_i 和 d_i 分别表示个体的上下边界; r_i 是 $[0, 1]$ 之间的随机数。

策略 1 群体攻击。当它们发现猎物时, 会以猎物为中心形成包围圈, 然后分工狩猎, 该捕猎行为如下式

$$\mathbf{X}_i(t+1) = \beta_i \sum_{k=1}^d \frac{\boldsymbol{\varphi}_k(t) - \mathbf{X}_i(t)}{d} - \mathbf{X}_*(t) \quad (17)$$

式中: t 为当前的迭代次数; $\mathbf{X}_i(t+1)$ 是野犬新位置 (表征野犬的移动); d 是在 $[2, N/2]$ 的逆序中生成的随机整数; N 是野犬种群的总规模; $\boldsymbol{\varphi}_k(t)$ 是野犬种群的子集 (将要发动攻击的野犬), 其中 $\varphi \in W$, W 是随机生成的野犬种群; $\mathbf{X}_i(t)$ 是当前野犬的位置; $\mathbf{X}_*(t)$ 是上一次迭代中发现的最佳野犬; β_i 是在 $[-2, 2]$ 内均匀生成的随机数, 它是一个用来控制野犬狩猎轨迹的比例因子。

策略 2 迫害攻击。该捕猎行为如下式

$$\mathbf{X}_i(t+1) = \mathbf{X}_*(t) + \beta_1 e^{\beta_2} (\mathbf{X}_{r_1}(t) - \mathbf{X}_i(t)), \quad i \neq r_1 \quad (18)$$

式中: β_2 是在 $[-1, 1]$ 区间内均匀生成的随机数; r_1 是在从 1 到最大搜索代理 (野犬) 大小的区间内生成的随机数; $\mathbf{X}_{r_1}(t)$ 是随机选择的第 r_1 个野犬。

策略 3 食腐行为。食腐行为即野犬在它们的栖息地寻找腐肉充饥的行为, 该捕猎行为如下式

$$\mathbf{X}_i(t+1) = \frac{1}{2} [e^{\beta_2} \mathbf{X}_{r_1}(t) - (-1)^\sigma \mathbf{X}_i(t)] \quad (19)$$

式中 σ 由算法随机生成, $\sigma \in \{0, 1\}$ 。

策略 4 野犬的存活率。根据大自然优胜劣汰原则, 考虑野犬存活率。通过设定存活率 S 阈值, 来淘汰、更新野犬种群。继而根据野犬的存活率来执行不同操作, 该捕猎行为如下式

$$S(i) = \frac{f_{\max} - f(i)}{f_{\max} - f_{\min}} \quad (20)$$

式中: $f(i)$ 是第 i 个野犬的当前适应度, 生存向量包含 $[0, 1]$ 区间内的归一化适应度。

当存活率小于或等于 0.3, 即 $S \leq 0.3$ 时, 用下式更新野犬位置

$$\mathbf{X}_i(t) = \mathbf{X}_*(t) + \frac{1}{2} [\mathbf{X}_{r_1}(t) - (-1)^\sigma \mathbf{X}_{r_2}(t)] \quad (21)$$

式中: r_2 是在从 1 到最大搜索代理 (野犬) 大小的区

间内生成的随机数, $r_1 \neq r_2$; $\mathbf{X}_{r_2}(t)$ 是随机选择的第 r_2 个野犬。

澳洲野犬通过 4 种觅食策略进行合作, 然后找到猎物。在 DOA 优化过程中, 首先在搜索空间中随机生成种群。设置参数 $P=0.5$, $Q=0.7$ 。随机数 $r < P$ 则执行攻击策略, 若 $r < Q$, 则执行群体攻击, 否则执行迫害攻击。随机数 $r \geq P$ 则执行食腐行为。计算野犬的存活率, 将 $S \leq 0.3$ 的野犬进行位置更新, 并重新计算适应度, 选择适应度最好的野犬位置作为最优野犬位置, 直到迭代结束, 返回最优的野犬位置。

2.2 改进野犬优化算法

传统野犬算法策略多变, 在测试函数的测试效果上远超其他算法, 但仍然有提升空间, 尤其是局部探索能力不足。本文将无限折叠混沌映射、精英反向学习以及差分进化算法融入 DOA 中, 以提高算法的搜索能力。

针对随机初始化种群分布不均的问题, 采用无限折叠迭代混沌映射和精英反向学习策略对 DOA 的种群进行初始化, 以提高种群的质量。研究表明, 相较于其他混沌映射, 无限折叠混沌映射最为均匀^[25], 其表达式如下

$$x_{n+1} = \sin(c\pi/x_n) \quad (22)$$

经过多次试验, 控制参数 c 设置为 0.7。

无限折叠混沌映射生成 N_p 个 d 维初始解 $\mathbf{X}_{i,j}$ ($i=1, 2, \dots, N_p$; $j=1, 2, \dots, d$) 后, 对初始解进行排序, 根据每个初始解对应的极点生成精英个体, 根据下式生成精英反向解 $\hat{\mathbf{X}}_{i,j}^e$, 对越界的位置点设置动态边界

$$\hat{\mathbf{X}}_{i,j}^e = k(\alpha_j + \beta_j) - \mathbf{X}_{i,j}^e \quad (23)$$

$$\hat{\mathbf{X}}_{i,j}^e = r(\alpha_j, \beta_j), \quad \hat{\mathbf{X}}_{i,j}^e < \alpha_j \cup \hat{\mathbf{X}}_{i,j}^e < \beta_j \quad (24)$$

式中: $k \in (0, 1)$ 为精英反向系数; $\alpha_j = \min(\mathbf{X}_{i,j}^e)$; $\beta_j = \max(\mathbf{X}_{i,j}^e)$ 。

最后通过将无限折叠混沌映射生成的初始解与精英反向解混合排序, 初始种群选取为前 N_p 个较优解。

差分进化算法具有不易陷入局部最优的优点, 因此将其用于优化野犬种群中适应度小于平均适应度的个体, 通过变异、交叉和选择操作来重组种群, 保留最优野犬个体, 进而得到差分进化后的野犬种群的最优个体, 然后通过原始种群个体和变异个体的竞争来获取每一代野犬种群中的最优个体, 引导搜索过程向最优解逼近。算法流程如图 2 所示。

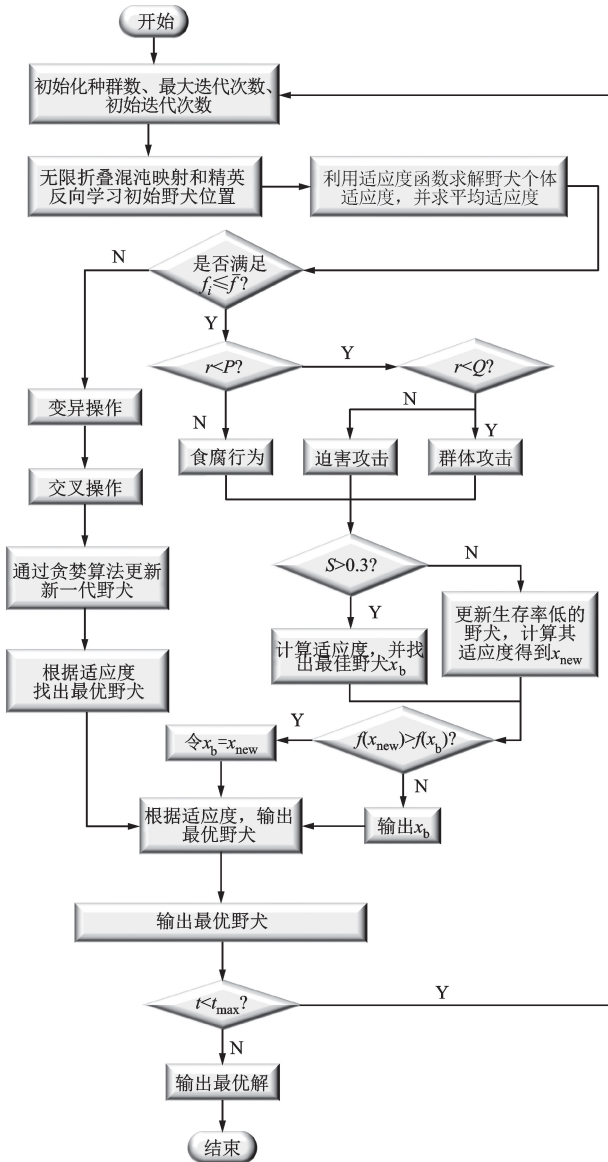


图 2 IDOA 流程图

Fig. 2 Flow chart of IDOA

3 建立 IDOA-DHKELM 短期风电功率预测模型

在 DHKELM 模型中,输入层与隐含层之间的输入权值和偏置正交参数是随机生成的,不同核函数的核参数以及权值的确定也会影响模型的预测精度。不可避免的会出现产生的随机值并非最优的情况,一方面 DHKELM 相较于其他 ELM 网络需要更多的隐含层神经元,另一方面使得到的隐含层输出矩阵出现较大偏差,进而使模型的泛化能力大大减弱,输出也极不稳定,因此对 DHKELM 的核函数权值、核参数、隐含层数、隐含层神经元数正则化系数、惩罚系数进行优化十分必要。

本文采用 IDOA 算法对 DHKELM 参数进行优化,利用 IDOA 算法的全局及局部寻优能力,找出以上 8 个参数的最优值。将优化后的参数代入到 DHKELM 中对短期风电功率进行预测。IDOA-DHKELM 预测模型建立步骤如下。

步骤 1 参数初始化。设置野犬种群数、种群维数、最大迭代次数, $P=0.5$, $Q=0.7$ 。

步骤 2 初始化野犬种群。使用无限折叠混沌映射和精英反向学习对野犬种群初始化。

步骤 3 计算野犬种群平均适应度 $\bar{f}(\cdot)$ 。将所有权值和偏置的组合作为每个野犬个体的位置,并计算每一只野犬的适应度。求得平均适应度后,通过判断个体适应度分别执行不同的策略。适应度差的个体通过执行变异、交叉操作来更新个体位置。适应度好的个体执行捕食策略。

步骤 4 考虑野犬个体存活率。策略 4 中存活率阈值 $S=0.3$, 并利用式(21)更新野犬种群个体的位置为 x_{new} 。

步骤 5 通过对比进化策略和捕食策略得出的最佳适应度 x_b 来得出最佳个体的位置。

步骤 6 计算野犬新位置的适应度,并与上一次迭代的适应度比较。若适应度优于上一代,则更新,否则不变。

步骤 7 判断是否已达到最大迭代次数 t_{max} 。若是则结束,并输出最优参数的组合,否则跳转至步骤 3。

步骤 8 利用优化后的 DHKELM 模型对短期风电功率进行预测,并对结果进行分析。

基于 IDOA-DHKELM 的短期风电功率预测流程如图 3 所示。

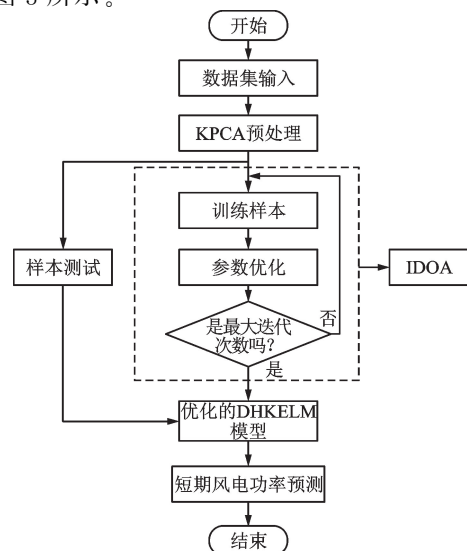


图 3 基于 IDOA-DHKELM 的短期风电功率预测流程图

Fig. 3 Flow chart of short-term wind power forecasting based on IDOA-DHKELM

4 算例分析

本文采用国外某风电场官网提供的实测数据集,该风电场数据包括风机 10 m、50 m、90 m 共 3 个高度处的 NWP 数据和风电场实际发电功率值。分析了风电场风速、风向正弦值、风向余弦值、温度、空气密度、压强 6 个类别的气象因素,分别位于 3 个高度位置共 18 维特征的 NWP 数据。设定累计贡献率阈值为 90%,在 KPCA 预处理后前 5 个特征值的累计贡献率 η 可达 90.37%,降为 5 个主成分,协方差矩阵特征值及累计贡献率如图 4 所示。选取 2020 年 2 月 1 日至 2020 年 3 月 31 日的风电功率数据,采样周期为 15 min,数据集共计 5 760 条风电功率数据,每天共产生 96 条数据。将由主成分和历史功率构成的 5 760 条数据划分为训练集和测试集。将前 57 天数据用于 IDOA-DHKELM 模型训练,最后 3 天的数据用于测试。对预测值与风电功率真实值进行对比分析。

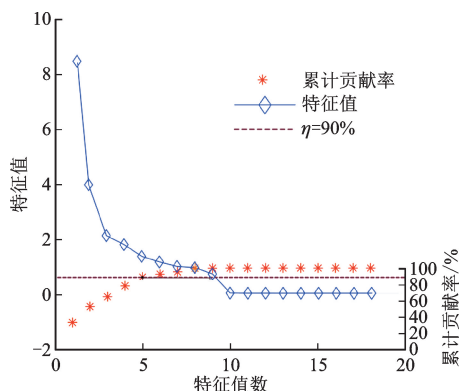


图 4 协方差矩阵特征值及累计贡献率

Fig. 4 Eigenvalues of the covariance matrix and its cumulative contribution

选择常用的风电功率预测评价指标,并将本文所提的 IDOA-DHKELM 预测模型所得结果与传统的 ELM、SKELM、HKELM、PSO-DHKELM、DE-DHKELM 以及 DOA-DHKELM 模型得到的预测结果进行对比分析,验证本文所提方法的有效性和优越性。

4.1 数据归一化及评价指标

为了避免因不同变量间的数量级产生的误差,因此对数据进行预处理十分必要。为防止将未来风电功率数据泄露到训练集中,将划分后的训练集和测试集分别归一化,且测试集的归一化是建立在训练集的基础上的,测试集最终采用与训练集相同的

归一化参数。本文利用最大最小归一化,统一将数据缩小至 $[0,1]$ 之内,归一化公式如下式

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (25)$$

式中: $x_{\min}$ 和 $x_{\max}$ 分别为 x 变量中的最大值和最小值; x 为待归一化的变量; x' 为归一化后的变量。

为了更好地评估不同风电功率预测模型的精度,本文将平均绝对误差、均方根误差和平均绝对误差百分比作为评价指标,表达式如下

$$e_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |\hat{p}_i - p_i| \quad (26)$$

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{p}_i - p_i)^2} \quad (27)$$

$$e_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{p}_i - p_i|}{p_i} \times 100\% \quad (28)$$

式中: n 为风电功率测试集样本数; p_i 为风电功率真实值; $\hat{p}_i$ 为风电功率预测值。 e_{MAE} 、 e_{RMSE} 和 e_{MAPE} 越小,则表示拟合效果越好,即模型的预测精度越高,预测性能越强。

4.2 IDOA 的优化性能测试

为验证上述 IDOA 算法的性能,选取 Schwefel 单峰测试函数 f_1 和 Rastrigin 多峰测试函数 f_2 作为本文的两个测试函数。

f_1 测试函数外部区域分布很多局部最小值,中心凹进去的位置是整个函数的极小点。函数定义域为 $x \in [-100, 100]$,表达式如下

$$f_1(x) = \max_i (|x_i|), \quad 1 \leq i \leq n \quad (29)$$

f_2 测试函数在定义域内分布大量局部最小值,对寻找位于 $(0,0)$ 处的全局最小值增加了难度。函数定义域为 $x \in [-5.12, 5.12]$,表达式如下

$$f_2(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10] \quad (30)$$

不同测试函数的取值范围、最优解等信息如表 1 所示,不同算法的参数设置如表 2 所示。

通过以上两种测试函数检验 IDOA 的性能,并与传统 PSO、DE 及 DOA 进行比较分析。测试结果如图 5、图 6 所示。从图 5、6 可以看出,在 $f_1(x)$ 和 $f_2(x)$ 测试中,4 种算法都随着迭代次数的增加逐渐收敛。PSO 的收敛速度明显慢于 DOA 和 IDOA,需要的迭代次数更多;IDOA 较 DOA、DE 和 PSO 的收敛速度快,迭代次数少,由此可以看出 IDOA 算法具有更好的稳定性和寻优能力,能更快速且准确地得到最优解。

表 1 测试函数

Table 1 Test function

函数	类型	维度	取值范围	最优解
f_1	Schwefel	30	$[-100, 100]$	0
f_2	Rastrigin	30	$[-5, 5]$	0

表 2 不同算法的参数设置

Table 2 Parameter settings of different algorithms

算法	参数
PSO	$c_1 = c_2 = 2, \omega = 0.8$
DE	$F = 0.6, C_R = 0.1$
DOA	$\beta_1 \in [-2, 2], \beta_2 \in [-1, 1],$ $P = 0.5, Q = 0.7, S = 0.3$
IDOA	$\beta_1 \in [-2, 2], \beta_2 \in [-1, 1], c = 0.7,$ $P = 0.5, Q = 0.7, S = 0.3, F = 0.6, C_R = 0.1$

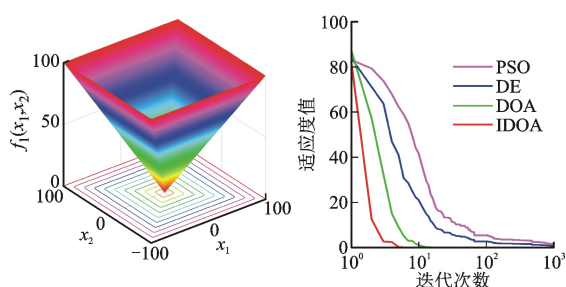


图 5 Schwefel 函数测试结果

Fig. 5 Test results of generalized Schwefel's function

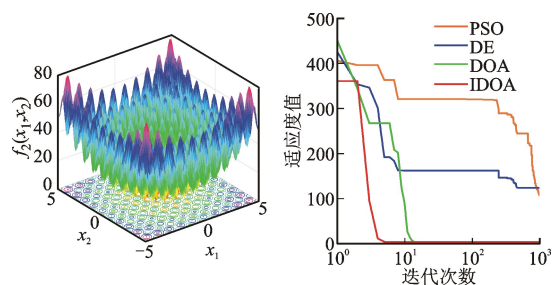


图 6 广义 Rastrigin 函数测试结果

Fig. 6 Test results of generalized Rastrigin's function

4.3 单一预测模型对比实验及分析

为了验证 DHKELM 模型的有效性,分别建立反向学习神经网络、支持向量机、长短时记忆网络、极限学习机、单核极限学习机 5 种对比模型,其网络参数均由交叉验证法所得。表 3 为不同模型的参数设置,表 4 为不同模型的预测误差,图 7 为不同模型的预测曲线。

表 3 不同模型的参数设置

Table 3 Parameter settings of different models

模型	参数	数值
BPNN	隐含层数	1
	神经元数	100
SVM	惩罚参数	0.5
	核参数	0.5
LSTM	学习率	0.005
	隐藏单元数	10
ELM	隐含层数	1
	神经元数	100
SKELM	核函数	RBF
	核参数	100
	正则化系数	100
DHKELM	核函数	RBF, Poly
	权重	0.5
	RBF 核参数	100
	Poly 核参数	100, 0.5
	隐含层数量	5
	神经元个数	50
	正则化系数	100
	惩罚系数	100

表 4 不同模型的预测误差

Table 4 Prediction errors of different models

模型	e_{RMSE}/MW	e_{MAE}/MW	$e_{MAPE}/\%$
BPNN	829.055 7	738.011 8	19.605 1
SVM	781.165 8	648.226 3	18.572 1
LSTM	637.739 8	561.511 6	14.905 9
ELM	607.793 6	482.815 5	13.637 4
SKELM	573.144 3	395.743 6	10.937 8
HKELM	387.609 4	339.263 2	8.556 4
DHKELM	290.755 5	241.323 6	6.353 4

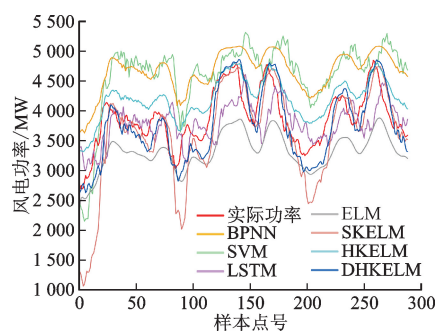


图 7 不同模型的预测曲线

Fig. 7 Prediction curves of different base models

由表 4 可知,在 5 种常见的神经网络模型中,BPNN 模型得到的预测结果最差,这是由于其在训练时已陷入局部最优,导致其预测结果很不稳定,而且误差较大;SVM 和 LSTM 的预测结果仅优于 BPNN,这是由于其预测性能受关键参数的影响较为严重,即使利用交叉验证法得到的参数也不是性能最好的。ELM 的预测结果较为稳定,该模型结构简单,隐藏层仅为单层,只需要人为设定神经元数这一个参数,在单一模型中,ELM 对风电功率数据的训练和预测所花费时间较短,且精度较高。模型的 e_{MAPE} 在单一模型中最低,综合考虑后选用 ELM 模型。在 ELM 中引入 RBF 核函数的 SKELM 模型的预测结果更为稳定,模型的 e_{MAPE} 降为 10.937 8%,这是由于将传统 ELM 随机映射替代成了核映射,进而将高复杂低维的空间问题转化为高维空间内积运算问题,使得 SKELM 模型输出稳定性和泛化能力大大提高。

在所有模型中,DHKELM 的预测精度最高,并且预测结果很稳定,模型的 e_{MAPE} 仅为 6.353 4%,相比于 BPNN、SVM、LSTM、ELM、SKELM 和 HKELM 分别降低了 13.251 7%、12.218 7%、8.552 5%、7.284%、4.584 4%和 2.203%。值得注意的是,HKELM 与 SKELM 的模型结构类似,HKELM 内含两个核函数,而 SKELM 只含有单一的核函数,MKELM 的最终预测结果明显好于 SKELM,这是由于单一的核函数存在学习能力与泛化能力冲突的问题,多种核函数结合能够弥补单一核函数的缺陷。综上所述,DHKELM 模型能够更好地挖掘风电功率内部的规律,具有更强的预测性能,所得结果更为精准且稳定。

4.4 IDOA 优化性能实验及分析

为了验证 IDOA-DHKELM 模型的有效性和优越性。对单一的 DHKELM 模型参数进行优化预测,同时将粒子群优化算法、差分进化算法和未改进的野犬算法作为对比算法。为实现各优化算法的对比分析,设置所有算法的种群数为 30,最大迭代次数为 100,DHKELM 的 8 个参数寻优范围分别为 $[0, 1]$ 、 $[1, 10^3]$ 、 $[1, 10^3]$ 、 $[10^{-2}, 10]$ 、 $[1, 10]$ 、 $[1, 10^2]$ 、 $[1, 10^3]$ 和 $[1, 10^3]$ 。

表 5 为不同算法优化 DHKELM 模型的预测误差结果。由表 5 可知,利用 IDOA 优化 DHKELM 得到的预测误差最小,由此也进一步说明 IDOA 能够更有效地提高 DHKELM 的稳定性和预测性能。

表 5 不同算法优化 DHKELM 的预测误差对比

Table 5 Comparison of prediction errors of DHKELM model optimized by different algorithms

模型	e_{RMSE}/MW	e_{MAE}/MW	$e_{MAPE}/\%$
PSO-DHKELM	194.700 2	177.428 2	4.557 2
DE-DHKELM	185.900 5	143.381 0	3.516 1
DOA-DHKELM	102.304 5	84.870 4	2.153 3
IDOA-DHKELM	57.058 9	44.646 9	1.174 0

图 8 为 IDOA-DHKELM 模型得到的预测结果,其中红线代表真实值,黄线代表预测值。由图 8 可知,相比其他模型来说,IDOA-DHKELM 表现出良好的精度,得到的拟合曲线也十分接近真实曲线,由此也验证了所提模型的有效性。

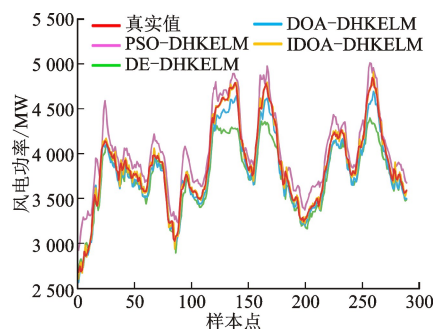


图 8 不同算法优化 DHKELM 的预测曲线

Fig. 8 Different algorithms optimize the prediction curve of DHKELM

4.5 多步预测实验

为了进一步验证所提模型的优越性,在单步预测的基础上,又进行了多步预测实验。单步预测是指当前 k 时刻的数据来预测第 $k+1$ 时刻的功率,而多步预测是用当前 k 时刻的数据来预测第 $k+2$ 、 $k+3$ 等时刻的功率,多步预测比单步预测更能有效地捕捉风电功率序列未来的动态性能,可以进一步分析模型的性能。各风电功率预测模型的误差对比如表 6 所示,多步预测曲线如图 9 所示。

随着预测步数增多,所有模型的预测精度均有所下降,DHKELM 模型在 1 步、2 步以及 4 步的 e_{MAE} 分别为 241.323 6 MW、282.756 7 MW 以及 337.208 3 MW。这是由于多步预测的难度高于单步预测。基于算法优化的模型得到的预测精度要优于单一的模型。相较于 DHKELM 模型,在 1 步、2 步以及 4 步中 IDOA-MKELM 得到的 e_{MAPE} 分别下降了 5.179 4%、5.736 9%和 6.201 9%。

表6 各风电功率预测模型多步误差

Table 6 Multi-step error of each wind power prediction model

模型	预测步数	e_{RMSE}/MW	e_{MAE}/MW	$e_{MAPE}/\%$
DHKELM	1步	290.755 5	241.323 6	6.353 4
	2步	357.195 4	282.756 7	7.884 7
	4步	425.962 1	337.208 3	9.720 8
DOA-DHKELM	1步	102.304 5	84.870 4	2.153 3
	2步	168.495 2	129.845 3	3.237 1
	4步	198.599 1	173.811 6	4.910 9
IDOA-DHKELM	1步	57.058 9	44.646 9	1.174 0
	2步	105.490 3	81.795 7	2.147 8
	4步	177.224 8	136.217 0	3.518 9

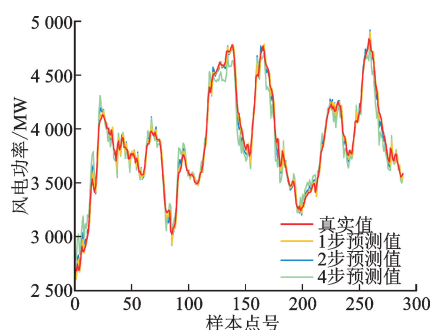


图9 IDOA-DHKELM的多步预测曲线

Fig. 9 Multi-step prediction curves of IDOA-DHKELM

以最小化均方误差为目标,利用改进野犬优化算法优化 DHKELM 模型的 8 个参数,模型的预测误差进一步下降,且其下降幅度比基本的野犬优化算法大,从而进一步验证了 IDOA 算法的优越性。在 1 步、2 步以及 4 步中 IDOA-DHKELM 的 e_{MAPE} 相较于 DOA-DHKELM 模型分别下降了 0.979 3%、1.089 3% 和 1.393 1%。

5 结 论

本文提出了一种基于 KPCA 特征选择和改进野犬算法优化深层混合核极限学习机的短期风电功率预测方法,结合实际算例验证,得出具体结论如下:

(1)利用 KPCA 方法对特征进行优选得到的最优特征集,既能表达风电功率的有效信息,也能避免冗余信息的出现,有利于 DHKELM 模型的学习与训练,同时也降低了模型的复杂度,提高了风电功率预测精度。

(2)针对 DHKELM 人工调参量大且准确性差的问题,利用多种策略对 DOA 进行改进,使算法的搜索能力增强,能够跳出局部最优值。仿真验证 IDOA 能够克服 DOA 算法易陷入局部最优解的缺点,较 PSO、DE 等单一的算法具有更强的稳定性和寻优能力。利用 IDOA 算法对 DHKELM 的 8 个参数进行优化,获得参数的全局最优解,充分发挥深层混合核极限学习机的性能。通过实例分析及对比实验,KPCA-IDOA-DHKELM 模型获得的 4 个评价指标均优于其他模型,具有更好的预测精度和泛化能力。

参考文献:

- [1] 黄雨涵,丁涛,李雨婷,等. 碳中和背景下能源低碳化技术综述及对新型电力系统发展的启示 [J]. 中国电机工程学报, 2021, 41(z1): 28-51.
HUANG Yuhuan, DING Tao, LI Yuting, et al. Decarbonization technologies and inspirations for the development of novel power systems in the context of carbon neutrality [J]. Proceedings of the CSEE, 2021, 41(z1): 28-51.
- [2] 帅轩越,王秀丽,王志成,等. 计及风电消纳下基于纳什议价理论的多综合能源系统协同运行 [J]. 西安交通大学学报, 2022, 56(3): 187-196.
SHUAI Xuanyue, WANG Xiuli, WANG Zhicheng, et al. Cooperative operation for multiple integrated energy systems considering wind power consumption by Nash bargaining convention [J]. Journal of Xi'an Jiaotong University, 2022, 56(3): 187-196.
- [3] 王丽婕,冬雷,高爽. 基于多位置 NWP 与主成分分析的风电功率短期预测 [J]. 电工技术学报, 2015, 30(5): 79-84.
WANG Lijie, DONG Lei, GAO Shuang. Wind power short-term prediction based on principal component analysis of NWP of multiple locations [J]. Transactions of China Electrotechnical Society, 2015, 30(5): 79-84.
- [4] 杨家然,王兴成,罗晓芬,等. 基于数值天气预报及模糊聚类的风电功率智能组合预测 [J]. 太阳能学报, 2017, 38(3): 669-675.
YANG Jiaran, WANG Xingcheng, LUO Xiaofen, et al. Intelligent combined prediction of wind power based on numerical weather prediction and fuzzy clustering [J]. Acta Energiae Solaris Sinica, 2017, 38(3): 669-675.
- [5] ZHENG Linyu, HE Hongmei. Share price prediction of aerospace relevant companies with recurrent neural

- networks based on PCA [J]. *Expert Systems with Applications*, 2021, 183: 115384.
- [6] 吴定安, 钟建伟, 王新磊, 等. 主成分分析和长短期记忆网络的电力负荷预测 [J]. *物联网技术*, 2021, 11(8): 47-51.
- WU Ding'an, ZHONG Jianwei, WANG Xinlei, et al. Principal component analysis and long and short-term memory networks for electric load forecasting [J]. *Internet of Things Technologies*, 2021, 11(8): 47-51.
- [7] 张云钦, 程起泽, 蒋文杰, 等. 基于 EMD-PCA-LSTM 的光伏功率预测模型 [J]. *太阳能学报*, 2021, 42(9): 62-69.
- ZHANG Yunqin, CHENG Qize, JIANG Wenjie, et al. Photovoltaic power prediction model based on EMD-PCA-LSTM [J]. *Acta Energetica Sinica*, 2021, 42(9): 62-69.
- [8] 向玲, 李京蓄, 王朋鹤, 等. 基于 VMD-FIG 和参数优化 GRU 的风速多步区间预测 [J]. *太阳能学报*, 2021, 42(10): 237-242.
- XIANG Ling, LI Jingxu, WANG Penghe, et al. Wind speed multistep interval forecasting based on VMD-FIG and parameter-optimized GRU [J]. *Acta Energetica Sinica*, 2021, 42(10): 237-242.
- [9] 徐耀松, 王德兴. 基于加权灰色关联投影的 Bagging-Blending 多模型融合短期电力负荷预测 [J]. *信息与控制*, 2022, 51(2): 188-200.
- XU Yaosong, WANG Dexing. Bagging-Blending multi-model fusion short-term power load forecasting based on weighted gray relational projection [J]. *Information and Control*, 2022, 51(2): 188-200.
- [10] 盛四清, 金航, 刘长荣. 基于 VMD-WSGRU 的风电场发电功率中短期及短期预测 [J]. *电网技术*, 2022, 46(3): 897-904.
- SHENG Siqing, JIN Hang, LIU Changrong. Short-term and mid-short-term wind power forecasting based on VMD-WSGRU [J]. *Power System Technology*, 2022, 46(3): 897-904.
- [11] QU Xiaoyun, KANG Xiaoning, ZHANG Chao, et al. Short-term prediction of wind power based on deep long short-term memory [C]//2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC). Piscataway, NJ, USA: IEEE, 2016: 1148-1152.
- [12] GREFF K, SRIVASTAVA R K, KOUTNÍK J, et al. LSTM: a search space odyssey [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(10): 2222-2232.
- [13] 商立群, 李洪波, 侯亚东, 等. 基于特征选择和优化极限学习机的短期电力负荷预测 [J]. *西安交通大学学报*, 2022, 56(4): 165-175.
- SHANG Liqun, LI Hongbo, HOU Yadong, et al. Short-term power load forecasting based on feature selection and optimized extreme learning machine [J]. *Journal of Xi'an Jiaotong University*, 2022, 56(4): 165-175.
- [14] 商立群, 李洪波, 侯亚东, 等. 基于 VMD-ISSA-KELM 的短期光伏发电功率预测 [J/OL]. *电力系统保护与控制* [2022-07-02]. DOI: 10.19783/j.cnki.pspc.220140.
- SHANG Liqun, LI Hongbo, HOU Yadong, et al. Short-term photovoltaic power generation prediction based on VMD-ISSA-KELM [J/OL]. *Power System Protection and Control* [2022-07-02]. DOI: 10.19783/j.cnki.pspc.220140.
- [15] CHU Yunfei, FENG Chunyan, GUO Caili, et al. Network embedding based on deep extreme learning machine [J]. *International Journal of Machine Learning and Cybernetics*, 2019, 10(10): 2709-2724.
- [16] ZHU Shunpeng, KESHTEGAR B, BEN SEGHER M E A, et al. Hybrid and enhanced PSO: novel first order reliability method-based hybrid intelligent approaches [J]. *Computer Methods in Applied Mechanics and Engineering*, 2022, 393: 114730.
- [17] DHIMAN G, KAUR A. STOA: a bio-inspired based optimization algorithm for industrial engineering problems [J]. *Engineering Applications of Artificial Intelligence*, 2019, 82: 148-174.
- [18] RAMESH KHAPARDE A, ALASSERY F, KUMAR A, et al. Differential evolution algorithm with hierarchical fair competition model [J]. *Intelligent Automation & Soft Computing*, 2022, 33(2): 1045-1062.
- [19] ALMAZÁN-COVARRUBIAS J H, PERAZA-VÁZQUEZ H, PEÑA-DELGADO A F, et al. An improved dingo optimization algorithm applied to SHE-PWM modulation strategy [J]. *Applied Sciences*, 2022, 12(3): 992.
- [20] XIA Xin, WANG Xiaolu. A novel hybrid model for short-term wind speed forecasting based on twice decomposition, PSR, and IMVO-ELM [J]. *Complexity*, 2022, 2022: 4014048.
- [21] ABUROMMAN A A, IBNE REAZ M B. A novel SVM-kNN-PSO ensemble method for intrusion detection system [J]. *Applied Soft Computing*, 2016, 38: 360-372.
- [22] ABUROMMAN A A, IBNE REAZ M B. A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection

- systems [J]. Information Sciences, 2017, 414: 225-246.
- [23] JIA Jianfang, YUAN Shufang, SHI Yuanhao, et al. Improved sparrow search algorithm optimization deep extreme learning machine for lithium-ion battery state-of-health prediction [J]. iScience, 2022, 25 (4): 103988.
- [24] KHAN J, FAYAZ M, HUSSAIN A, et al. An improved alpha beta filter using a deep extreme learning machine [J]. IEEE Access, 2021, 9: 61548-61564.
- [25] 黄倩, 刘升, 李萌萌, 等. 多策略黑猩猩优化算法研究及其工程应用 [J/OL]. 计算机工程与应用[2022-06-09]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20210806.1055.011.html>.
- HUANG Qian, LIU Sheng, LI Mengmeng, et al. Multi-strategy chimp optimization algorithm and its application of engineering problem [J/OL]. Computer Engineering and Applications [2022-06-09]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20210806.1055.011.html>.
- (编辑 武红江 亢列梅)