

徐武<sup>1</sup> 范鑫豪<sup>1</sup> 沈智方<sup>1</sup> 刘洋<sup>1</sup> 刘武<sup>2</sup>

# 基于 ikPCA-FABAS-KELM 的短期风电功率预测

## 摘要

为了增强在短期风电功率预测领域中传统数据驱动机器学习模型的精度,提出基于 ikPCA-FABAS-KELM 的短期风电功率预测模型.首先,对主成分分析进行改进,提出可逆核主成分分析(ikPCA),在保证数据特征的同时,降低输入数据的复杂度,以提升模型运行速度;其次,引入萤火虫个体吸引策略对天牛须算法(BAS)进行改进,提出 FABAS 算法;最后,利用 FABAS 算法对核极限学习机(KELM)的正则化参数  $C$  和核参数  $\gamma$  进行寻优,降低人为因素对模型盲目训练的影响,提高模型预测精度.仿真结果显示,提出的预测模型有效提高了传统模型的预测精度.

## 关键词

短期风电功率预测;萤火虫算法;天牛须算法;核主成分分析;核极限学习机

中图分类号 TM614

文献标志码 A

收稿日期 2023-08-10

资助项目 国家自然科学基金(U1802271)

作者简介

徐武,男,教授,研究方向为电力系统安全与控制、虚拟现实技术.1369925132@qq.com

1 云南民族大学 电气信息工程学院,昆明,650031

2 新疆独山子石化公司供水供电公司,克拉玛依,834000

## 0 引言

构建清洁低碳、安全稳定的新型电力系统已成为电力行业的中中之重,风电更是被业界誉为实现“双碳”目标的主力军,受到广泛的关注<sup>[1]</sup>.由于风能极易受到风速、风向、气压和温度等自然条件的影响而具有一定的间歇性、波动性和随机性,所以,当风电大规模并网时,会危及电网的安全和稳定运行<sup>[2]</sup>.

为降低大规模风电并网带来的影响,短期风电功率预测成为研究热点,涌现出多种物理模型和统计模型<sup>[3]</sup>.物理模型要求对大气条件和风电场的物理特性有准确的数学描述,然后通过求解计算,最终得到风电场的预测功率.而统计模型依赖大量历史统计数据的天气状况与风电功率之间的关系进行风电输出功率预测.两种模型各有优劣,物理方法数据需求不高,但计算复杂,统计模型计算速度快,但数据需求高.基于数据驱动的统计学模型,如支持向量机<sup>[4-5]</sup>、极限学习机<sup>[6-7]</sup>、神经网络<sup>[8-11]</sup>等,因其具备计算效率较高、模型简单的优势,得到了广泛应用.

已有的研究表明,传统的单一模型在进行短期风电功率预测时,普遍存在对复杂数据处理能力有限、模型训练效果不佳等问题,因此,考虑气象因素数据处理的组合预测模型被重点关注<sup>[12-13]</sup>.文献[14]利用变分模态分解对数据进行平稳化处理,提升了模型的预测精度.文献[15]利用弹性网稀疏核主成分分析方法,降低数据复杂性,提升了预测精度.文献[16]先利用自适应智能灰色系统进行风速预测,并和数值天气预报(Numerical Weather Prediction, NWP)数据整合为预测样本,随后采用遗传算法优化核极限学习机模型并进行风电功率预测,一定程度上提升了预测精度.

针对传统模型在短期风电功率预测方面存在的不足,本文提出一种基于 ikPCA-FABAS-KELM 组合预测模型.首先,引入傅里叶随机特征近似主成分分析的核函数,提出可逆核主成分分析(invertible kernel Principal Component Analysis, ikPCA),在保证数据特征的基础上降低数据复杂度,便于后续机器学习模型预测处理;其次,利用萤火虫算法(Firefly Algorithm, FA)的吸引策略改进天牛须算法(Beetle Antennae Search, BAS),提出 FABAS 算法;最后,利用 FABAS 算法对核极限学习机(Kernel Extreme Learning Machine, KELM)进行参数寻优,提升模型的预测性能和泛化能力.实例分析结果表明,上述模型在

短期风电功率预测时具有一定优势.

## 1 可逆核主成分分析

### 1.1 核主成分分析原理

针对传统机器学习预测模型存在难以处理复杂数据的问题,本文首先利用核主成分分析(kernel Principal Component Analysis, kPCA)对输入数据进行降维<sup>[17-18]</sup>.

kPCA的原理是遵从某种非线性映射关系,将非线性数据组从原始特征空间映射到高维的特征空间,随后进行主成分分析实现非线性数据的降维.非线性映射 $\varphi$ 即为核函数:

$$\varphi: \mathbf{x} \rightarrow \varphi(\mathbf{x}). \quad (1)$$

假设矩阵 $\mathbf{X}$ 的数据样本为 $\mathbf{x}_i \in \mathbf{R}^{m \times n}$ ,那么在映射到高维特征空间后,协方差矩阵 $\mathbf{C}$ 可表示为

$$\mathbf{C} = \frac{1}{m} \sum_{i=1}^m \varphi(\mathbf{x}_i) \varphi^T(\mathbf{x}_i). \quad (2)$$

矩阵 $\mathbf{C}$ 的特征值 $\lambda_i$ 和特征向量 $\mathbf{P}_i$ 为

$$\mathbf{C}\mathbf{P}_i = \lambda_i \mathbf{P}_i, \quad (3)$$

$$\mathbf{P}_i = \sum_{i=1}^m \alpha_i \varphi(\mathbf{x}_i). \quad (4)$$

式中, $\alpha_i$ 为线性系数矩阵.

计算向量 $\varphi(\mathbf{x})$ 在特征向量 $\mathbf{P}_i$ 上的投影,以提取非线性主成分:

$$\varphi(\mathbf{x})\mathbf{P}_i = \sum_{i=1}^m \alpha_i \varphi(\mathbf{x}) \varphi(\mathbf{x}_i). \quad (5)$$

$$\text{降维后的数据矩阵 } \hat{\mathbf{X}} \text{ 和方差贡献率 } \alpha \text{ 为} \\ \hat{\mathbf{X}} = \mathbf{X} \times \mathbf{P}_i, i = 1, 2, \dots, m, \quad (6)$$

$$\alpha = \frac{\lambda_i}{\sum_{i=1}^m \lambda_i}, i = 1, 2, \dots, m. \quad (7)$$

### 1.2 可逆核主成分分析

研究发现,kPCA实现数据降维时,如何从高维特征空间映射到输入空间的过程是难以准确解释的,这就可能导致降维后的数据出现数据特征丢失的问题.为尽可能地保证数据特征的完备性,提升降维特性,本文提出可逆核主成分分析方法(ikPCA),

其实现过程如图1所示.

改进策略如下:

首先,将前 $i$ 个分量的特征映射表示为 $\varphi(\mathbf{x}) = [\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x}), \dots, \varphi_i(\mathbf{x})]$ 来近似核,则 $k(\mathbf{x}, \mathbf{y}) \approx [\varphi(\mathbf{x}), \varphi(\mathbf{y})]$ .随机傅里叶特征认为,对于一个任意的平移不变核都可以用非线性函数 $\sigma$ 和 $\mathbf{W} \in \mathbf{R}^{r \times n}$ 特征映射近似表示.

将核函数近似为如式(8)所示的形式:

$$\varphi(\mathbf{x}) = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b}), \quad (8)$$

则可按照如下计算过程实现降维:

$$\begin{cases} \mathbf{v} = \mathbf{W}\mathbf{x} + \mathbf{b}, \\ \beta = \sigma(\mathbf{v}), \\ \hat{\mathbf{X}} = \beta\mathbf{P}. \end{cases} \quad (9)$$

为更好地解释高维特征空间映射到输入空间的过程,可对上述过程进行逐步反演.通常,近似的核函数 $\sigma$ 在整个定义域上是不可逆的,但在其子定义域内可能可逆.在逐步反演时,绕过不可逆部分,得:

$$\hat{\beta} = \mathbf{P}^{-1} \hat{\mathbf{X}}, \quad (10)$$

$$\hat{\mathbf{v}} = \sigma^{-1}(\hat{\beta}) + \bar{\mathbf{v}}, \quad (11)$$

$$\hat{\mathbf{x}} = \arg_x \min \|\mathbf{W}\mathbf{x} + \mathbf{b} - \hat{\mathbf{v}}\|^2 + \lambda \|\mathbf{x}\|^2. \quad (12)$$

式(10)实现反转降维,其原理和PCA方法相同;投影矩阵 $\mathbf{P}$ 使重构误差 $\|\hat{\beta} - \mathbf{P}^T \hat{\mathbf{X}}\|$ 最小;式(11)实现非线性函数 $\sigma$ 的反演,并加上绕过的不可逆部分 $\bar{\mathbf{v}}$ ;式(12)通过求解Ridge回归问题反转线性映射 $\mathbf{x} \rightarrow \mathbf{W}\mathbf{x} + \mathbf{b}$ .

通过上述降维过程的反演,可以对数据从特征空间映射到输入空间的过程进行解释,对保证数据特征的完备具有一定意义.

### 1.3 实验验证

为了验证ikPCA方法的性能,本研究选取鸢尾花数据集(www.kaggle.com)作为实验案例进行实验验证.通过对花瓣和花萼特征进行分析,将鸢尾花数据分为3个类别:setosa、versicolor和virginica.分别使用PCA、kPCA和ikPCA对数据集进行重构,相应的结果如图2—4所示.

由图2可知,经PCA降维后,setosa分类的分界

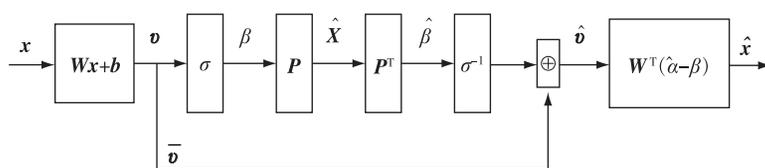


图1 ikPCA实现过程

Fig. 1 Implementation process of ikPCA

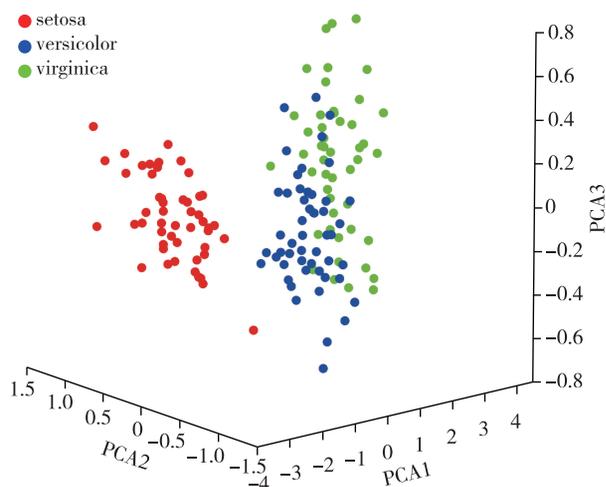


图2 PCA 重构主成分

Fig. 2 Principal components reconstructed by PCA

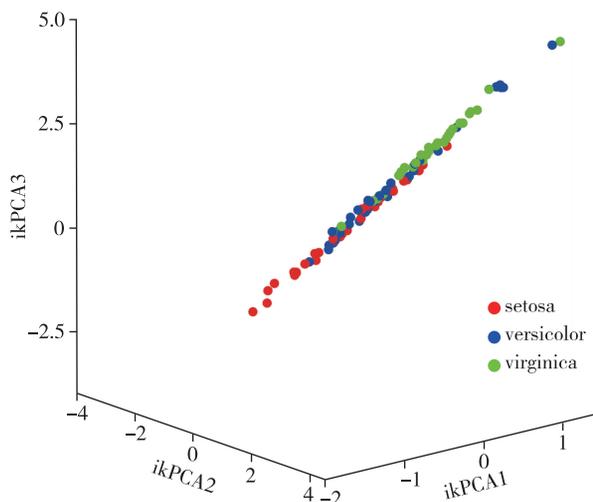


图4 ikPCA 重构主成分

Fig. 4 Principal components reconstructed by ikPCA

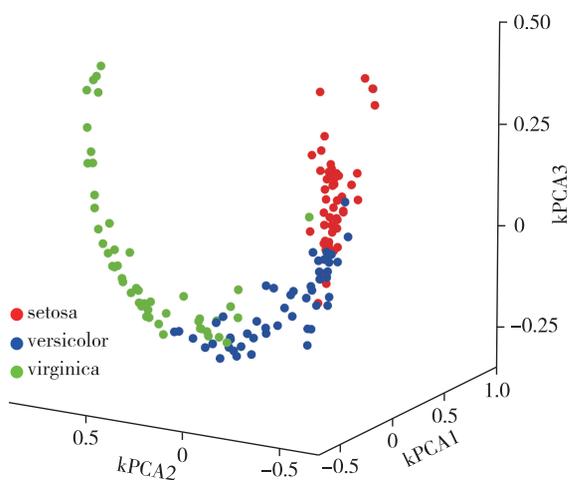


图3 kPCA 重构主成分

Fig. 3 Principal components reconstructed by kPCA

线比其他两种类别明显,而其他两种类别分布重叠部分较多;由图3可知,经kPCA降维后,3种分类的分界线都较为明显,且versicolor和virginica两类的重叠部分相较于PCA降维明显减少,表明引入核方法能提升PCA的降维性能;由图4可知,经ikPCA降维后,3种分类已达到近似线性可分,表明引入傅里叶随机特征近似核函数后,降维性能得到提高,达到了将非线性数据转化为线性数据、降低输入数据复杂度的目的。

## 2 FABAS 算法优化核极限学习机

### 2.1 天牛须搜索算法原理

天牛须搜索算法(Beetle Antennae Search, BAS)是一种仿生类智能优化算法,模仿了天牛在觅食时

的行为<sup>[19]</sup>.天牛主要依靠嗅觉感知食物气味的强弱来辨别食物位置。

将天牛独特的觅食方式用数学描述,其实现步骤如下:

- 1)生成天牛须朝向随机向量  $\mathbf{d}$ :

$$\mathbf{d} = \frac{\text{rand}(k, 1)}{\|\text{rand}(k, 1)\|} \quad (13)$$

- 2)创建天牛须空间坐标:

$$\begin{cases} x_{lm} = x_m - d_m \times \mathbf{d}/2, \\ x_{rm} = x_m + d_m \times \mathbf{d}/2. \end{cases} \quad (14)$$

式中:  $x_{lm}, x_{rm}, x_m$  分别为第  $m$  次迭代时的左须坐标、右须坐标和质心坐标;  $d_m$  为两须间距。

- 3)求解左右须感知的食物气味,即适应度值.其中,适应度函数为  $f(x)$ ,迭代更新天牛位置信息.

$$\begin{cases} x_{m+1} = x_m + \delta_m \times \mathbf{d} \times \text{sign}(f(x_m) - f(x_{lm})), \\ \delta_m = 0.95\delta_{m-1}. \end{cases} \quad (15)$$

式中:  $\delta_m$  为步长因子;  $\text{sign}$  为符号函数。

### 2.2 天牛须搜索算法改进策略

- 1)天牛须多向感知策略

传统BAS算法中的天牛触须只有两个感知方向,天牛个体只会在某一位置的两个方向计算适应度大小和判断下一次迭代移动方向,很可能忽略当前位置附近其他方向上的更优解,进而影响算法寻优特性.事实上,天牛须在觅食过程中,并不只有两个感知方向,因此,利用式(16)生成天牛触须的多向感知随机向量,构建如图5所示的多向感知模型。

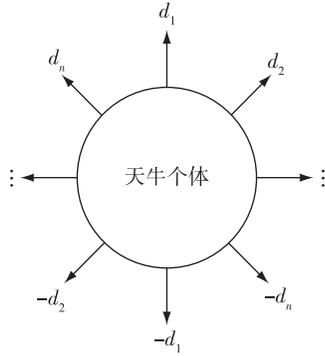


图5 多向感知模型

Fig.5 Orientation perception model

$$\vec{d} = \frac{\text{rand}(k, n)}{\|\text{rand}(k, n)\|}, \quad (16)$$

则改进后的天牛须空间坐标为

$$\begin{cases} x_{1m} = x_m - d_m \times \vec{d}, \\ x_{2m} = x_m + d_m \times \vec{d}. \end{cases} \quad (17)$$

## 2) 萤火虫算法吸引策略

BAS 算法是单体搜索算法,虽然结构简单、运算速度快,但对高维问题的寻优能力有限,且极易陷入局部最优。而群体智能优化算法的优势就是通过多个个体间的信息共享来降低陷入局部最优的风险。萤火虫算法<sup>[20]</sup>可以通过个体间的吸引力逼近最优解。为提高 BAS 算法的寻优能力,本文借鉴群体智能算法的优越性,将单个天牛设定为天牛群,并将天牛的位置更新与萤火虫算法吸引策略结合,提出 FABAS 算法。改进后天牛个体位置更新方式为

$$x_i = x_i + \beta_0 e^{-\mu r^2} (x_j - x_i) \times \vec{d} + \alpha(\text{rand} - 0.5). \quad (18)$$

式中:  $x_i, x_j$  为天牛种群任意两个个体的位置信息,且  $f(x_i) < f(x_j)$ ;  $r$  为两者之间的笛卡尔距离;  $\beta_0$  为个体自身的吸引力,设定值为 1;  $\mu$  为吸引力衰减系数,取值为  $[0, +\infty)$ ;  $\alpha$  为随机扰动系数,取值为  $[0, 1]$ ;  $\text{rand}$  为  $[0, 1]$  之间的均匀随机数。

FABAS 算法实现步骤如下:

步骤 1: 初始化天牛种群,设置种群规模参数、感知系数以及吸引力系数等初始数值。

步骤 2: 对于每一只天牛个体,计算其适应度值,并找出种群中的最优个体,即适应度值最高的个体。

步骤 3: 根据式(16)建立感知模型,运用式(17)生成天牛须的空间位置坐标。比较各个方向上相邻个体的适应度值,判断天牛个体应该前往的方向。

步骤 4: 计算个体之间的距离,利用式(18)进行位置更新。

步骤 5: 检查是否满足终止条件,如果满足则输出得到的最优解。否则,循环执行步骤 2 至步骤 4。

## 2.3 FABAS 算法优化 KELM

极限学习机是一类训练单隐层前向神经网络的机器学习算法,相较于传统的单隐层前馈神经网络(Single-hidden Layer Feedforward Neural network, SLFN)<sup>[21]</sup>,其优势在于随机选取隐含层输入权值  $\omega$  及隐含层阈值  $b$ ,而输出权值则根据广义逆矩阵理论求解。

针对 ELM 模型训练时存在结果波动性和泛化能力弱的问题, Huang 等<sup>[22]</sup>提出了核极限学习机(KELM)。该模型将核函数引入 ELM 模型中,在保证 ELM 模型优势的同时,提升了模型性能和泛化能力。根据 mercer 条件定义得:

$$\begin{cases} \Omega_{\text{ELM}} = \mathbf{H}\mathbf{H}^T, \\ \Omega_{i,j} = h(\mathbf{x}_i) \cdot h(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j). \end{cases} \quad (19)$$

式中:  $\mathbf{x}_i, \mathbf{x}_j$  为输入向量;  $\Omega_{\text{ELM}}$  为核函数矩阵;  $\mathbf{H}\mathbf{H}^T$  为 ELM 模型的随机矩阵。

核函数选用 RBF 核函数:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\{-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2\}. \quad (20)$$

式中,  $\sigma$  为 RBF 核函数宽度系数。

通过引入正则化系数  $C$  和单位矩阵  $\mathbf{I}$ ,可以得到输出权值的最小二乘解为

$$\beta = \mathbf{H}^T \left( \mathbf{H}\mathbf{H}^T + \frac{\mathbf{I}}{C} \right) \mathbf{Y}. \quad (21)$$

式中,  $\mathbf{Y}$  为期望的输出向量。

综上所述, KELM 模型的输出函数可表示为

$$f(\mathbf{x}) = \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_n) \end{bmatrix} \left( K_{\text{ELM}} + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{Y}. \quad (22)$$

由式(22)可以看出, KELM 模型的输出受到正则化参数  $C$  和核函数宽度系数  $\sigma$  的影响。因此,在模型训练过程中,正确选择参数  $(C, \sigma)$  可以提升模型性能和泛化能力。本文采用 FABAS 算法对 KELM 的参数进行寻优,得到最优的模型参数组合。图 6 右侧为 FABAS 算法优化 KELM 参数的流程。

## 3 ikPCA-FABAS-KELM 的组合预测模型

本文预测模型(ikPCA-FABAS-KELM)流程如图 6 所示,其实现步骤如下:

1) 对采集到的历史风电功率和气象数据进行预

处理.

2) 利用 ikPCA 对历史数据进行降维处理, 得到一种在空间线性可分的输入数据, 降低输入复杂度, 同时降低数据波动对预测精度的影响.

3) 初始化 FABAS 算法和 KELM 模型.

4) KELM 模型参数编码到天牛群中, 确定 FABAS 算法的适应度函数, 选择适应度值最小的个体作为最优个体. 本文的 FABAS 算法选择的适应度函数如下:

$$f_{\text{fitness}} = \frac{\sum_{i=1}^m (p_i - \bar{p}_i)^2}{m}. \quad (23)$$

式中:  $p_i$  为训练集实际风电功率;  $\bar{p}_i$  为训练集预测值;  $m$  为训练集样本数.

5) 计算天牛多向感知系统各方向的适应度值.

6) 依据改进的天牛位置更新方式进行寻优, 天牛群向适应度值最小的方向移动, 直到满足 FABAS 算法终止条件.

7) 对适应度值最小的天牛个体进行解码, 以获得 KELM 模型的最佳参数.

8) 输入测试集进行预测, 并进行预测误差分析.

## 4 风电功率预测算例分析

### 4.1 数据来源及评价标准

本文短期风电功率预测实验以某风电场的 1 个

月实测风电输出功率, 采样时间间隔为 15 min, 共 2 880 组数据为数据集. 数据包括采样时刻的风速、风向、温度、气压、湿度 5 项气象数据以及相应的风力发电历史数据. 2 784 组数据进行模型训练, 对其余 96 组数据进行预测测试.

本文采用预测领域通用评价标准: 平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE) 和均方根误差 (RMSE). 即:

$$\sigma_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n | \bar{p}_i - p_i |, \quad (24)$$

$$\sigma_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^n \frac{| \bar{p}_i - p_i |}{p_i} \times 100\%, \quad (25)$$

$$\sigma_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{p}_i - p_i)^2}. \quad (26)$$

式中:  $p_i$  为实际风电输出功率;  $\bar{p}_i$  为预测值.

### 4.2 ikPCA 降维实验

为了验证本文所述的可逆核主成分分析数据降维的优越性, 建立 kPCA 和 ikPCA 模型, 对采集的气象和风电场输出功率数据集进行降维实验验证.

两种方法的降维结果如表 1 和表 2 所示, 表中的  $X_1 \sim X_5$  分别代表风速、风向、温度、气压和湿度. 根据表 1 和表 2, 两种方法第 1 特征向量的方差贡献率分别为 80.12% 和 84.59%, 表明 ikPCA 的第 1 主成分包含更多信息, 且其余 4 个主成分方差贡献率

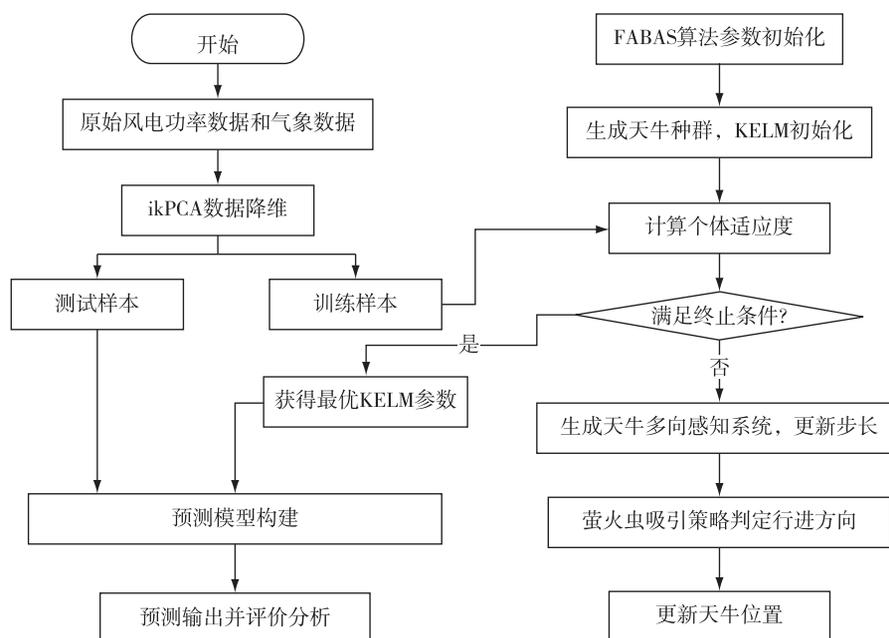


图 6 ikPCA-FABAS-KELM 风电功率预测流程

Fig. 6 Wind power prediction process via the proposed ikPCA-FABAS-KELM

有相同的规律.同时, ikPCA 模型方差贡献率小于 1%的主成分个数比 kPCA 模型要多,表明 ikPCA 模型的重构数据具有更低维度,运算复杂度更低.

表 1 kPCA 特征向量及方差贡献率

Table 1 kPCA eigenvectors and variance contributions

特征向量	kPCA1	kPCA2	kPCA3	kPCA4	kPCA5
$X_1$	0.58	0.52	-0.12	0.76	0.00
$X_2$	0.07	-0.61	0.50	0.55	-0.06
$X_3$	-0.60	-0.15	0.01	0.00	0.70
$X_4$	0.07	0.05	0.01	0.07	0.74
$X_5$	-0.05	0.46	0.85	0.05	0.06
$\alpha/\%$	80.12	21.56	8.60	3.12	1.40
累计 $\alpha/\%$	82.59	85.45	94.26	98.60	100.00

表 2 ikPCA 特征向量及方差贡献率

Table 2 ikPCA eigenvectors and variance contributions

特征向量	ikPCA1	ikPCA2	ikPCA3	ikPCA4	ikPCA5
$X_1$	-0.68	-0.52	0.25	0.06	0.50
$X_2$	0.76	-0.45	0.18	0.13	0.50
$X_3$	-0.06	0.05	-0.84	0.20	0.50
$X_4$	0.02	0.25	0.01	0.01	0.50
$X_5$	-0.05	0.66	0.45	0.44	0.50
$\alpha/\%$	84.59	13.56	2.29	0.01	0
累计 $\alpha/\%$	84.59	94.45	98.26	100.00	100.00

将累计方差贡献率阈值设定为 95%,最终经 kPCA 和 ikPCA 降维后的维度分别为 3 和 2,有效地降低了数据的复杂度,验证了本文 ikPCA 方法在数据降维方面的优越性.

构建 KELM、VMD-KELM、kPCA-KELM 和 ikPCA-KELM 4 种预测模型进行预测实验,以验证数据降维后的数据特征是否完备.结果如图 7 所示,预测误差评价如表 3 所示.

根据预测结果可知,在预测前对原始数据进行处理可以使 KELM 模型运算速率得到提升.值得注意的是,经过 VMD 处理后的数据,对模型预测精度的提升是有限的,而且由于此方法自身的缺陷,导致预测效率降低.经过 kPCA 处理后的数据,虽然预测时效率有所提升,但预测精度不高,表明此方法降维过程中丢失了部分数据特征.而经过 ikPCA 处理后,不仅预测效率得到提升,而且预测误差更小,曲线拟合性也更好.表明 ikPCA 避免了降维过程存在的特征丢失问题,保证了原始数据特征的完备性.而重构主成分的空间线性可分性,在一定程度上降低了风速随机性和波动性的影响,进而提升了预测精度.

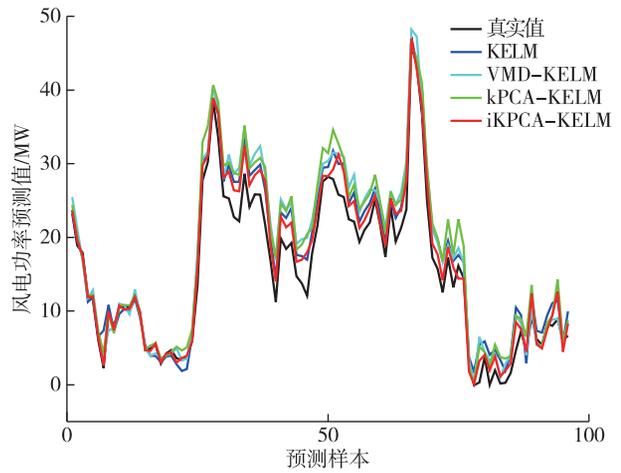


图 7 基于不同降维方法的模型预测结果

Fig. 7 Model prediction results based on different dimensionality reduction methods

表 3 基于不同降维方法的模型预测误差评价

Table 3 Prediction error evaluation of models based on different dimensionality reduction methods

模型	MAE/MW	MAPE	RMSE/MW	时间/s
KELM	2.38	0.038	3.09	184.5
VMD-KELM	2.36	0.038	3.05	223.9
kPCA-KELM	2.61	0.035	3.31	164.6
ikPCA-KELM	1.35	0.024	1.91	150.8

### 4.3 FABAS 算法测试实验

为了评估本文提出的 FABAS 算法的有效性和优越性,选择了如表 4 所示的 6 个 CECBenchmark 基准函数进行寻优测试.同时,为了了解 FABAS 算法与其他智能优化算法的寻优性能,将其与粒子群算法(PSO)、差分进化算法(DE)以及天牛须算法(BAS)进行对比.测试结果如表 5 所示.

表 4 测试基准函数

Table 4 Test benchmark functions

类型	编号	函数名	理论值	寻优范围
单峰函数	F1	Sphere	0	$[-100, 100]^D$
	F2	Rosenbrock	0	$[-30, 30]^D$
多峰函数	F3	Ackley	0	$[-32, 32]^D$
	F4	Griewank	0	$[-600, 600]^D$
偏移函数	F5	Shifted Schwefel1.2	-450	$[-100, 100]^D$
	F6	Shifted Rosenbrock	390	$[-100, 100]^D$

注:  $D$  表示寻优区域的空间维度.

为降低参数设置对测试结果的影响,参与测试的算法种群规模均设置为 40,最大迭代次数 100, PSO 算法学习因子  $c_1 = c_2 = 0.5$ 、惯性权重因子  $w =$

0.9; DE 算法缩放因子  $F$  和交叉概率  $P_c$  分别为:  $F_{\max} = P_{c, \max} = 0.9, F_{\min} = P_{c, \min} = 0.1$ ; BAS 算法和 FABAS 算法天牛须间距  $d_m = 0.001$ , 步长因子初始值  $\delta_0 = 1$ , FABAS 算法的多向感知系数  $n = 8$ , 自身吸引力参数  $\beta_0 = 1$ , 吸引力衰减系数  $\mu = 0.98$ , 扰动系数  $\alpha = 0.5$ .

表 5 PSO、DE、BAS、FABAS 算法测试结果

Table 5 Test results of PSO, DE, BAS, and FABAS

编号	对比项	PSO	DE	BAS	FABAS
F <sub>1</sub>	平均值	0.00E+00	6.42E-32	3.13E+02	0.00E+00
	标准差	0.00E+00	0.00E+00	1.34E+02	0.00E+00
F <sub>2</sub>	平均值	3.99E-03	1.79E+00	1.55E+06	6.97E-01
	标准差	1.26E+00	2.03E+00	1.02E+06	1.01E+00
F <sub>3</sub>	平均值	3.91E-15	4.55E-03	1.50E+01	3.10E-15
	标准差	1.12E-15	3.09E-03	3.19E+00	1.12E-15
F <sub>4</sub>	平均值	9.20E-02	5.03E-0	1.05E+00	8.01E-02
	标准差	1.97E-02	3.09E+00	4.44E-02	1.12E-02
F <sub>5</sub>	平均值	2.99E-28	5.46E-25	1.21E+05	5.58E-29
	标准差	2.99E-28	1.05E-24	1.33E+05	8.76E-29
F <sub>6</sub>	平均值	7.99E-01	7.97E-01	1.27E+10	3.99E-01
	标准差	1.68E+00	1.68E+00	1.23E+09	1.26E+00
+/-		0/6	0/6	0/6	

注:“+/-”表示比 FABAS 算法寻优性能更优劣的测试函数个数。

由表 5 可知,相较于其他 3 种对比算法, FABAS 算法的寻优能力更好.具体表现为,在 6 个测试函数上, FABAS 算法均能以更高的精度收敛到全局最优

解.为更加直观反映 FABAS 算法的优越性能,图 8 给出 4 种算法在求解测试函数时的收敛情况.可清晰地看到,与对比算法相比, FABAS 算法收敛速度明显更快,收敛精度也有所提高,表明对天牛须算法的改进策略有效地解决了寻优能力不足的问题.

在上述测试实验的基础上,利用 4 种寻优算法对 KELM 模型参数进行寻优,以降维实验得到的数据作为模型输入进行预测实验.预测结果如图 9 所示.

由图 9 可看到,通过智能优化算法对 KELM 模型参数寻优后,预测曲线与真实值的拟合程度更好,表明对 KELM 模型参数寻优可避免对模型的盲目训练,提高传统模型的预测能力.应该注意到,4 种算法中, FABAS 算法优化后的模型预测曲线更贴合真实值.

#### 4.4 ikPCA-FABAS-KELM 功率点预测实验

为验证所述模型在短期风电功率预测方面的优越性能,分别构建文献[14]提出的 VMD-ISSA-KELM 模型、文献[15]提出的 EN-SKPCA-LSTMNN 模型和本文提出的 ikPCA-FABAS-KELM 模型进行短期风电功率点预测对比实验.

模型相关参数设置如下:KELM 初始核密度系数  $\sigma = 500$ , 正则化参数  $C = 100$ ; FABAS 算法天牛须间距  $d_m = 0.001$ , 多向感知系数  $n = 8$ , 自身吸引力  $\beta_0 = 1$ , 吸引力衰减系数  $\mu = 0.98$ , 扰动系数  $\alpha = 0.5$ . 对比模型参数设置和原文献一致.

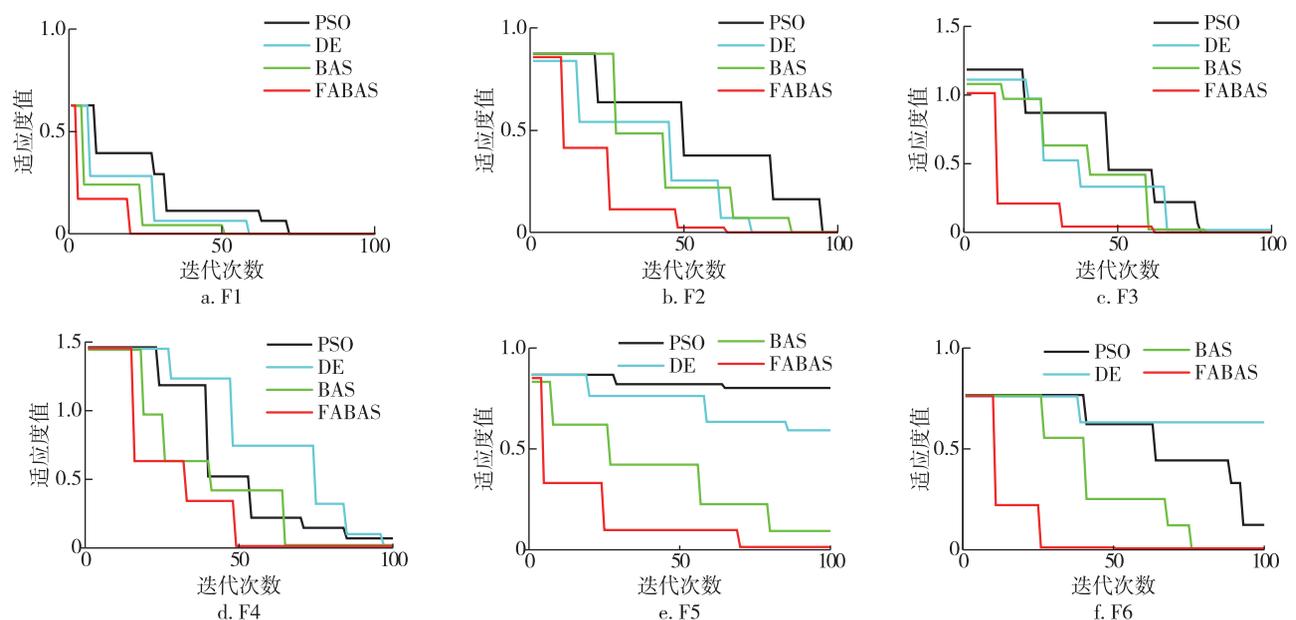


图 8 4 种算法的迭代收敛曲线

Fig. 8 Iterative convergence curves of four algorithms

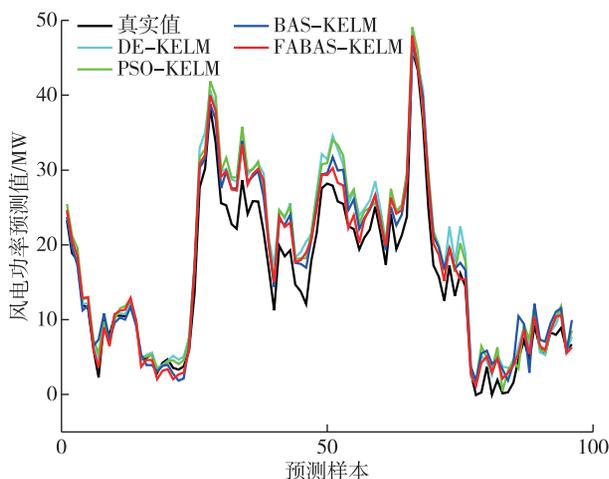


图9 基于不同算法优化 KELM 模型预测结果

Fig. 9 Prediction results of different optimized KELM models

图 10 展示了预测结果,图 11 为模型预测绝对误差.可知,参与对比实验的 3 种模型中,本文的 ikPCA-FABAS-KELM 模型绝对值误差相对较小,且总体上的点预测结果曲线与实际值的拟合度更高,更接近真实值.

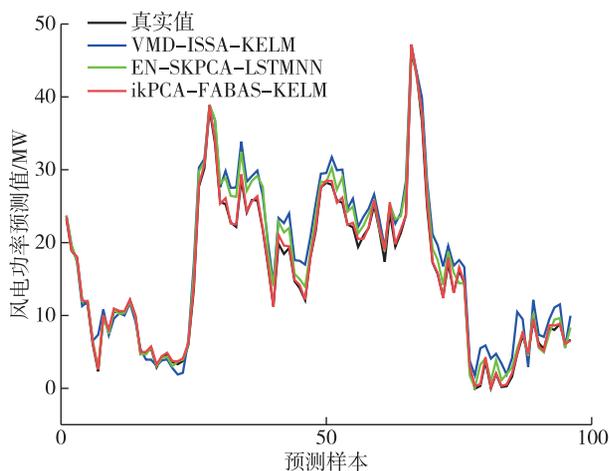


图 10 不同模型的点预测结果曲线

Fig. 10 Point prediction results of different models

表 6 给出了不同模型的点预测误差情况.由表 6 可知:本文的 ikPCA-FABAS-KELM 模型的 MAE 误差为 0.038,与对比的 VMD-ISSA-KELM 模型和 EN-SKPCA-LSTMNN 模型相比降低了约 0.181 和 0.084;MAPE 误差为 0.007,与对比的 VMD-ISSA-KELM 模型和 EN-SKPCA-LSTMNN 模型相比降低了约 0.038 和 0.019;RMSE 误差为 0.052,与对比的 VMD-ISSA-KELM 模型和 EN-SKPCA-LSTMNN 模型相比提升了约 0.212 和 0.115.

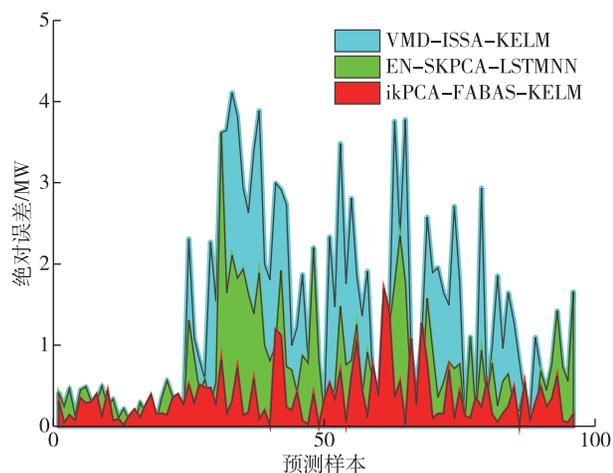


图 11 不同模型点预测绝对误差

Fig. 11 Absolute errors of point prediction for different models

相较于对比模型,本文模型在短期风电功率点预测方面表现出了更好的预测能力,精度更高,并且预测误差符合短期预测标准,表明本文针对传统预测模型不足的改进方法是有效的:一是利于 ikPCA 进行数据处理,降低预测输入的复杂性;二是 FABAS 算法对 KELM 模型参数寻优,可避免人为因素导致的盲目训练问题.

表 6 不同模型的点预测误差评价

Table 6 Evaluation of point prediction errors for different models

预测模型	MAE/MW	MAPE	RMSE/MW
VMD-ISSA-KELM	0.219	0.045	0.264
EN-SKPCA-LSTMNN	0.122	0.026	0.167
ikPCA-FABAS-KELM	0.038	0.007	0.052

#### 4.5 ikPCA-FABAS-KELM 多步预测实验

为验证本文模型在预测精度上的优越性,对上述对比模型进行了多步预测实验,预测步长分别为 2 步和 4 步,预测样本数为 48.

图 12 和图 13 所示分别为 2 步预测和 4 步预测结果.根据预测结果显示,随着预测步长的增加,对比模型的曲线在预测后期出现显著偏离,然而本文提出的模型的预测曲线整体上与真实值曲线保持一致,这说明本文模型的预测精度明显优于对比模型.

对比模型的预测误差评价指标如表 7 所示.可知:ikPCA-FABAS-KELM 模型的 2 步预测的 MAE 误差为 0.032, MAPE 误差为 0.011, RMSE 误差为 0.039,相比 VMD-ISSA-KELM 模型分别降低了 0.169、0.005 和 0.216,相比 EN-SKPCA-LSTMNN 模型分别降低了 0.112、0.007 和 0.164;4 步预测的

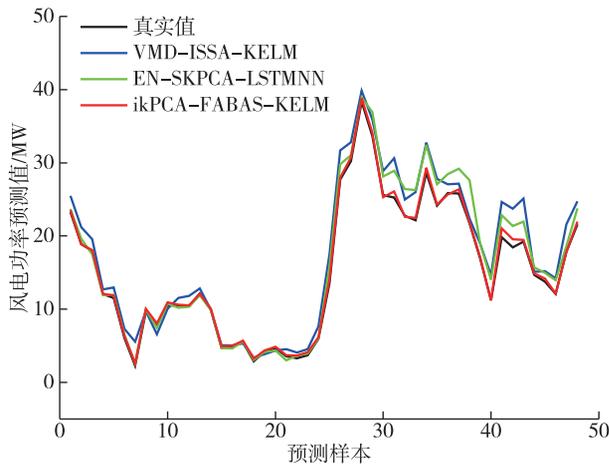


图 12 2步预测结果

Fig. 12 Two-step prediction results

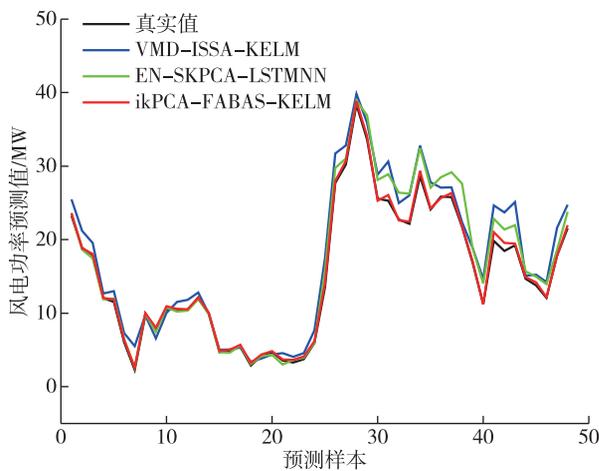


图 13 4步预测结果

Fig. 13 Four-step prediction results

MAE 误差为 0.102, MAPE 误差为 0.013, RMSE 误差为 0.141, 相比 VMD-ISSA-KELM 模型分别降低了 0.194、0.007 和 0.23, 相比 EN-SKPCA-LSTMNN 模型分别降低了 0.084、0.003 和 0.122.

表 7 多步预测误差评价指标

Table 7 Evaluation indicators for multi-step prediction error

预测步长	模型	MAE/MW	MAPE	RMSE/MW
2 步	VMD-ISSA-KELM	0.201	0.016	0.255
	EN-SKPCA-LSTMNN	0.144	0.018	0.203
	ikPCA-FABAS-KELM	0.032	0.011	0.039
4 步	VMD-ISSA-KELM	0.296	0.020	0.371
	EN-SKPCA-LSTMNN	0.186	0.016	0.263
	ikPCA-FABAS-KELM	0.102	0.013	0.141

由上述实验结果可知,增加预测步长后,各模型

预测误差均有不同程度的增大.相比之下,本文的 ikPCA-FABAS-KELM 模型 2 步预测和 4 步预测的误差相差较小,预测稳定性更好.

## 5 结论

为了增强传统数据驱动机器学习模型在短期风电功率预测方面的准确性,本文提出了 ikPCA-FABAS-KELM 预测模型,通过实验验证,得出如下结论:

1) 本文首先提出 ikPCA 模型,在保证完备数据特征的前提下降低了输入复杂性,弥补了传统机器学习模型对复杂问题处理能力的不足;

2) 针对天牛须算法易陷入局部最优问题,本文将萤火虫算法的个体吸引策略和传统天牛须算法结合,提出 FABAS 算法,提升了算法的寻优能力;

3) 采用 FABAS 算法对 KELM 模型的正则化参数和核密度系数进行寻优,避免了对模型的盲目训练造成的泛化能力差的问题;

4) 通过对比仿真验证,本文模型的点预测误差和多步预测能力均符合短期风电功率预测标准,有效地提高了传统模型短期风电功率预测精度.

## 参考文献

### References

- [1] 万灿,宋永华.新能源电力系统概率预测理论与方法及其应用[J].电力系统自动化,2021,45(1):2-16  
WAN Can, SONG Yonghua. Theories, methodologies and applications of probabilistic forecasting for power systems with renewable energy sources[J]. Automation of Electric Power Systems, 2021, 45(1): 2-16
- [2] 姜惠兰,周照清,蔡继朝.风电接入比例对电力系统暂态功角稳定性影响的分析方法[J].电力自动化设备, 2020,40(7):53-67  
JIANG Huilan, ZHOU Zhaoqing, CAI Jizhao. Analysis method of influence of wind power access proportion on transient power angle stability of power system [J]. Electric Power Automation Equipment, 2020, 40(7): 53-67
- [3] 王伟胜,王铮,董存,等.中国短期风电功率预测技术现状与误差分析[J].电力系统自动化,2021,45(1): 17-27  
WANG Weisheng, WANG Zheng, DONG Cun, et al. Status and error analysis of short-term forecasting technology of wind power in China[J]. Automation of Electric Power Systems, 2021, 45(1): 17-27
- [4] Li L L, Chang Y B, Tseng M L, et al. Wind power prediction using a novel model on wavelet decomposition-support vector machines-improved atomic search algorithm [J]. Journal of Cleaner Production, 2020, 270:121817

- [ 5 ] 岳晓宇,彭显刚,林俐.鲸鱼优化支持向量机的短期风电功率预测[J].电力系统及其自动化学报,2020,32(2):146-150  
YUE Xiaoyu, PENG Xiangang, LIN Li. Short-term wind power forecasting based on whales optimization algorithm and support vector machine[J]. Proceedings of the CSU-EPSSA, 2020, 32(2): 146-150
- [ 6 ] Tan L, Han J, Zhang H T. Ultra-short-term wind power prediction by salp swarm algorithm-based optimizing extreme learning machine [J]. IEEE Access, 2020, 8: 44470-44484
- [ 7 ] 龙干,黄媚,方力谦,等.基于改进多元宇宙算法优化 ELM 的短期电力负荷预测[J].电力系统保护与控制,2022,50(19):99-106  
LONG Gan, HUANG Mei, FANG Liqian, et al. Short-term power load forecasting based on an improved multi-verse optimizer algorithm optimized extreme learning machine [J]. Power System Protection and Control, 2022, 50(19): 99-106
- [ 8 ] 史加荣,赵丹梦,王琳华,等.基于 RR-VMD-LSTM 的短期风电功率预测[J].电力系统保护与控制,2021,49(21):63-70  
SHI Jiarong, ZHAO Danmeng, WANG Linhua, et al. Short-term wind power prediction based on RR-VMD-LSTM[J]. Power System Protection and Control, 2021, 49(21): 63-70
- [ 9 ] 赵凌云,刘友波,沈晓东,等.基于 CEEMDAN 和改进时间卷积网络的短期风电功率预测模型[J].电力系统保护与控制,2022,50(1):42-50  
ZHAO Lingyun, LIU Youbo, SHEN Xiaodong, et al. Short-term wind power prediction model based on CEEMDAN and an improved time convolutional network [J]. Power System Protection and Control, 2022, 50(1): 42-50
- [ 10 ] 苗长新,李昊,王霞,等.基于数据驱动和深度学习的超短期风电功率预测[J].电力系统自动化,2021,45(14):22-29  
MIAO Changxin, LI Hao, WANG Xia, et al. Data-driven and deep-learning-based ultra-short-term wind power prediction[J]. Automation of Electric Power Systems, 2021, 45(14): 22-29
- [ 11 ] 钟琳,颜七笙.基于误差修正和 VMD-ICPA-LSSVM 的短期风速预测建模[J].南京信息工程大学学报,2024,16(2):247-260  
ZHONG Lin, YAN Qisheng. Short term wind speed prediction modeling based on error correction and VMD-ICPA-LSSVM[J]. Journal of Nanjing University of Information Science & Technology, 2024, 16(2): 247-260
- [ 12 ] 叶林,赵金龙,路朋,等.考虑气象特征与波动过程关联的短期风电功率组合预测[J].电力系统自动化,2021,45(4):54-62  
YE Lin, ZHAO Jinlong, LU Peng, et al. Combined prediction of short-term wind power considering correlation of meteorological features and fluctuation process [J]. Automation of Electric Power Systems, 2021, 45(4): 54-62
- [ 13 ] 胡帅,向月,沈晓东,等.计及气象因素和风速空间相关性的风电功率预测模型[J].电力系统自动化,2021,45(7):28-36  
HU Shuai, XIANG Yue, SHEN Xiaodong, et al. Wind power prediction model considering meteorological factor and spatial correlation of wind speed [J]. Automation of Electric Power Systems, 2021, 45(7): 28-36
- [ 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].电力系统保护与控制,2022,50(21):138-148  
SHANG Liqun, LI Hongbo, HOU Yadong, et al. Short-term photovoltaic power generation prediction based on VMD-ISSA-KELM [J]. Power System Protection and Control, 2022, 50(21): 138-148
- [ 15 ] 张淑清,杨振宁,姜安琦,等.基于 EN-SKPCA 降维和 FPA 优化 LSTMNN 的短期风电功率预测[J].太阳能学报,2022,43(6):204-211  
ZHANG Shuqing, YANG Zhenning, JIANG Anqi, et al. Short term wind power prediction based on EN-SKPCA dimensionality reduction and FPA optimizing LSTMNN [J]. Acta Energetica Solaris Sinica, 2022, 43(6): 204-211
- [ 16 ] 王浩,王艳,纪志成.基于 SAIGM-KELM 的短期风电功率预测[J].电力系统保护与控制,2020,48(18):78-87  
WANG Hao, WANG Yan, JI Zhicheng. Short-term wind power forecasting based on SAIGM-KELM [J]. Power System Protection and Control, 2020, 48(18): 78-87
- [ 17 ] Wang H, Peng M J, Yu Y, et al. Fault identification and diagnosis based on KPCA and similarity clustering for nuclear power plants [J]. Annals of Nuclear Energy, 2021, 150: 107786
- [ 18 ] Hu Q, Qin A S, Zhang Q H, et al. Fault diagnosis based on weighted extreme learning machine with wavelet packet decomposition and KPCA [J]. IEEE Sensors Journal, 2018, 18(20): 8472-8483
- [ 19 ] Lin M J, Li Q H, Wang F, et al. An improved beetle antennae search algorithm and its application on economic load distribution of power system [J]. IEEE Access, 2020, 8: 99624-99632
- [ 20 ] Li J, Wei X Y, Li B, et al. A survey on firefly algorithms [J]. Neurocomputing, 2022, 500: 662-678
- [ 21 ] 李娜,刘冰,王伟.基于单隐层前馈神经网络的优化算法[J].科学技术与工程,2019,19(1):136-141  
LI Na, LIU Bing, WANG Wei. An optimization algorithm based on single hidden layer feedforward neural networks [J]. Science Technology and Engineering, 2019, 19(1): 136-141
- [ 22 ] Huang G B, Zhou H M, Ding X J, et al. Extreme learning machine for regression and multiclass classification [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2012, 42(2): 513-529

## Short-term wind power prediction based on ikPCA-FABAS-KELM

XU Wu<sup>1</sup> FAN Xinhao<sup>1</sup> SHEN Zhifang<sup>1</sup> LIU Yang<sup>1</sup> LIU Wu<sup>2</sup>

<sup>1</sup> School of Electrical and Information Engineering, Yunnan Minzu University, Kunming 650031, China

<sup>2</sup> Water Supply and Power Supply Company of Xinjiang Dushanzi Petrochemical Company, Karamay 834000, China

**Abstract** A prediction model based on ikPCA-FABAS-KELM is proposed to improve the short-term wind power prediction by traditional data-driven machine learning models. First, the principal component analysis is improved and the reversible kernel Principal Component Analysis (ikPCA) is proposed to reduce the complexity of input data while ensuring data features, with the purpose to advance the model in running speed. Second, the individual attraction strategies for Firefly Algorithm (FA) are used to improve the Beetle Antennae Search (BAS) thus a FABAS algorithm is proposed. Finally, the FABAS algorithm is used to optimize the regularization parameter  $C$  and kernel parameters  $\gamma$  of the Kernel Extreme Learning Machine (KELM), which can reduce the impact of manual parameter setting on blind model training thus improve model prediction accuracy. The simulation results show that the proposed model effectively improves the short-term wind power prediction accuracy.

**Key words** short-term wind power prediction; firefly algorithm (FA); beetle antennae search (BAS); kernel principal component analysis; kernel extreme learning machine (KELM)