



电网技术  
Power System Technology  
ISSN 1000-3673, CN 11-2410/TM

## 《电网技术》网络首发论文

题目: 基于 Stacking 融合的短期风速预测组合模型  
作者: 李永刚, 王月, 刘丰瑞, 吴滨源  
DOI: 10.13335/j.1000-3673.pst.2020.0283  
网络首发日期: 2020-06-24  
引用格式: 李永刚, 王月, 刘丰瑞, 吴滨源. 基于 Stacking 融合的短期风速预测组合模型. 电网技术. <https://doi.org/10.13335/j.1000-3673.pst.2020.0283>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于 Stacking 融合的短期风速预测组合模型

李永刚<sup>1</sup>, 王月<sup>1</sup>, 刘丰瑞<sup>2</sup>, 吴滨源<sup>1</sup>

(1. 新能源电力系统国家重点实验室(华北电力大学), 河北省 保定市 071003;

2. 东北电力大学 电气工程学院, 吉林省 吉林市 132012)

## Combination Model of Short-term Wind Speed Prediction Based on Stacking Fusion

LI Yonggang<sup>1</sup>, WANG Yue<sup>1</sup>, LIU Fengrui<sup>2</sup>, WU Binyuan<sup>1</sup>

(1. State Key Laboratory of New Energy Power System (North China Electric Power University), Baoding 071003, Hebei Province, China; 2. School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, Jilin Province, China)

**ABSTRACT:** Accurate wind speed prediction is of great significance to the stable operation of the power grid when large scale wind power is connected to the grid. In order to improve wind speed prediction accuracy and model generalization ability, a combined wind speed prediction model based on improved Stacking algorithm is proposed. First, A kernel ridge regression model based on different kernel functions is established. Then key parameters are selected using an improved firefly algorithm. Improve the global search ability and convergence speed of the algorithm by introducing adaptive parameters, global search and Levy flight. Finally, the independent models are fused by the Stacking algorithm to enhance the generalization of the model, and the cross-validation is used to further improve the prediction accuracy. The measured results of different wind fields and seasons were selected to simulate the prediction effect of the proposed model, and the prediction accuracy and generalization ability of the proposed model were verified through comparative analysis.

**KEY WORDS:** wind speed prediction; kernel ridge regression; improved firefly algorithm; Stacking algorithm; cross-validation

**摘要:** 大规模风电并网时, 准确的风速预测对电网稳定运行具有重要意义。为提高风速预测精度及预测模型泛化能力, 提出基于改进 Stacking 算法的风速组合预测模型。首先建立基于不同核函数的核岭回归模型; 然后利用改进的萤火虫算法对模型关键参数进行选取, 通过引入自适应参数、全局搜索及 Levy 飞行提高算法的全局搜索能力及收敛速度; 最后通过 Stacking

算法将相互独立的各模型进行融合, 以增强模型泛化性, 并采用交叉验证进一步提高预测精度。选择不同风场、不同季节的实测数据对所提模型的预测效果进行仿真, 通过对比分析验证了所提模型的预测精度和泛化能力。

**关键词:** 风速预测; 核岭回归; 改进萤火虫算法; Stacking 算法; 交叉验证

**DOI:** 10.13335/j.1000-3673.pst.2020.0283

## 0 引言

风场的风速预测可提高含可再生能源并网的电网调度经济性和风电场的运行安全性<sup>[1-2]</sup>, 但风速的波动性、间接性及低能量密度等特点也会降低电力系统运行的可靠性<sup>[3-4]</sup>。因此, 准确的风速预测对于风电并网及电力系统的运行变得越发重要<sup>[5-6]</sup>。

现有的风速及风功率预测方法包括物理方法<sup>[7-8]</sup>、数学方法<sup>[9-10]</sup>、人工智能方法<sup>[11-12]</sup>等。目前, 大多数的国内外学者更关注于基于人工智能的风速预测方法, 其中 BP 神经网络、极限学习机、核岭回归(kernel ridge regression, KRR)等方法应用较多。文献[13]运用天牛须算法优选 BP 神经网络参数从而提出组合风速预测方法, 验证了参数优选的有效性。文献[14]基于互信息约简构建正则化极限学习机组预测模型, 得到了较好的预测效果。KRR 参数设置简单、学习能力强、计算速度快且易收敛, 但传统 KRR 的正则化系数及核参数随机生成, 对预测效果的影响较大。文献[15]对萤火虫优化算法在各个领域中的应用进行了分析, 表明其可有效提高机器学习模型对非线性序列的学习能力, 同时存在易陷入局部最优等问题。研究表明<sup>[16]</sup>, 人工智能及机器学习的预测效果较好, 但仍存在以下不足: 1) 单一预测模型学习能力受随机性限制, 导致泛化能

**基金项目:** 国家自然科学基金项目(51777075)。

Project Supported by National Natural Science Foundation of China (51777075).

力不佳<sup>[17]</sup>。2) 基于加权平均的组合预测算法也不能够体现样本集内单个样本预测误差对整体权重的影响<sup>[18]</sup>。由此, 本文选用 Stacking 算法<sup>[19]</sup>对多个独立的预测模型进行融合, 以提高模型精度和泛化能力。

本文对多个不同的核岭回归模型通过 Stacking 算法进行融合以构建组合风速预测模型。首先, 利用改进的萤火虫算法(improved firefly algorithm, IFA)对各核岭回归模型进行参数优选, 以避免参数随机设置导致的预测精度下降; 然后利用 Stacking 算法融合相互独立的核岭回归模型, 并采用交叉验证方法对基础模型的输入进行划分, 防止过拟合。最后将预测结果与实际风场数据进行对比, 验证所提模型的有效性。

## 1 基于 IFA 的核岭回归预测模型

### 1.1 核岭回归

核岭回归<sup>[20]</sup>基于核技巧将时间序列数据模式非线性地转换为由满足 Mercer 条件的核函数所确定的某些高维特征空间。设给定训练集为 $\{(x_1, y_1), \dots, (x_N, y_N)\}$ , 非线性映射函数 $\varphi(x_i)$ 将原始输入空间转换为高维特征空间。线性回归模型表示为:

$$y_i = \beta_i \cdot \varphi(x_i) + \xi_i, i = 1, 2, \dots, N \quad (1)$$

式中:  $\xi_i$  为松弛变量;  $\beta_i$  组成权重向量  $\beta$ 。

$$\beta = [\beta_1, \beta_2, \dots, \beta_N]^T \quad (2)$$

核岭回归使用正则化最小二乘法通过最小化目标函数来获得权重向量  $\beta$ :

$$\arg \min_{\beta} L_{KRR} = \frac{1}{2} \|\beta\|^2 + \frac{1}{2} C \sum_{i=1}^N \xi_i^2 \quad (3)$$

式中  $C$  为正则化系数。其中

$$\xi_i = y_i - \beta_i \cdot \varphi(x_i), i = 1, 2, \dots, N \quad (4)$$

引入拉格朗日乘数  $\alpha_i$  可得到:

$$L_{KRR} = \frac{1}{2} \|\beta\|^2 + \frac{1}{2} C \sum_{i=1}^N \xi_i^2 + \sum_{i=1}^N \alpha_i (y_i - \beta \cdot \varphi(x_i) - \xi_i) \quad (5)$$

通过求取  $L_{KRR}$  相对于  $\beta$  的导数, 并且令该导数结果为零, 得到输出权重向量  $\beta$  为

$$\beta = (\varphi^T \varphi + \frac{1}{C} I)^{-1} \varphi^T Y \quad (6)$$

式中  $Y = [y_1, y_2, \dots, y_N]^T$ 。

向  $N$  阶矩阵的对角元素添加  $I/C$  的正值产生  $\beta$  的稳定解, 可增强其泛化能力。其中  $I$  为  $N$  阶单位矩阵。

在特征空间中, 由 Mercer 条件<sup>[21]</sup>, 内积可以由核形式表示:

$$K(x_i, x_j) = \varphi(x_i) \varphi(x_j)^T \quad (7)$$

所以, 得到相应的回归方程:

$$y = Y^T \left( K + \frac{1}{C} I \right)^{-1} k \quad (8)$$

式中  $K$  是核心矩阵。

$$K = \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) & \dots & K(x_1, x_N) \\ K(x_2, x_1) & K(x_2, x_2) & \dots & K(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_N, x_1) & K(x_N, x_2) & \dots & K(x_N, x_N) \end{bmatrix}_{N \times N} \quad (9)$$

式中:  $k$  是项为  $k = \langle \varphi(x_i); \varphi(x_j) \rangle$  的向量。

在训练过程中, 满足 Mercer 条件的核函数通过将数据映射到高维特征空间来提高计算能力, 增强数据的线性可分性, 从而提高了模型回归和分类的准确性、稳定性及泛化能力。

内核函数包含全局和局部两类, 其中多项式内核为全局内核函数, 其最显著特征是距离较远的数据点也影响内核函数的值; 高斯内核为局部内核函数, 只有紧邻的数据点会影响内核函数的值; 除此之外, 内核函数还包括具有强大的函数拟合能力以及泛化能力的两种小波核函数, Morlet 小波核函数及墨西哥帽小波核函数。内核函数具体表达式如下:

1) 多项式核函数(polynomial kernel)。

$$K_p(x_i, x_j) = (1 + x_i^T x_j / \sigma^2)^d \quad (10)$$

式中:  $\sigma$  为幅宽;  $d$  为多项式度。

2) 高斯核函数(Gaussian kernel)。

$$K_g(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / (2\sigma^2)) \quad (11)$$

3) 正切双曲核函数(tangent hyperbolic kernel)。

$$K_{TH}(x_i, x_j) = \tanh(b x_i^T x_j + c) \quad (12)$$

式中  $b$  和  $c$  为变量系数。

4) Morlet 小波核函数(Morlet Wavelet kernel)。

$$K_w(x_i, x_j) = \cos\left(\frac{1.675 \|x_i - x_j\|}{e}\right) \exp\left(-\frac{\|x_i - x_j\|^2}{2e^2}\right) \quad (13)$$

5) 墨西哥帽小波核函数(Mexican hat wavelet kernel)。

$$K_{MH}(x_i, x_j) = \left[ 1 - \left( \frac{\|x_i - x_j\|}{f} \right)^2 \right] \exp\left[ -\frac{1}{2} \left( \frac{\|x_i - x_j\|}{f} \right)^2 \right] \quad (14)$$

式中  $f$  为大于 0 的核参数。

在 KRR 回归模型中, 正则化系数  $C$  代表预测模型对误差的宽容度,  $C$  越高, 说明越不能容忍出现误差, 容易过拟合;  $C$  越小, 则容易出现欠拟合。  $C$  过大或过小, 都会导致泛化能力变差。针对不同核函数的核参数选取也不一样, 如多项式核函数中的  $\sigma$ , 设置不当将会影响模型训练与预测的效率及精度。

### 1.2 改进的萤火虫算法

由于不同的预测模型的误差特性不同, 导致各个模型结构和参数不一致且难以调整, 因此本文采用萤火虫算法优化不同核函数的 KRR 模型参数。

传统的萤火虫算法(firefly algorithm, FA)通过萤火虫的光强度来进行萤火虫  $M$  与萤火虫  $J$  间的位置更新, 利用两只萤火虫之间的距离  $r$  来计算他们之间的吸引力  $L(r)$ 。每只萤火虫根据式(15)和式(16)更新位置和吸引力:

$$X_M^T = X_M^{T-1} + L_0 e^{-\phi r^2} (X_M^{T-1} - X_J^{T-1}) + \chi R \quad (15)$$

$$L(r) = L_0 e^{-\phi r^2} \quad (16)$$

式中:  $L_0$  为初始吸引力;  $\phi$  为介质对光的吸收系数;  $\chi$  为初始步长;  $R$  为 $[-0.5, 0.5]$ 之间的随机数;  $T$  为迭代次数。

FA 的核心即每只萤火虫都向着更亮的萤火虫飞行, 将导致群体迅速收敛。当群体全都集中于一个位置时, 算法的搜索能力会迅速下降且无法跳出局部最优。由此, 本文在以下方面对 FA 算法进行改进:

1) 过大的随机化参数将有助于萤火虫探索新的搜索空间但对全局解的收敛无效, 而小的随机化易使算法陷入局部最优。同时为避免算法受制于空间规模且易陷入局部最优, 使每一代的萤火虫能够在大空间中搜索可能的解决方案, 并找到克服局部最优的全局最优解, 本文添加自适应随机化参数及全局最佳概念以提高全局搜索能力。

$$\chi(T+1) = \left( \frac{1}{2T} \right)^{\frac{1}{r}} \chi(T) \quad (17)$$

$$L(r) = (L_{\max} - L_{\min}) e^{-\phi r^2} + L_{\min} \quad (18)$$

式中  $L_{\max}$  和  $L_{\min}$  分别为萤火虫中最大及最小吸引力。

2) 引入 Levy 飞行随机切换长步长与短步长使算法跳出局部最优, 并提高搜索速度, 使模型尽快得到最优解:

$$f_{\text{Levy}}(\kappa) \sim u = T^{-\kappa}, \quad 1 < \kappa < 3 \quad (19)$$

式中  $u$  为服从正态分布的随机步长。

由此式(15)改为

$$X_J^{T'} = X_J^T + o \otimes f_{\text{Levy}}(\kappa) \quad (20)$$

式中:  $o$  为步长缩放因子;  $\otimes$  为卷积运算。

本文基本参数设置如下: 最大迭代次数  $T_{\max}$  为 50,  $L_0$  为 1,  $L_{\max}$ 、 $L_{\min}$  分别为 2.0 和 0.2,  $\phi$  为 1,  $\chi$  为 0.97, 萤火虫种群数量为 20。采用 IFA 优化各个 KRR 模型的参数, 并将多个 KRR 模型通过 Stacking 进行融合。

## 2 基于 Stacking 融合的风速预测模型

已有研究表明<sup>[22]</sup>, 单一模型的预测结果精确度有限, 因此采用 Stacking 模型融合多个预测模型。Stacking 模型具有较强的异构数据处理能力及泛化能力, 可同时降低机器学习模型的偏差及方差。

### 2.1 Stacking 融合模型

本文利用 Stacking<sup>[23]</sup>模型融合多个不同的 KRR 模型, 以提升风速预测精度。其融合系统设计为 2 层结构, 在强化学习效果的同时又避免了预测模型的冗余复杂, 保证预测精度的同时缩短了运算时间。模型第一层即系统的基础模型层, 由  $n$  个不同的基础模型集成; 第二层的元模型与第一层的基础模型进行融合,

共同构成 Stacking 模型。

当模型输入为  $X_i$ , 第一层第  $h$  个基础模型为  $F_h$ , 第二层预测模型为  $F$ 。则第一层第  $h$  个基础模型的输出为  $F_h(X_i)$ , 并将其作为第二层预测模型的输入, 最终预测结果的  $z_i$ , 如式(21)所示。针对融合过程的 2 个阶段进行分析。

$$z_i = F(F_1(X_i), \dots, F_h(X_i), \dots, F_n(X_i)) \quad (21)$$

第 1 阶段, 指定训练集及预测集以实现基础模型层的训练及预测, 最终得到针对元模型第 1- $n$  组的预测值。

第 2 阶段, 根据上一阶段得到的  $n$  组基础模型层输出的预测值和预测集中的数据, 分别构建元模型的训练集及测试集, 用于第二层预测模型的预测, 获得最终预测值。

传统的 Stacking 模型中, 相同的训练数据集由基础模型拟合, 以便为元模型准备输入, 这可能导致信息泄漏和过度拟合。因此, 本文对传统的 Stacking 算法进行改进, 通过使用交叉验证方法<sup>[24]</sup>以平衡过拟合和预测误差之间的关系。具体模型如图 1 所示。

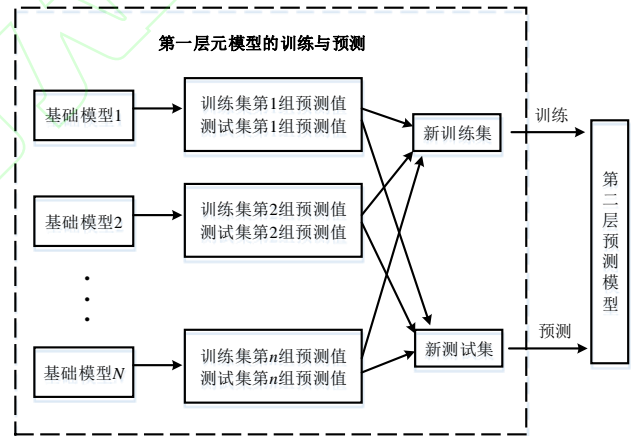


图 1 Stacking 模型融合过程  
Fig.1 Stacking model fusion process

基础模型的数量与融合效果强相关, 模型数量较少达不到各模型的互补融合, 数量过多则会造成冗余, 增加系统参数复杂度及预测时间。一般来说, 基础模型数量为 3~5 个为宜, 本文经过大量仿真及人工经验, 选用含有 4 个模型时 Stacking 模型, 此时的融合效果最佳。因此, 本文将 4 个不同 KRR 模型作为 Stacking 模型融合系统第一层的基础模型, 将 GKRR 模型作为第二层预测模型, 使 Stacking 模型的预测性能最强化。

### 2.2 风速组合预测模型流程及评价指标

为验证所提模型的有效性, 选取以下两种评价指标对预测结果进行评价。

$$\varepsilon_R = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - z_i)^2} \quad (22)$$



$$\varepsilon_M = \frac{1}{N} \sum_{t=1}^N \frac{|x_t - z_t|}{|x_t|} \times 100\% \quad (23)$$

式中： $z_t$ 为 $t$ 时刻的预测值； $x_t$ 为 $t$ 时刻的实际值； $N$ 为预测集中时间点总数。

绘制本文提出的风速组合预测模型具体流程如图2所示。

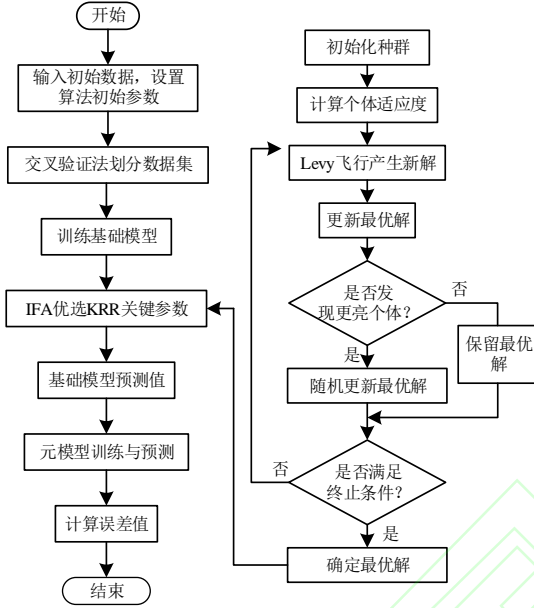
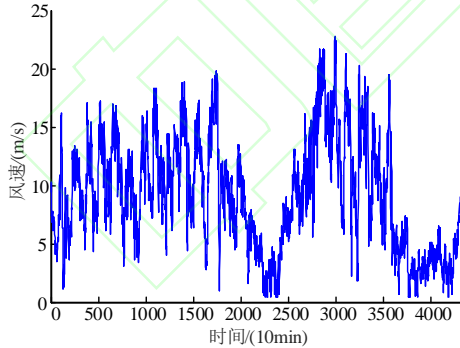
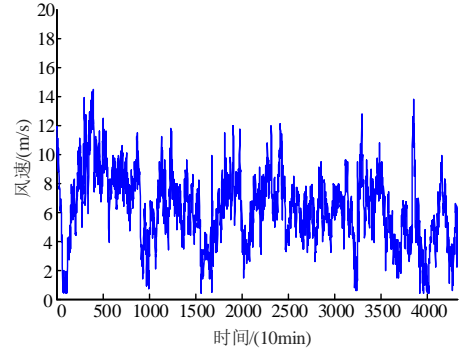


图2 组合模型流程图

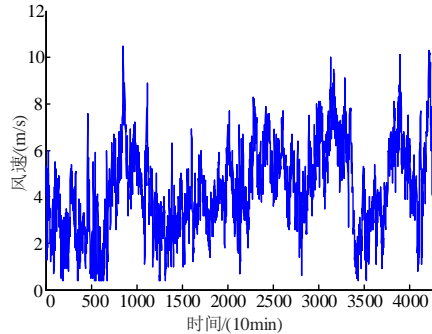
Fig.2 Combination model flow chart



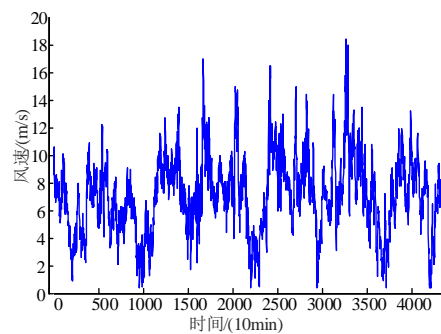
(a)3月风速



(b)6月风速



(c)9月风速



(d)12月风速

图3 南方风场不同季节典型月风速序列

Fig.3 Sequence of typical monthly wind speed in different seasons of southern wind field

### 3 仿真分析

选取南方电网实际风场风速数据开展研究, 取该风场2012年各季度典型月原始风速序列进行仿真分析, 并对原始数据进行预处理, 采用相邻数据互补法替换缺失及异常风速值。原始数据的采样间隔为10 min, 共计4320个点, 取每个月前20天数据为训练集, 21—25日数据为测试集, 26—30日为检验集(用以检验模型参数设定是否合适)。

#### 3.1 不同方法预测结果分析

2012年各季度典型月风速数据如图3所示, 可知该风场在不同季节风速波动变化差异较大。其中春季波动最为剧烈, 不同时刻风速差较大。而冬季风速波动情况比较规律, 为周期性增强或减弱。夏秋两季风速变化规律相似且波动较平缓。

为检验和说明 IFA-KRR 的预测性能, 选用极限学习机(extreme learning machine, ELM)和BP神经网络两种人工智能算法与之进行对比, 对各季度典型月风速进行单步滚动的24h短期预测。

其中, ELM 采用径向基函数(radical basis function, RBF)作为核函数。BP 使用单隐层网络结构, 激活函数为 sigmoid 函数。本文的训练及测试均在 Matlab 2016b 环境下运行, 采用 Intel(R)Core(TM) i5-8500, CPU@3.00GHz, RAM 8.00 GB 的微机平台。

由图4可以看出,类似3月21日10~40及12月21日0~60之间的风速波动较平缓时段,各模型均能得到较好的预测效果。但风速波动较剧烈时,如图4(c)中的120~144之间,BP神经网络由于并没有对参数进行优化导致对非线性序列的学习能力较弱且容易过拟

合,导致预测结果与实际偏差较大。ELM能够很好预测出风速序列的变化趋势,但在预测值仍与实际值仍存在一定差距。而本文方法通过IFA进行参数优化后,表现出优异的预测性能,对风速的随机波动性具有较好的拟合效果。具体的预测误差见表1。

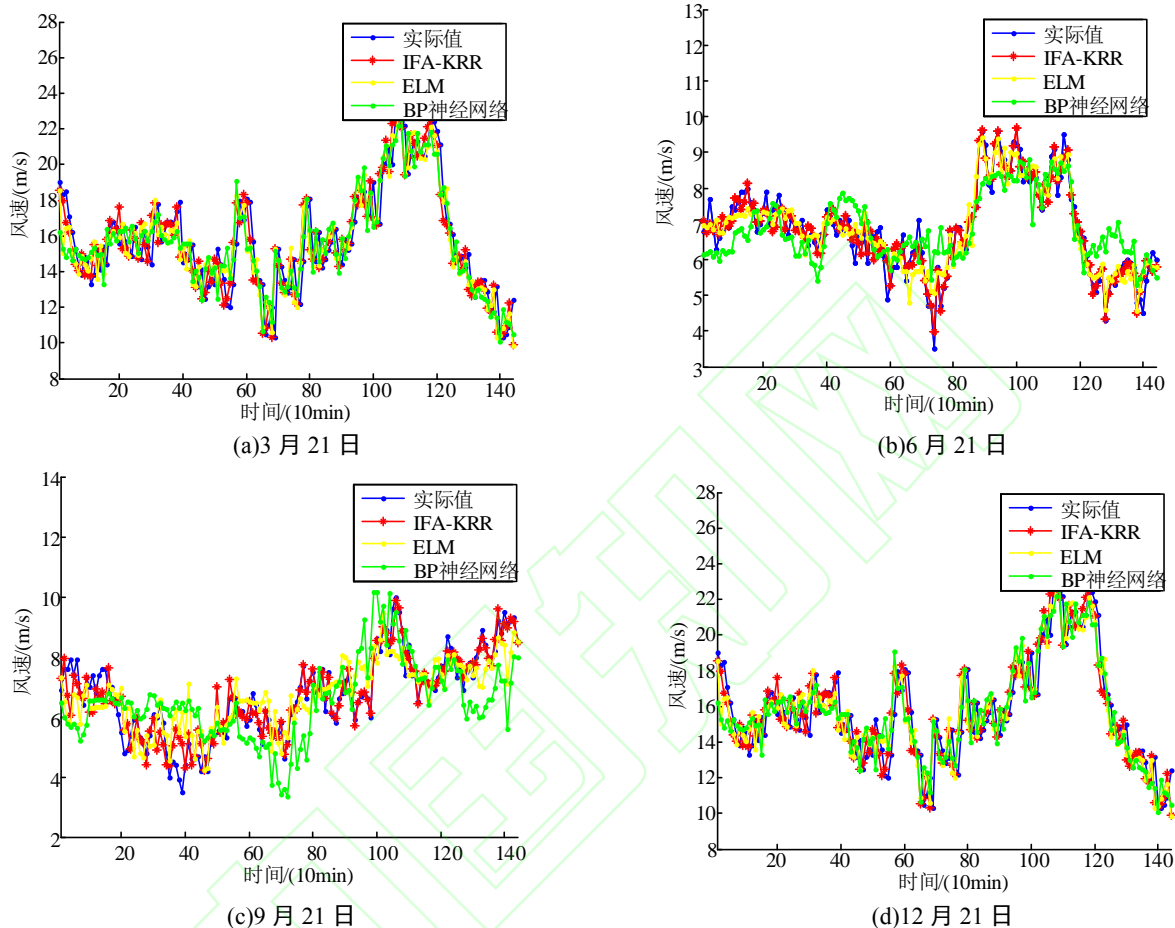


图4 不同方法短期预测对比结果

Fig.4 Contrast of short-term prediction results of different methods

表1 不同方法短期预测误差

Tab.1 Short-term prediction error results of different methods

日期	模型	$\varepsilon_R/\%$	$\varepsilon_M/(\text{m/s})$
3月21日	IFA-KRR	0.315	8.685
	ELM	0.507	11.995
	BP	0.586	13.79
6月21日	IFA-KRR	0.341	9.076
	ELM	0.535	11.933
	BP	0.683	14.675
9月21日	IFA-KRR	0.384	9.317
	ELM	0.544	12.075
	BP	0.713	15.672
12月21日	IFA-KRR	0.329	8.476
	ELM	0.512	11.524
	BP	0.597	12.021

由表1可知,与BP神经网络相比,ELM模型在3月21日的平均绝对误差为12%左右,具有更高的预测精度和预测效果。而IFA-KRR通过改进的萤火虫算法对KRR参数进行了快速的全局搜索,在一定程度上提高了KRR的预测性能,有效避免了由参数设置不合理而引发过拟合等问题。

### 3.2 对比模型预测结果分析

首先对改进的FA和传统萤火虫算法进行对比,其

误差曲线如图5所示。

由图5可知本文改进的FA明显优于原始萤火虫算法。FA虽有较快的搜索速度但易陷入局部最优,并没有进行全局搜索。而引入全局最佳概念及自适应参数可有效提高模型的寻优能力,Levy飞行也可提高搜索速度,使模型尽快得到最优解,减小融合模型的时间复杂度,有效提高了算法的收敛性能。

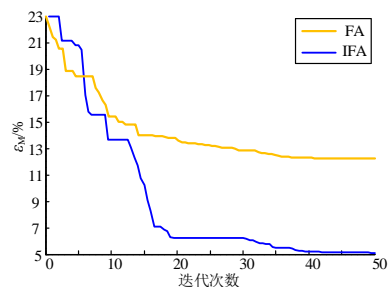


图5 误差曲线对比

Fig.5 Contrast of error curves

为验证本文所提Stacking融合模型的准确性,选择不同组合模型对25日风速进行滚动预测。其中,模

型 1 为 Sta-IFA-KRR，模型 2 为 Sta-KRR，模型 3 为基于高斯内核的 KRR，结果见图 6。

由图 6 中可以看出，模型 1 和模型 2 在各季节的预测效果均优于模型 3，证明 Stacking 模型可有效提高组合预测模型的泛化性。其中，对于冬季（12 月）风速为规律性的大幅波动时，单一模型的 KRR 预测性能明显弱于通过 Stacking 模型融合

明通过 Stacking 模型融合多个相互独立的预测模型，具有更强的非线性表述能力，可有效降低泛化误差。而模型 2 与模型 1 相比，由于没有对参数进行优选，虽能对风速变化趋势进行很好的预测，但在细节上（如 9 月 25 日 110~130 及 3 月 25 日 0~10 之间）容易出现过拟合现象导致预测结果与实测数据偏差较大。

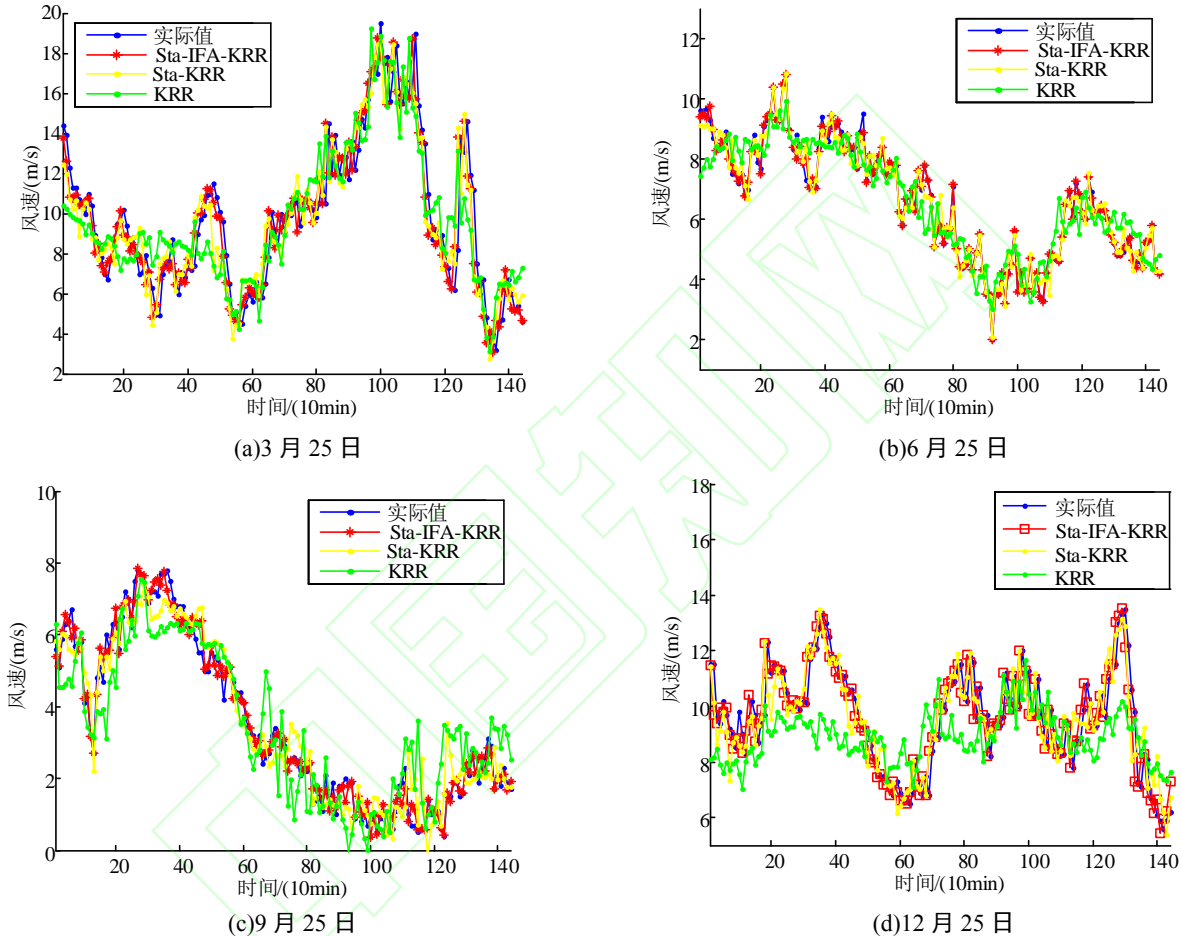


图 6 不同组合模型短期预测对比结果

Fig.6 Comparison of short-term prediction results of different combination models

对各模型不同季度风速预测进行误差分析，表 2 结果表明，模型 1 的误差均小于对比模型，这是因为模型 2 并没有对 KRR 参数进行优选，使得预测模型针对不同的风速波动情况的泛化能力有所下降。以 3 月份为例，模型 1 的平均绝对百分比误差相对于模型 2 减小了 34%，进一步验证了 IFA 进行参数优选的有效性。而模型 2 相对于模型 3 的误差较小是由于方法 3 并未融合多个预测模型结果，单一模型对非线性序列的学习能力有限，导致预测精度下降。

表 2 不同组合模型短期预测误差

Tab.2 Short-term prediction error of different combination models

日期	模型	$\varepsilon_R/(m/s)$	$\varepsilon_M/\%$
3 月 25 日	模型 1	0.238	6.474
	模型 2	0.437	9.834
	模型 3	0.576	14.93

6 月 25 日	模型 1	0.165	5.941
	模型 2	0.293	8.379
	模型 3	0.507	10.607
9 月 25 日	模型 1	0.194	6.249
	模型 2	0.401	9.318
	模型 3	0.525	11.108
12 月 25 日	模型 1	0.153	5.433
	模型 2	0.384	9.287
	模型 3	0.608	15.018

3.3 泛化验证分析

为避免所选地区经纬度及气候特殊性的影响，结合美国风能资源中心(39.91°N, 105.23°W)所提供的数据，进一步验证本文所提模型的适用性。以 2018 年 11 月历史风速数据为例，数据采样间隔为 10 min，风速波动时间序列见图 7。

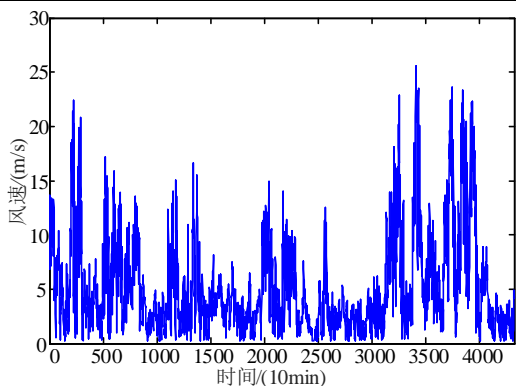


图7 风速波动时序图

Fig.7 Schematic diagram of wind speed

该风场受地理位置和气候环境的影响,在不同季度的风速波动特征与我国西南地区存在较大差异。根据当地实测数据选用不同组合模型对11月21日风速进行预测,结果如图8及表3所示。

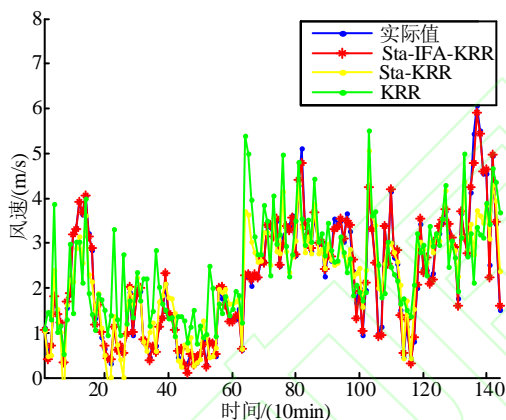


图8 泛化验证对比结果

Fig.8 Generalization verification comparison results

表3 不同模型泛化验证误差

Tab.3 Generalization verification error of different models

模型	$\varepsilon_R/(m/s)$	$\varepsilon_M/\%$
模型1	0.148	5.375
模型2	0.403	9.318
模型3	0.517	11.015

由图8可知,模型1可以很好的学习验证风场的风速变化,而模型3的预测值与实际风速具有较大误差,检验了Stacking模型的灵活性及泛化性。进一步对该预测结果进行误差分析,模型1的 $\varepsilon_R$ 为0.148m/s,  $\varepsilon_M$ 为5.375%,表明本文所提模型能够较为准确地反映风速变化趋势,各项预测指标均符合标准要求<sup>[25]</sup>,进而验证了该方法具有较强的适用性和泛化能力。

## 4 结论

基于核岭回归,本文构建Stacking融合预测模型,结合实际风场数据进行短期风速预测,得出以下结论:

1) Stacking算法充分考虑了基于不同核函数的核

岭回归预测效果差异性,对多个独立预测模型进行融合以获得更好的泛化能力,并采用交叉验证模型进一步提高模型预测精度。

2) 对原始萤火虫算法进行改进有效提高搜索速度,使模型尽快得到全局最优解,提高了算法的收敛性能。基于改进的萤火虫优化算法优选核岭回归模型参数,可以避免参数设置不当而引起的过拟合问题。

3) 通过构建组合风速预测模型对不同风场各季节典型月进行预测,结果表明所提模型具有较好的准确性与适用性。

## 参考文献

- [1] 张丽英,叶廷路,辛耀中,等.大规模风电接入电网的相关问题及措施[J].中国电机工程学报,2010,30(25):1-9.  
Zhang Liying, Ye Tinglu, Xin Yaozhong, et al. Problems and measures of power grid accommodating large scale wind power[J]. Journal of Chinese Electrical Engineering Science, 2010, 30(25): 1-9(in Chinese).
- [2] 殷豪,曾云,孟安波,等.基于奇异谱分析-模糊信息粒化和极限学习机的风速多步区间预测[J].电网技术,2018,42(5):1467-1474.  
Yin Hao, Zeng Yun, Meng Anbo, et al. Wind speed multi-step interval prediction based on singular spectrum analysis fuzzy information granulation and extreme learning machine[J]. Power System Technology, 2018, 42(5): 1467-1474(in Chinese).
- [3] 杨楠,黄禹,叶迪,等.基于NACEMD和改进非参数核密度估计的风功率波动性概率分布研究[J].电网技术,2019,43(3):910-917.  
Yang Nan, Huang Yu, Ye Di, et al. Study on Probability Distribution of Wind Power Fluctuation Based on NACEMD and Improved Nonparametric Kernel Density Estimation[J]. Power System Technology, 2019, 43(3): 910-917(in Chinese).
- [4] 熊音笛,刘开培,秦亮,等.基于时序数据动态天气划分的短期风电功率预测方法[J].电网技术,2019,43(9):3353-3359.  
Xiong Yindi, Liu Kaipei, Qin Liang, et al. Short-term wind power prediction method based on dynamic wind power weather division of time sequence data[J]. Power System Technology, 2019, 43(9): 3353-3359(in Chinese).
- [5] 张妍,韩璞,王东风,等.基于变分模式分解和LSSVM的风电场短期风速预测[J].太阳能学报,2018,39(1):194-202.  
Zhang Yan, Han Yan, Wang Dongfeng, et al. Short-term wind speed prediction of wind farms based on variational mode decomposition and LSSVM[J]. Journal of Solar Energy, 2018, 39(1): 194-202.
- [6] Deng Ying, Wang Bofu, Lu Zhiming. A hybrid model based on data preprocessing strategy and error correction system for wind speed forecasting[J]. Energy Conversion and Management, 2020, 212.
- [7] 牛东晓,纪会争.风电功率物理预测模型引入误差量化分析方法[J].电力系统自动化,2020,44(8):57-67.  
Niu Dongxiao, Ji Huizheng. Wind power physical prediction model introduces error quantitative analysis method[J]. Automation of Electric Power Systems, 2020, 44(8): 57-67.
- [8] 冯双磊,王伟胜,刘纯,等.基于物理原理的风电场短期风速预测研究[J].太阳能学报,2011,32(5):611-616.  
Feng Shuanglei, Wang Weisheng, Liu Chun, et al. Short-term wind speed prediction based on physical principle[J]. Journal of Solar Energy, 2011, 32(5): 611-616.
- [9] 潘超,秦本双.基于实测数据的大型风电场功率模型研究[J].电网技术,2017,41(8):2625-2631.  
Pan Chao, Qin Benshuang. Research on power model of large-scale wind farm based on measured data[J]. Power System Technology, 2017, 41(8): 2625-2631.



- 2625-2631.
- [10] Yang Mao, Chen Xinxin, Huang Binyang, et al. Ultra-short-term multi-step wind power prediction based on fractal scaling factor transformation[J]. Journal of Renewable and Sustainable Energy, DOI: 10.1063/1.5042795.
- [11] 向玲, 邓泽奇, 赵玥. 基于 LPF-VMD 和 KELM 的风速多步预测模型[J]. 电网技术, 2019, 43(12): 4461-4467.  
Xiang Ling, Deng Zeqi, Zhao Yue. Multi-step wind speed prediction model based on LPF-VMD and KELM [J]. Power System Technology, 2019, 43(12): 4461-4467.
- [12] Lu Hongfang, Ma Xin, Kun Huang, et al. Prediction of offshore wind farm power using a novel two-stage model combining kernel-based nonlinear extension of the Arps decline model with a multi-objective grey wolf optimizer[Z]. Renewable and Sustainable Energy Reviews, 2020, 127.
- [13] 李大中, 李昉, 张克延. 一种利用混合算法选择变量的天牛须优化神经网络风速预测方法[J]. 电力科学与工程, 2019, 35(12): 28-33.  
Li Dazhong, Li Yan, Zhang Keyan. A method for predicting the wind speed of the beetles optimization neural network using hybrid algorithm to select variables[J]. Electric Power Science and Engineering, 2019, 35 (12): 28-33.
- [14] 潘超, 秦本双, 何瑶, 等. 基于优化 FCM 聚类的 RELM 风速预测[J]. 电网技术, 2018, 42(3): 842-848.  
Pan Chao, Qin Benshuang, He Yao, et al. RELM Wind Speed Prediction Based on Optimized FCM Clustering[J]. Power System Technology, 2018, 42(3): 842-848.
- [15] Fister I, Yang X S, Brest J. A comprehensive review of firefly algorithms[J]. Swarm and Evolutionary Computation, 2013, 13: 34-46.
- [16] Katuwal R, Suganthan P N. An ensemble of kernel ridge regression for multi-class classification[J]. Procedia Comput. Sci. DOI: 10.1016/j.procs.2017.05.109.
- [17] 胡梦月, 胡志坚, 仇梦林, 等. 基于改进 AdaBoost.RT 和 KELM 的风功率预测方法研究[J]. 电网技术, 2017, 42(2): 536-542.  
Hu Mengyue, Hu Zhijian, Wu Menglin, et al. Research on wind power prediction methods based on improved AdaBoost. RT and KELM [J]. Power System Technology, 2017, 42(2): 536-542.
- [18] 刘克文, 蒲天骄, 周海明, 等. 风电日前发电功率的集成学习预测模型[J]. 中国电机工程学报, 2013, 33(34): 130-135.  
Liu Kewen, Pu Tianjiao, Zhou Haiming, et al. A short term wind Power forecasting model based on combination algorithms[J]. Proceedings of the CSEE, 2013, 33(34): 130-135.
- [19] 刘波, 秦川, 鞠平, 等. 基于 XGBoost 与 Stacking 模型融合的短期母线负荷预测[J]. 电力自动化设备, 2020, 40(3): 147-153.  
Liu Bo, Qin Chuan, Ju Ping, et al. Short-term Bus Load Forecasting Based on Fusion of XGBoost and Stacking Model[J]. Electric Power Automation Equipment, 2020, 40 (3): 147-153.
- [20] Li Y Z, Luo D Y, Liu S Q. Face recognition using neighborhood preserving maximal margin analysis of kernel ridge regression[J]. Moshir Shiebi Yu Rengong Zhineng/pattern Recognition & Artificial Intelligence, 2010, 23(1): 23-28.
- [21] Naik J, Bisoi R, Dash P K. Prediction interval forecasting of wind speed and wind power using modes decomposition based low rank multi-kernel ridge regression[J]. Renewable energy, 129(2018): 357-383.
- [22] 史佳琪, 张建华. 基于多模型融合 Stacking 集成学习方式的负荷预测方法[J]. 中国电机工程学报, 2019, 39(14): 4032-4042.  
Shi Jiaqi, Zhang Jianhua. Load Forecasting Based on Multi-model by Stacking Ensemble Learning [J]. Chinese Journal of Electrical Engineering, 2019, 39(14):4032-4042.
- [23] 张雨金, 周杭霞. Stacking-SVM 的短期光伏发电功率预测[J]. 中国计量大学学报, 2018, 29(2): 121-127.  
Zhang Yujin, Zhou Hangxia. Stacking-SVM short-term photovoltaic power prediction [J]. Journal of China Jiliang University, 2018, 29(2): 121-127.
- [24] Zhang Y, Yang Y. Cross-validation for selecting a model selection procedure[J]. Journal of Econometrics, 2015, 187(1): 95-112.
- [25] 中华人民共和国国家能源局. 风电场功率预测预报管理暂行办法[Z]. 中华人民共和国国家能源局, 2011.

收稿日期: 2020-04-07.

作者简介:

李永刚(1967), 男, 教授, 博士生导师, 通信作者, 研究方向为电气设备诊断、新能源电力系统, E-mail: lygzxm0@163.com;

王月(1995), 女, 硕士研究生, 研究方向为新能源电力系统, E-mail: 1071859080@qq.com;

刘丰瑞(1998), 男, 硕士研究生, 研究方向为综合能源系统, E-mail: lfr0007@sina.com.



李永刚

(责任编辑 王金芝)