

DOI: 10.19783/j.cnki.pspc.191347

# 基于 SAIGM-KELM 的短期风电功率预测

王浩, 王艳, 纪志成

(江南大学物联网技术应用教育部工程研究中心, 江苏 无锡 214122)

**摘要:** 针对时序下风电功率的随机性和波动性问题, 提出一种基于自适应智能灰色系统(SAIGM)和遗传算法优化核极限学习机(GA-KELM)的混合风电功率预测模型。首先, 以灰色关联性分析不同季度下风向量与数值天气预报(NWP)对风电功率的影响为基础, 采用自适应智能灰色系统预测风速, 并将预测的风速与相连时序下的风向和NWP有效整合作为预测样本。其次, 利用遗传算法优化核极限学习机搭建风电功率预测模型, 并将实际风向量与NWP有效整合作为预测模型的训练样本。最后, 利用优化后的预测模型实现不同季节的风电功率预测。实验表明混合预测模型可实现对风电功率的短期预测, 预测结果具有准确性和可靠性。

**关键词:** 风电功率; 灰色关联性; 自适应智能灰色系统; 遗传算法; 核极限学习机

## Short-term wind power forecasting based on SAIGM-KELM

WANG Hao, WANG Yan, JI Zhicheng

(Engineering Research Center of Internet of Things Technology Applications, Ministry of Education,  
Jiangnan University, Wuxi 214122, China)

**Abstract:** Given the problem caused by the randomness and volatility of wind power under time series, a hybrid wind power forecasting model based on a Self-Adaptive Intelligence Grey Predictive Model with Alterable Structure (SAIGM) and Genetic Algorithm Optimized Kernel Extreme Learning Machine (GA-KELM) is proposed. First, the influence of wind vector and Numerical Weather Prediction (NWP) on wind power in different seasons is analyzed by grey correlation, and the wind speed is predicted by an adaptive intelligent grey system. The predicted wind speed is effectively integrated with the actual wind vector and NWP in the adjacent time series as prediction samples. Secondly, the optimized kernel extreme learning machine based on a genetic algorithm is used to build the wind power prediction model, and the actual wind vector with NWP are also effectively integrated as training samples of the forecasting model. Finally, the optimized prediction model is used to achieve wind power forecasting in different seasons. Experiments demonstrate that the hybrid forecasting model can realize short-term wind power forecasting and the results are accurate and reliable.

This work is supported by National Natural Science Foundation of China (No. 61973138).

**Key words:** wind power; grey correlation analysis; SAIGM; genetic algorithm; KELM

## 0 引言

随着工业化的发展, 化石燃料的污染程度高, 对环境产生不利影响的缺点日益凸显, 导致对可再生能源如风能、水能、太阳能等的需求量急剧增加。风能作为自然界中因空气流动形成的清洁、天然、循环可再生的能源, 在电力系统中增长速度最快。但风速的随机性给风电场风能的预测带来困难和挑战, 对电网的计划和调度造成不利影响。因此, 准

确预测风电功率对电力系统的运行极其重要<sup>[1-2]</sup>。

根据预测的时间尺度以及相应的控制策略, 可将风电功率预测分为超短期、短期、中期以及长期预测等<sup>[3]</sup>; 风电功率预测的方法可分为统计方法、物理方法、组合预测方法、空间相关性预测方法等<sup>[4]</sup>。统计方法通常根据对时间尺度下的历史数据进行统计分析, 找到输入因素与风电功率之间的关联后进行预测, 在风电功率的超短期与短期预测方面应用广泛。传统的线性回归统计方法如灰色模型<sup>[5]</sup>、卡尔曼滤波<sup>[6]</sup>、自回归积分滑动模型<sup>[7]</sup>等已经无法满足风电功率的非线性特征。因此, 国内外学

者提出了更先进的能找出输入与输出非线性关系的统计方法, 如人工神经网络<sup>[8-9]</sup>、小波分析<sup>[10-11]</sup>、深度学习<sup>[12]</sup>、支持向量机<sup>[13-14]</sup>等, 同时研究采用多种预测模型的优点以建立具有更好泛化性和精确性的组合预测模型<sup>[15-16]</sup>, 使得在时间序列下对风电功率的超短期与短期预测具有更好的预测效果。

极限学习机(Extreme Learning Machine, ELM)是文献[17]提出的一种新型单隐层前馈神经网络(Single-hidden Layer Feedforward Neural networks, SLFNs)的学习算法, 与传统神经网络相比, ELM 的特点在于输入权值与隐藏层的偏置随机产生且不需要改变, 隐藏层的节点数目可选择调整, 训练中只需确定输出权重, 使得 ELM 具有快速的学习能力和良好的泛化性能。文献[18]提出将极限学习机(ELM)与 Bootstrap 方法相结合用于风电概率的间隔预测, 利用不同季节下的实际数据验证得到满意的结果。

在 ELM 的基础上, 以核函数代替未知的隐藏层特征映射形成核极限学习机(KELM)。与 ELM 相比, KELM 不需要确定隐藏层的节点数, 其通过选择合适的核参数与惩罚因子即可计算出输出权值。文献[19]提出一类 O-KELM 方法, 分别基于 GA、SA、DE 的进化算法来优化核参数与正则化系数。文献[20]提出一种将时域分析与传感器数据融合的方法优化 KELM 的分类算法, 成功应用于风力涡轮机的故障诊断, 并获得较高的精度。

本文针对风电功率时序下的随机性与波动性, 提出基于自适应灰色模型(SAIGM)和遗传算法优化核极限学习机(GA-KELM)的风电功率混合预测模型。首先, 将实际风电场中的历史数据进行提取筛选样本数据, 并进行归一化和灰色关联分析, 得出对风电功率变化影响最大的因素为风速; 然后, 搭建 SAIGM-GAKELM 预测模型, 将训练数据放入模型中获得最优的核参数和惩罚因子, 对预测数据中的风速使用 SAIGM 进行预测; 最后, 预测模型获得下个时序下的风电功率。仿真结果表明, 在相同条件下与 GA-ELM 相比, GA-KELM 具有更好的预测性能, 进一步验证了本文所提出方法的有效性。

## 1 核极限学习机与自适应灰色系统

### 1.1 核极限学习机

核极限学习机是文献[21]通过将极限学习机(ELM)与支持向量机(SVM)的原理进行对比后的非线性延伸, 本节在 ELM 的基础上介绍 KELM 方法。

ELM 的本质是一种单隐层前馈神经网络, 对于具有  $L$  个隐藏层节点的 SLFNs, 假定给出  $N$  组训练样本数据  $\{\mathbf{x}_j, \mathbf{t}_j\} \in \mathbf{R}^n \times \mathbf{R}^m$ , 其中  $\mathbf{x}_j = [x_{j1}, x_{j2}, \dots, x_{jn}]^T$ ,  $\mathbf{t}_j = [t_{j1}, t_{j2}, \dots, t_{jm}]^T$ , 则其节点输出为

$$\mathbf{y}(\mathbf{x}_j) = \sum_{i=1}^L \beta_i h(\mathbf{x}_j) = \sum_{i=1}^L \beta_i G(\mathbf{x}_j; \mathbf{w}_i, b_i), \quad (1)$$

$$j = 1, \dots, N$$

若具有  $L$  个隐藏层节点的 SLFNs 能够零误差地逼近任意  $N$  个样本, 即

$$\sum_{i=1}^N \|\mathbf{y}_i - \mathbf{t}_i\| = 0$$

此时有

$$\mathbf{t}_j = \sum_{i=1}^L \beta_i G(\mathbf{x}_j; \mathbf{w}_i, b_i), \quad j = 1, 2, \dots, N \quad (2)$$

式(1)可改写成矩阵形式:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \quad (3)$$

且有

$$\mathbf{H} = \begin{bmatrix} h(\mathbf{x}_1) \\ \vdots \\ h(\mathbf{x}_N) \end{bmatrix} = \begin{bmatrix} G(\mathbf{x}_1; \mathbf{w}_1, b_1) & \cdots & G(\mathbf{x}_1; \mathbf{w}_L, b_L) \\ \vdots & & \vdots \\ G(\mathbf{x}_N; \mathbf{w}_1, b_1) & \cdots & G(\mathbf{x}_N; \mathbf{w}_L, b_L) \end{bmatrix}_{N \times L};$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m}; \quad \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N \times m}.$$

式中:  $\mathbf{H}$  为隐藏层节点的输出矩阵;  $\boldsymbol{\beta}$  为输出层与隐藏层之间的权值矩阵;  $\mathbf{T}$  为期望输出矩阵。

当激励函数  $G(\mathbf{x}_j; \mathbf{w}_i, b_i)$  在任意区间无限微分时, SLFNs 隐藏层节点的输入权值和偏差值可以随机产生, 且训练时保持不变, 当隐藏层节点数  $L$  小于样本数  $N$  时, SLFNs 可以以极小的训练误差逼近训练样本, 此时 ELM 的训练模型可等价于求解线性系统  $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$  的最小二乘解  $\hat{\boldsymbol{\beta}}$ 。

$$\left\| \mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_L, b_1, \dots, b_L) \hat{\boldsymbol{\beta}} - \mathbf{T} \right\| = \min_{\boldsymbol{\beta}} \left\| \mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_L, b_1, \dots, b_L) \boldsymbol{\beta} - \mathbf{T} \right\| \quad (4)$$

在 ELM 算法中, 为了增强网络的稳定性和泛化能力, 根据岭回归思想, 在  $\mathbf{H}\mathbf{H}^T$  的对角线上加上小的正数  $\mathbf{I}/C$ , 则输出权值的最小二乘解为

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^T (\mathbf{H}\mathbf{H}^T + \mathbf{I}/C)^{-1} \mathbf{T} \quad (5)$$

式中:  $\mathbf{I}$  是单位对角矩阵;  $C$  是惩罚因子。因此 ELM 的输出函数为  $\mathbf{y}(\mathbf{x}) = h(\mathbf{x}) \hat{\boldsymbol{\beta}}$ 。

由于 ELM 中的  $h(\mathbf{x})$  与 SVM 中的核函数非常相

似, 特征映射函数  $h(\mathbf{x})$  未知时, 参考 SVM 中核函数的映射, 根据 Mercer 条件核矩阵的定义为

$$\mathbf{\Omega}_{ELM} = \mathbf{H}\mathbf{H}^T : \mathbf{\Omega}_{ELM,i,j} = h(\mathbf{x}_i) \cdot h(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$$

结合式(5), KELM 算法的输出函数表示为

$$\mathbf{y}(\mathbf{x}) = h(\mathbf{x})\hat{\boldsymbol{\beta}} = h(\mathbf{x})\mathbf{H}^T(\mathbf{H}\mathbf{H}^T + \mathbf{I}/C)^{-1}\mathbf{T} = \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_N) \end{bmatrix} \left( \frac{\mathbf{I}}{C} + \mathbf{\Omega}_{ELM} \right)^{-1} \mathbf{T} \quad (6)$$

KELM 算法中常用的核函数主要为多项式核函数、RBF 核函数、线性核函数、小波核函数等, 本文核函数的类型取高斯核函数, 其表达式为

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{\gamma^2}\right) \quad (7)$$

相比较传统的极限学习机, KELM 算法引入的核函数具有很强的非线性映射能力, 可以更好地实现函数逼近功能, 并且规避传统 ELM 算法由于隐藏层节点的偏差和输入权值的随机设定带来的系统不稳定问题。引入核函数, 使得 KELM 算法在确保学习速度的同时, 提高系统的鲁棒性和结果的精确性。

### 1.2 自适应灰色系统

传统的 GM(1,1)模型由于单一固定的指数型结构导致系统适应性差, Zeng 等人根据此问题提出具有可变结构的自适应智能灰色预测模型 (Self-adaptive Intelligence Grey predictive model with alterable structure, SAIGM)<sup>[22]</sup>。该模型具有参数可调的优点, 其结构不仅适用于齐次或非齐次指数模型, 也适用于单变量线性自回归模型。该模型可自动求解相对最优参数, 以适应模型序列中的实际数据特征。

与 GM(1,1)的序列相同, 其初始序列  $X^{(0)}$ , 累计序列  $X^{(1)}$ , 紧邻均值生成序列  $Z^{(1)}$  可表示为

$$\begin{cases} X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(k)) \\ X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(k)) \\ Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(k)) \end{cases} \quad (8)$$

其中,

$$\begin{cases} x^{(0)}(k) \geq 0 (k=1, 2, \dots, n), \\ x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) (k=1, 2, \dots, n), \\ z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) (k=2, 3, \dots, n) \end{cases}$$

则 SAIGM 微分方程是 GM(1,1)的扩展形式, 新增变量  $c$  是 SAIGM 中可改变结构的主要参数。

$$x^{(0)}(k) + az^{(1)}(k) = kb + c \quad (9)$$

SAIGM 模型的最小二乘估计参数满足

$$\hat{\mathbf{a}} = (a, b, c)^T = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}, \quad \text{其中}$$

$$\mathbf{Y} = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -z^{(1)}(2) & 2 & 1 \\ -z^{(1)}(3) & 3 & 1 \\ \vdots & \vdots & \vdots \\ -z^{(1)}(n) & n & 1 \end{bmatrix}$$

当使用  $z^{(1)}(k) = 0.5 \times [\hat{x}^{(1)}(k) + \hat{x}^{(1)}(k-1)]$  与  $x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1)$  代替式(9)中的  $z^{(1)}(k)$  与  $x^{(0)}(k)$ , 可得

$$(1+0.5a)x^{(1)}(k) + (0.5a-1)x^{(1)}(k-1) = kb + c \quad (10)$$

求解式(10)的微分方程, 则 SAIGM 模型的时间响应函数  $\hat{x}^{(1)}(k+1)$  和还原函数  $\hat{x}^{(0)}(k+1)$  为

$$\hat{x}^{(1)}(t) = x^{(0)}(1) \left( \frac{1-0.5\hat{a}}{1+0.5\hat{a}} \right)^{t-1} + \sum_{i=0}^{t-2} \left[ (t-i) \frac{\hat{b}}{1+0.5\hat{a}} + \frac{0.5\hat{b}-\hat{c}}{1+0.5\hat{a}} \right] \left( \frac{1-0.5\hat{a}}{1+0.5\hat{a}} \right)^i \quad (11)$$

$$\hat{x}^{(0)}(t) = \hat{x}^{(1)}(t) - \hat{x}^{(1)}(t-1) =$$

$$\left[ x^{(0)}(1) \left( \frac{-a}{1+0.5a} \right) + \frac{2b+c}{1+0.5a} \right] \left( \frac{1-0.5a}{1+0.5a} \right)^t + \quad (12)$$

$$\sum_{g=0}^{t-3} \frac{b}{1+0.5a} \left( \frac{1-0.5a}{1+0.5a} \right)^g$$

## 2 基于优化极限学习机的预测模型

### 2.1 GA-KELM 算法

在 KELM 算法中, 系统性能主要会受到核参数  $\gamma$  和惩罚因子  $C$  选择的影响, 一般会采用交叉验证的方式进行参数确认。为了解决上述参数选择带来的问题, 将遗传算法群体搜索最优值的能力与 KELM 快速学习的能力相结合, 对模型的核参数  $\gamma$  和惩罚因子  $C$  择优进化, 进一步提升 KELM 算法的学习性能和泛化能力。因此, 本节提出一种 GA-KELM 算法。

待优化的决策变量核参数  $\gamma$  和惩罚因子  $C$  是实数变量, 为扩大搜索范围将分解的决策变量  $\boldsymbol{\gamma} = [\gamma_1 \ \gamma_2 \ \gamma_3]$  与  $\mathbf{C} = [C_1, C_2, \dots, C_n]$  映射为种群中的染色体, 可表示为  $\boldsymbol{\lambda} = [\lambda_1 \ \lambda_2 \ \dots \ \lambda_m]$ 。在 GA-KELM 的算法构造中, 选取均方根差(Root Mean Squard Error, RMSE)作为染色体适应度函数, 即

$$J_{\text{fitness}} = E_{\text{RMSE}}(y, y_d) = \sqrt{\frac{1}{N} \sum_{j=1}^N [y(j) - y_d(j)]^2} \quad (13)$$

式中:  $y(j)$  是模型的预测输出;  $y_d(j)$  是期望输出。

GA-KELM 通过选择、交叉、变异等操作, 将选取优化决策变量转化为求取最小适应度值下对应的最优染色体, 从而计算预测模型的输出, 具体算法实现步骤如下。

**Step1:** 种群初始化, 初始种群  $R$  采用线性插值的方式随机生成, 即

$$R = [\lambda_1, \lambda_2, \dots, \lambda_m] \quad (14)$$

式中:  $\lambda_i$  为第  $i$  个染色体;  $m$  为种群的规模。

**Step2:** 对染色体进行解码, 得到决策变量  $\gamma$  和  $C$ , 将其代入 KELM 的训练模型中训练。

**Step3:** 将训练参数代入到测试模型中, 根据公式(13)计算种群个体的适应度。

**Step4:** 根据适应度值大小通过轮盘赌算法选择种群中最好适应度的个体作为父代和母代, 并以概率  $p_c$  进行单点基因交叉, 从而得到新的种群。

其中变异操作以较小概率  $p_m$  对染色体的某个基因自适应均匀变异。精英保留策略将每代中最优染色体和适应度值保留至下一代。

**Step5:** 进化是否达到阈值, 若是, 进化终止; 否则, 继续进化, 转至 Step3。

**Step6:** 将进化后的最优染色体解码得到最优的核参数  $\gamma$  和惩罚因子  $C$  代入到测试模型中, 将预测输出与样本值选取合适的评价指标进行分析。

## 2.2 灰色关联度

在系统中, 多个变量的输入与输出之间的关联度是不一样的, 必然存在起到决定性影响的因素, 因此关联性分析显得尤其重要。本文采用灰色关联分析来分析风电功率与风速、风向和其他气象因素, 比如空气湿度、温度、地表气压之间的关系, 为关联度较大因素的评估做准备。

灰色关联分析是系统发展变化态势的量化描述和比较的方法, 通过确定参考数据列与若干比较数据列的变化趋势的相似性以衡量因素间的关联程度<sup>[23]</sup>。因此, 在分析影响风电功率输入因素的关联性时, 首先需要确定预测的风电功率作为参考序列  $X_0 = \{x_0(1), x_0(2), \dots, x_0(N)\}$ , 影响系统行为的风向量与其他气象数据(气温、气压、湿度)作为比较序列  $X_i = \{x_i(1), x_i(2), \dots, x_i(N)\}, i \in \{1, 2, \dots, D\}$ 。但系统中各因素序列采用不同的量纲导致计算比较困难, 因此必须将数据进行无量纲化处理, 本文采用均值化方法, 其公式为

$$r_i(k) = \frac{x_i(k)}{\frac{1}{N} \sum_{k=1}^N x_i(k)} \quad (15)$$

则可以分别计算每个比较序列与参考序列的关联系数, 其计算为

$$\gamma_i(k) = \frac{\min_i \min_k |r_0(k) - r_i(k)| + \rho \max_i \max_k |r_0(k) - r_i(k)|}{|\min_i \min_k |r_0(k) - r_i(k)| + \rho \max_i \max_k |r_0(k) - r_i(k)|} \quad (16)$$

式中:  $\gamma_i(k)$  是参考序列与第  $i$  组比较序列在时间  $k$  的关联系数;  $r_0(k)$  为参考序列在时间  $k$  的历史数据;  $r_i(k)$  为比较序列在时间  $k$  的历史数据;  $\rho$  表示分辨系数,  $0 < \rho < 1$ , 本文为突显最大影响因素的关联程度取  $\rho = 0.5$ <sup>[24]</sup>。

由于关联系数是参考序列与比较序列在各个时间的关联程度值, 是数组的形式, 过于分散不利于整体比较, 因此有必要将各个时刻的关联系数集中求取平均值, 以此作为整体关联性系数  $\gamma_i$ , 其公式为

$$\gamma_i = \frac{1}{N} \sum_{k=1}^N \gamma_i(k) \quad (17)$$

最后对所求的关联度系数进行排序, 数值越大则代表比较序列与参考序列越相似, 其影响效果越大, 对此进行预测估计的价值越大。

## 2.3 风电功率混合预测模型

根据自适应灰色系统自动调节参数以适应实际数据特征的突出优势, 本文提出一种基于自适应灰色系统和遗传算法优化核极限学习机(SAIGM-GAKELM)的风电功率混合预测模型, 该模型结构如图 1 所示, 具体步骤如下。

**Step1:** 采用灰色关联性方法通过计算不同季节下, 风电功率与风向量 and 气象因素的历史数据获取关联系数, 反映不同季节下对风电场输出功率影响最大的因素。

**Step2:** 对影响最大的因素采用自适应灰色系统进行预测估计, 作为测试集中的输入因素。

**Step3:** 使用 GA-KELM 算法对训练集下历史数据进行训练, 并引用遗传算法不断优化核参数和惩罚因子, 以此选取最优染色体作为预测模型的决策因素。

**Step4:** 将最优决策因素与测试集数据代入预测模型中, 求取不同季节下风电功率的预测输出。

**Step5:** 选择均方根误差和平均绝对误差 2 种误差评价指标对预测结果进行性能分析。

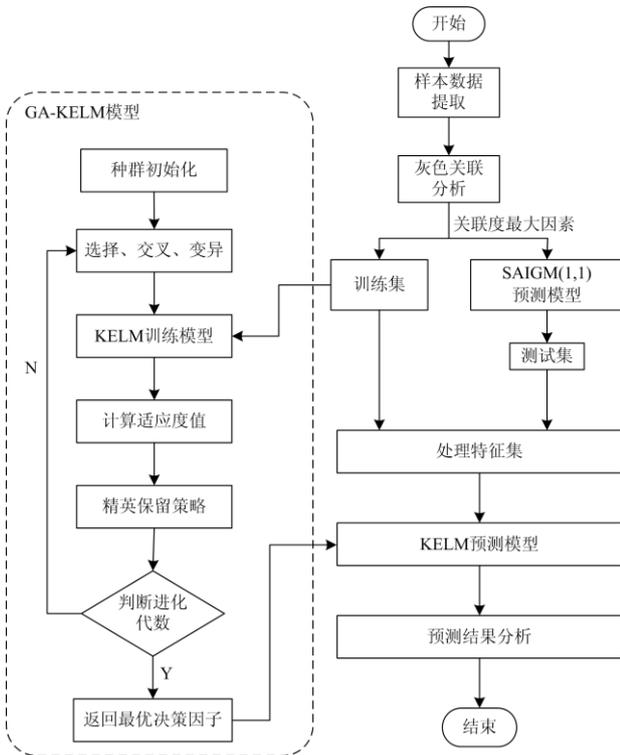


图 1 风电功率预测模型结构图

Fig. 1 Schematic diagram of wind power prediction model

### 3 算例分析

为了更好地分析风电功率预测值与期望输出值之间的误差，通过选取合适的评价指标均方根误差(RMSE)和平均绝对误差(Mean Absolute Error, MAE)对误差进行量化与分析，其公式为

$$E_{RMSE}(y, y_d) = \sqrt{\frac{1}{N} \sum_{t=1}^N [y(t) - y_d(t)]^2} \quad (18)$$

$$E_{MAE} = \frac{1}{N} \sum_{t=1}^N |y(t) - y_d(t)| \quad (19)$$

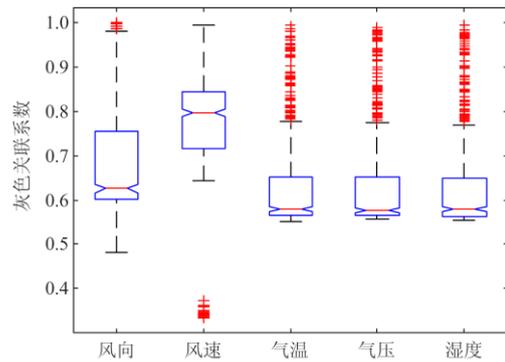
式中： $y(t)$  和  $y_d(t)$  分别为  $t$  时刻功率的实际值与预测值； $N$  为采样总时间。

#### 3.1 灰色关联性分析

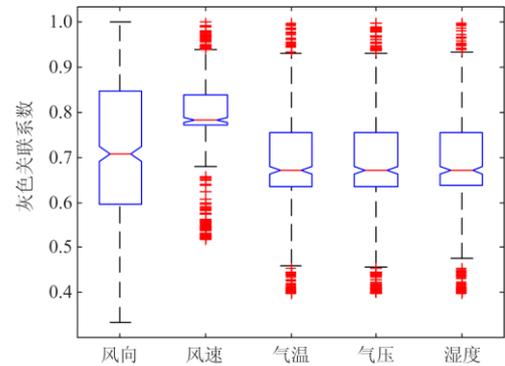
天气条件与风速在不同的季节变化很大，气候系统的混沌性与复杂性导致风力发电具有高度不确定性<sup>[25]</sup>。为了研究拟议算法的适用性与有效性，本文采用美国国家可再生能源实验室(NREL)提供的2012年科罗拉多州丹佛市西部方向风电场(经度： $-105.31$ ，纬度： $39.71$ )的真实数据集，并随机选取不同季节中连续11天的历史数据：1月21—31日，4月9—19日，7月9—19日，11月12—22日，分别对应冬天、春天、夏天和秋天，时间间隔为30 min的数据作为实验数据，因此无法保证风电场始终处

于最佳的工作状态。仿真环境是 Matlab2018b，一台配置为8 GB内存，且AMD FX-7500 2.1 GHz四核处理器的台式电脑。

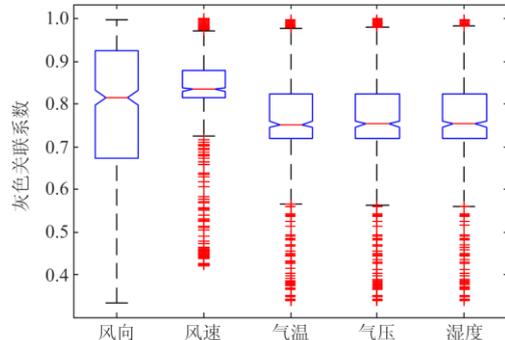
不同季节下风电功率与风速、风向、气象因素(气温、气压、空气湿度)的关联盒图如图2所示，表1列举了不同时刻关联系数集中计算得到的平均值。根据表1和图2可以看出，风速的关联系数最大，即风速与风电功率的变化发展趋势最紧密，其次是风向，气温、气压、湿度之间的关联系数基本一致。从盒图离散点的角度来看，风向在不同季节下关联系数的异常值都最少；从盒图的长度、上下隔间的形状以及胡须的长度来看，风速除在冬天的长度略大于气象因素外，在春天、夏天、秋天的长度都明显最小。



(a) 冬天



(b) 春天



(c) 夏天

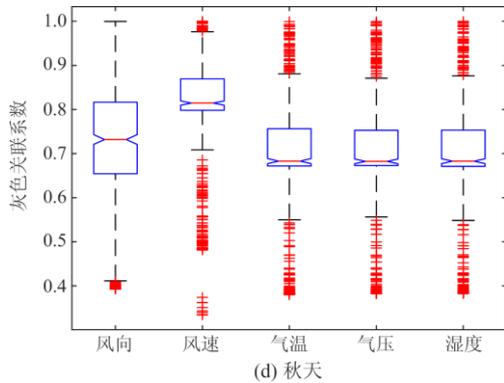


图 2 不同季节下风电功率与输入因素的关联性盒图

Fig. 2 Box diagram of wind power and input factors in different seasons

表 1 不同季节下风电功率与输入因素灰色关联系数表

Table 1 Grey correlation coefficient table of wind power and input factors in different seasons

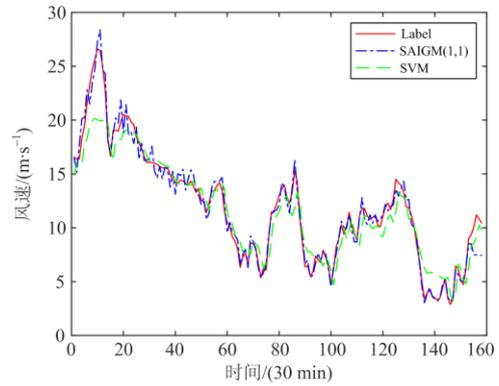
关联系数	风向	风速	气温	气压	湿度
冬天	0.682 3	0.783 8	0.628 9	0.628 3	0.627 9
春天	0.709 0	0.777 3	0.679 3	0.679 2	0.679 6
夏天	0.784 1	0.821 7	0.750 3	0.750 4	0.750 4
秋天	0.719 1	0.795 4	0.692 4	0.691 5	0.691 3

通过以上对风电功率和风向量、气象因素之间关联系数的分析, 可以直观得到与风电功率变化趋势关系最紧密的因素是风速, 因此, 提前预测风速对短期风电功率的预测更具有实际意义。

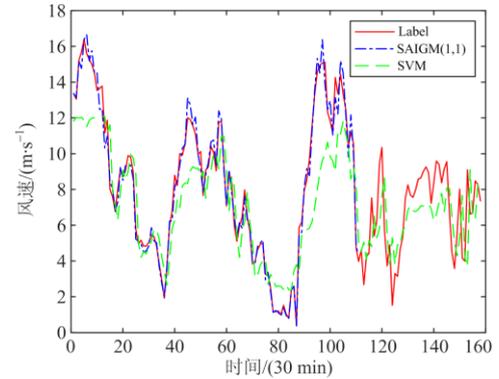
### 3.2 SAIGM-GAKELM 混合预测

在 GA-KELM 的实验中, KELM 的核函数选取 RBF 函数, GA 的种群规模为 100, 最大进化代数取 100 代, 交叉概率  $p_c$  取 0.8, 变异概率  $p_m$  取 0.1。随机选取不同季节连续 11 天风电场历史数据, 以 30 min 为单位进行采样, 即不同季节下连续 528 个数据样本。输入变量为风速、风向、气温、气压、湿度 5 个自然因素, 输出变量为风电场实际的输出风电功率, 将数据无量纲化处理, 归一化到  $[-1, 1]$  区间, 前 70% 的数据作为训练集, 后 30% 作为测试集。

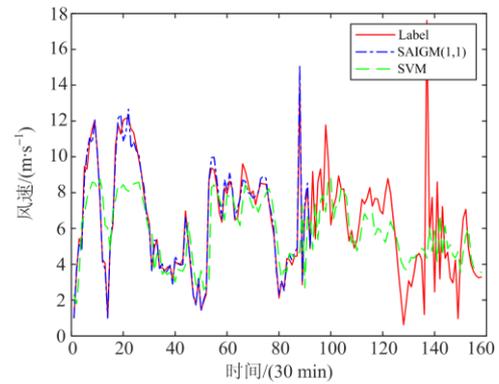
由前面的灰色关联分析实验可以直观得出, 与风电功率变化趋势关联最大的是风速, 因此, 对风速的提前预测具有重要意义。使用 SAIGM(1,1) 对测试集中的风速进行预测, 将风速实际值替换为预测值。如图 3 所示, 对不同季节下测试集的风速进行预测, 预测风速值与实际风速值基本拟合, 波动性较小, 拟合性能较好。采用 SVM 方法对不同季节下 528 个风速数据样本进行预测对比试验, 前 70% 作为训练数据, 后 30% 作为测试数据, 预测风速值



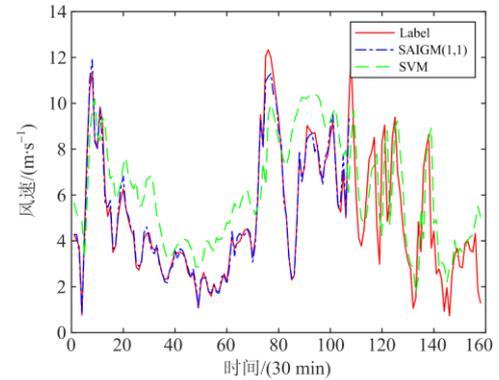
(a) 冬天



(b) 春天



(c) 夏天



(d) 秋天

图 3 SAIGM(1,1)与 SVM 对不同季节风速的预测对比  
Fig. 3 Comparison of wind speed prediction by SAIGM (1,1) and SVM in different seasons

与实际风速值虽然变化趋势相似,但是波动性较大,拟合性能一般。因此,选取 SAIGM(1,1)的方法对风速进行预测更具有实际价值。

从图 4 的进化曲线可以直观看出,冬天在第 48 代开始收敛,春天在第 58 代开始收敛,夏天在第 50 代开始收敛,秋天在第 33 代就开始收敛,总体收敛速度基本维持在 50 代左右,较稳定。从图 5 实际功率与预测功率的对比可以看出,预测曲线基本符合实际变化趋势,且拟合性较好,但秋天和冬天功率样本值在第 20 到 60 采样点之间较集中,特征区分不明显,使得预测功率波动较大。从图 6 的功率预测误差曲线可以看出,不同季节下风电功率的预测值与实际值误差在 $[-4, 4]$ 波动,基本符合实际风电功率的变化范围。

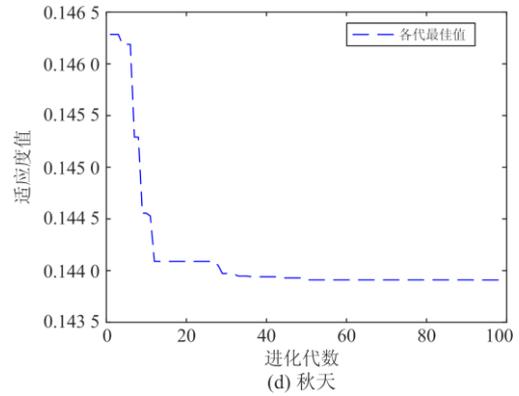
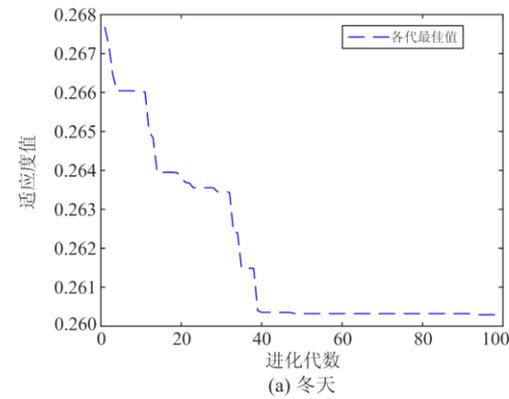
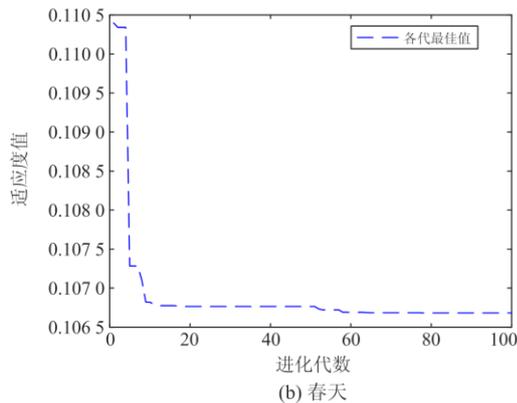


图 4 不同季节下适应度值的进化曲线

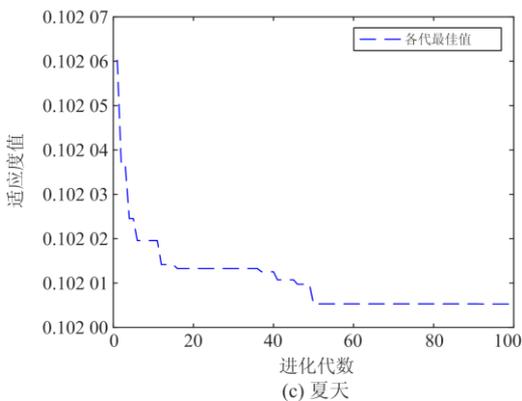
Fig. 4 Evolution curve of fitness value in different seasons



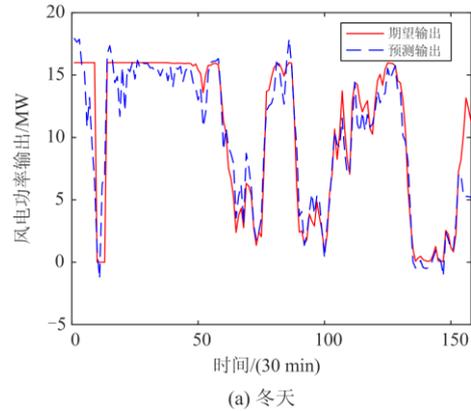
(a) 冬天



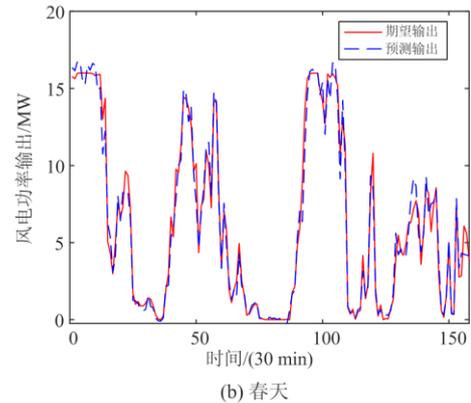
(b) 春天



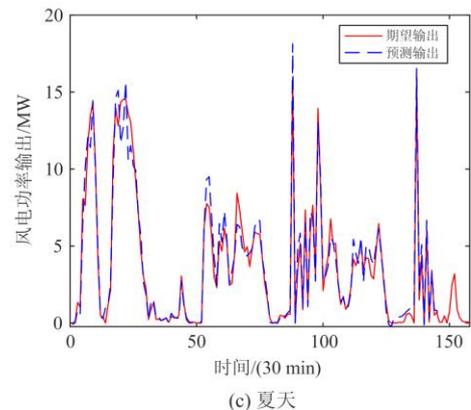
(c) 夏天



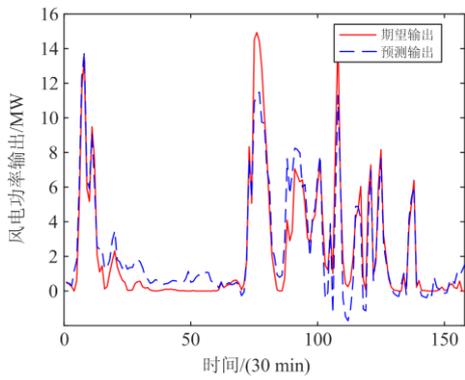
(a) 冬天



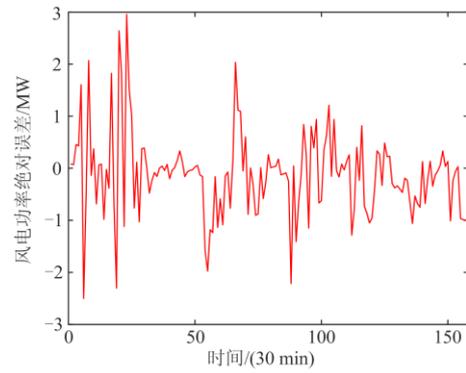
(b) 春天



(c) 夏天

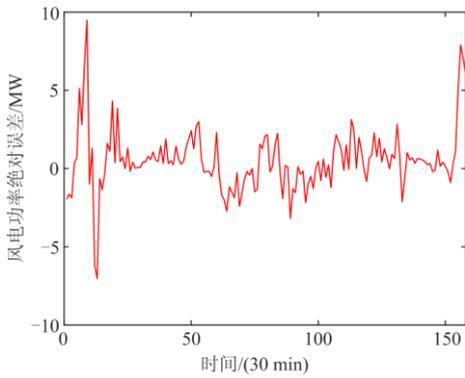


(d) 秋天

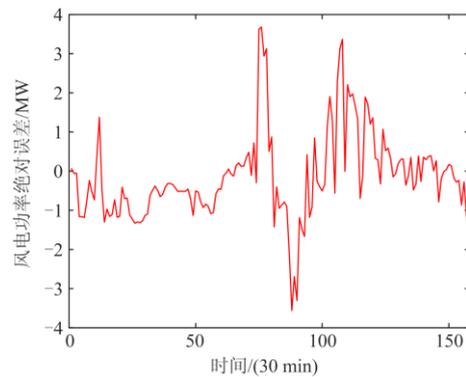


(c) 夏天

图 5 不同季节下测试集的实际功率与预测功率  
Fig. 5 Labels and prediction results of test samples in different seasons



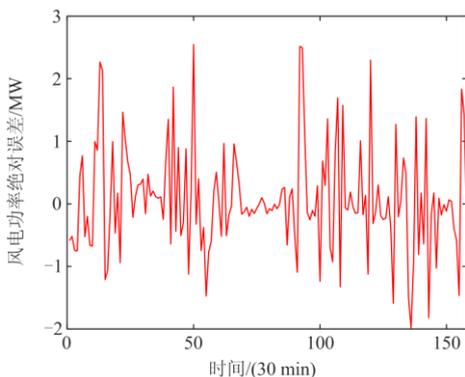
(a) 冬天



(d) 秋天

图 6 不同季节下测试集的功率预测误差

Fig. 6 Prediction errors of test samples in different seasons



(b) 春天

通过表 2 对比实验的数据可以看出, GAKELM 的评价指标整体均优于 GAELM, 表明带核函数的极限学习机学习效果更好, 更符合期望值。同时 SAIGM-GAELM 与 SAIGM-GAKELM 的评价指标均优于 SVM-GAELM 与 SVM-GAKELM, 表明 SAIGM 比 SVM 对风速的预测效果更符合实际值, 使得混合模型的预测功率与实际功率的波动更小。根据实验结果可知, SAIGM-GAKELM 具有最优的评价指标系数, 证明 SAIGM-GAKELM 预测模型性具有更好的预测精度以及最佳的泛化性能。

表 2 四种风电功率预测方法的对比

Table 2 Comparison of four wind power prediction methods

季节	评价指标	SAIGM-GAELM	SAIGM-GAKELM	SVM-GAELM	SVM-GAKELM
冬天	RMSE	0.529 4	0.260 3	0.534 3	0.395 1
	MAE	0.375 9	0.166 6	0.408 6	0.251 5
春天	RMSE	0.207 1	0.106 7	0.449 7	0.415 6
	MAE	0.170 9	0.076 8	0.345 6	0.306 0
夏天	RMSE	0.187 6	0.102 0	0.416 3	0.413 3
	MAE	0.133 9	0.070 1	0.268 8	0.264 9
秋天	RMSE	0.269 9	0.143 9	0.424 8	0.366 2
	MAE	0.236 7	0.105 2	0.361 2	0.265 0



## 4 结论

本文基于实际风电场时序下的历史数据进行灰色关联分析,利用 SAIGM(1,1)对关联最紧密的风速进行预测,并搭建 GA-KELM 的预测模型得到下个时序下的风电功率,从实验过程与结果可以得出如下结论。

(1) 灰色关联分析对多输入单输出系统发展变化趋势关联程度的描述比较具有重要意义,凸显决定性信息,为预测模型中关联度较大因素的评估做准备。

(2) 与 GM(1,1)相比,SAIGM(1,1)克服单一线性结构的局限性,使模型参数可自动调整求取最优解,适用于非线性系统的预测。从风速的仿真可以看出,SAIGM(1,1)对风速的变化具有更好的跟踪性。

(3) KELM 通过引入核函数,无需确认隐藏层节点的数目和特征映射函数,通过矩阵求解可获得输出权值。与 GA 的全局搜索能力和进化策略相结合,获得最优的核参数与惩罚因子,使得 GA-KELM 预测模型具有更好的泛化性能与预测精度。

本文所提预测模型具有较好自适应性与泛化性,在风电功率短期预测中具有良好的应用潜力。

## 参考文献

- [1] 薛禹胜, 郁琛, 赵俊华, 等. 关于短期及超短期风电功率预测的评述[J]. 电力系统自动化, 2015, 39(6): 141-151.
- XUE Yusheng, YU Chen, ZHAO Junhua, et al. A review on short-term and ultra-short-term wind power prediction[J]. Automation of Electric Power Systems, 2015, 39(6): 141-151.
- [2] REN Y, SUGANTHAN P N, SRIKANTH N. Ensemble methods for wind and solar power forecasting—a state-of-the-art review[J]. Renewable and Sustainable Energy Reviews, 2015, 50: 82-91.
- [3] FOLEY A M, LEAHY P G, MARVUGLIA A, et al. Current methods and advances in forecasting of wind power generation[J]. Renewable Energy, 2012, 37(1): 1-8.
- [4] JUNG J, BROADWATER R P. Current status and future advances for wind speed and power forecasting[J]. Renewable and Sustainable Energy Reviews, 2014, 31: 762-777.
- [5] 李俊芳, 张步涵, 谢光龙, 等. 基于灰色模型的风速-风电功率预测研究[J]. 电力系统保护与控制, 2010, 38(19): 151-159.
- LI Junfang, ZHANG Buhan, XIE Guanglong, et al. Grey predictor models for wind speed-wind power prediction[J]. Power System Protection and Control, 2010, 38(19): 151-159.
- [6] CASSOLA F, BURLANDO M. Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output[J]. Applied Energy, 2012, 99: 154-166.
- [7] SHI J, QU X, ZENG S. Short-term wind power generation forecasting: direct versus indirect ARIMA-based approaches[J]. International Journal of Green Energy, 2011, 8(1): 100-112.
- [8] LIU H, TIAN H, LI Y. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction[J]. Applied Energy, 2012, 98: 415-424.
- [9] 据焱, 祁林, 刘帅. 基于改进乌鸦算法和 ESN 神经网络的短期风电功率预测[J]. 电力系统保护与控制, 2019, 47(4): 58-64.
- JU Yao, QI Lin, LIU Shuai. Short-term wind power forecasting based on improved crow search algorithm and ESN neural network[J]. Power System Protection and Control, 2019, 47(4): 58-64.
- [10] 叶瑞丽, 郭志忠, 刘瑞叶, 等. 基于小波包分解和改进 Elman 神经网络的风电场风速和风电功率预测[J]. 电工技术学报, 2017, 32(21): 103-111.
- YE Ruili, GUO Zhizhong, LIU Ruiye, et al. Wind speed and wind power forecasting method based on wavelet packet decomposition and improved Elman neural network[J]. Transactions of China Electrotechnical Society, 2017, 32(21): 103-111.
- [11] 肖迁, 李文华, 李志刚, 等. 基于改进的小波-BP 神经网络的风速和风电功率预测[J]. 电力系统保护与控制, 2014, 42(15): 80-86.
- XIAO Qian, LI Wenhua, LI Zhigang, et al. Wind speed and power prediction based on improved wavelet-BP neural network[J]. Power System Protection and Control, 2014, 42(15): 80-86.
- [12] WANG H Z, LI G Q, WANG G B, et al. Deep learning based ensemble approach for probabilistic wind power forecasting[J]. Applied Energy, 2017, 188: 56-70.
- [13] ZHANG Q, LAI K K, NIU D, et al. A fuzzy group forecasting model based on least squares support vector machine (LS-SVM) for short-term wind power[J]. Energies, 2012, 5(9): 3329-3346.

- [14] YANG L, HE M, ZHANG J, et al. Support vector machine enhanced markov model for short-term wind power forecast[J]. IEEE Transactions on Sustainable Energy, 2015, 6(3): 791-799.
- [15] ZHENG D, ESEYE A T, ZHANG J, et al. Short-term wind power forecasting using a double-stage hierarchical ANFIS approach for energy management in microgrids[J]. Protection and Control of Modern Power Systems, 2017, 2(2): 136-145. DOI: 10.1186/s41601-017-0041-5.
- [16] HAQUE A U, NEHRIR M H, MANDAL P. A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting[J]. IEEE Transactions on Power Systems, 2014, 29(4): 1663-1672.
- [17] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(1): 489-501.
- [18] WAN C, XU Z, PINSON P, et al. Probabilistic forecasting of wind power generation using extreme learning machine[J]. IEEE Transactions on Power Systems, 2014, 29(3): 1033-1044.
- [19] 李军, 李大超. 基于优化核极限学习机的风电功率时间序列预测[J]. 物理学报, 2016, 65(13): 1-10.  
LI Jun, LI Dachao. Wind power time series prediction using optimized kernel extreme learning machine method[J]. Acta Physica Sinica, 2016, 65(13): 1-10.
- [20] LONG X, YANG P, GUO H, et al. A CBA-KELM-based recognition method for fault diagnosis of wind turbines with time-domain analysis and multisensor data fusion[J]. Shock and Vibration, 2019: 1-14.
- [21] HUANG G B. An insight into extreme learning machines: random neurons, random features and kernels[J]. Cognitive Computation, 2014, 6(3): 376-390.
- [22] ZENG B, MENG W, TONG M. A self-adaptive intelligence grey predictive model with alterable structure and its application[J]. Engineering Applications of Artificial Intelligence, 2016, 50: 236-244.
- [23] 刘思峰, 蔡华, 杨英杰, 等. 灰色关联分析模型研究进展[J]. 系统工程理论与实践, 2013, 33(8): 2041-2046.  
LIU Sifeng, CAI Hua, YANG Yingjie, et al. Advance in grey incidence analysis modelling[J]. Systems Engineering-Theory & Practice, 2013, 33(8): 2041-2046.
- [24] 东亚斌, 段志善. 灰色关联度分辨系数的一种新的确定方法[J]. 西安建筑科技大学学报: 自然科学版, 2008, 40(4): 589-592.  
DONG Yabin, DUAN Zhishan. A new determination method for identification coefficient of grey relational grade[J]. Journal of Xi'an University of Architecture & Technology: Natural Science Edition, 2008, 40(4): 589-592.
- [25] BELU R, KORACIN D. Wind characteristics and wind energy potential in western Nevada[J]. Renewable Energy, 2009, 34(10): 2246-2251.

收稿日期: 2019-10-29; 修回日期: 2020-01-17

作者简介:

王浩(1994—), 男, 硕士研究生, 研究方向为智能算法与预测控制; E-mail: 6171905064@stu.jiangnan.edu.cn

王艳(1978—), 女, 通信作者, 博士, 教授, 研究方向为制造系统能效优化; E-mail: wangyan@jiangnan.edu.cn

纪志成(1959—), 男, 博士, 教授, 研究方向为预测控制与新能源。E-mail: zcji@jiangnan.edu.cn

(编辑 许威)