



基于误差修正和 VMD-ICPA-LSSVM 的短期风速预测建模

摘要

精准的风速预测是将风能大规模应用到电力系统中的关键,而风速序列的随机性和波动性等特点使得风速预测难度增加。为增强风速序列的可预测性,采用 Logistic 混沌映射策略、自适应参数调整策略以及引入变异策略对食肉植物算法(CPA)进行改进,并提出了基于误差修正和 VMD-ICPA-LSSVM 的短期风速预测模型。首先将气象因子作为最小二乘支持向量机(LSSVM)的输入对风速进行预测,获得误差序列。再利用 K-L 散度自适应地确定变分模态分解(VMD)的参数,并对误差序列进行分解。结合改进食肉植物算法(ICPA)优化 LSSVM 可调参数的方法来预测分解的子序列。叠加各子序列预测结果后对原始预测序列进行误差修正,进而得到最终风速预测值。实验结果表明,与其他模型相比,所提模型有着更好的预测精度和泛化性能。

关键词

变分模态分解;食肉植物算法;最小二乘支持向量机;误差修正;风速预测

中图分类号 TP183;TM614

文献标志码 A

收稿日期 2023-04-21

资助项目 国家自然科学基金(71961001);东华理工大学研究生创新基金(DHYC-202225)

作者简介

钟琳,女,硕士生,研究方向为智能计算及其应用.874451912@qq.com

颜七笙(通信作者),男,博士,教授,主要研究方向为智能计算及其应用.yanqs93@126.com

0 引言

风能作为一种清洁的可再生能源,对缓解化石能源的消耗带来的环境问题有着重要意义^[1]。精准的风速预测是提高风力资源利用率、将风能大规模地应用到电力系统中的关键^[2-3]。

目前,风速预测模型主要可分为物理模型和统计模型^[4]两大类。物理模型需在气象因子及其他相关因素的基础上建模,模型较为复杂,普适性不高^[5]。统计模型则是通过建立数据特征间的映射关系来实现对风速的预测^[6-7]。常见的统计模型有时间序列分析^[8]、神经网络^[9-10]、支持向量机(Support Vector Machine, SVM)^[11]以及参数优化的预测模型^[12-13]。相比物理模型,统计模型虽简洁易实现,但也存在一定的不足,例如:由于风速的随机性,时间序列模型在非线性的风速序列预测中效果不佳,神经网络模型对于较复杂的序列学习程度不深,而支持向量机则对大样本数据预测精度不高。由于风速序列随机性,单一预测模型难以反映风速的变化规律。

为了提高风速的预测精度,集成各模型优点的混合模型成为研究焦点。混合模型实际上是采用各种分解算法与不同预测模型进行组合预测^[14-16],从而提高预测精度。Xu 等^[17]对风速序列进行变分模态分解(Variational Mode Decomposition, VMD),为消除残差序列异方差性,建立了改进差分自回归滑动平均模型对子序列进行预测,在短期风速预测上精度较高;桑茂景等^[18]对风速序列进行集合经验模态分解,利用样本熵选出不同频率分量,并用深度学习方法对其进行多步风速预测,结果表明预测性能较好;Ma 等^[19]提出一种局部均值分解(Local Mean Decomposition, LMD)和改进多元宇宙算法(Improve Multi Verse Optimizer, IMVO)优化最小二乘支持向量机(Least Squares Support Vector Machine, LSSVM)的风速预测模型,较大程度地提高了预测的准确性;摆玉龙等^[20]采用 VMD 对误差序列进行分解,利用长短期记忆神经网络(Long Short-Term Memory, LSTM)模型进行预测,结果表明对误差序列进行 VMD 分解可以提高误差修正策略的效果。运用分解算法对风速序列先分解再预测的组合模型未考虑到气象因子对风速的影响,在一定程度上影响了预测的精度。此外,优化算法对于预测结果的精度也存在一定的影响,性能好的优化算法可以较

1 东华理工大学 理学院,南昌,330013

好地提高预测精度。

针对以上问题,鉴于 VMD 算法能够有效地减少模态混叠现象,各分量能很好地反映数据规律^[21]以及 LSSVM 算法训练速度快、泛化能力强^[19]的优点,结合风速时序特点以及气象因素的影响,在改进 CPA (Carnivorous Plant Algorithm, 食肉植物算法)收敛快、寻优能力强的基础上,本文提出一种时间与特征融合的误差修正和 VMD-ICPA-LSSVM 短期风速预测模型。首先采用 LSSVM 进行多变量风速预测,再利用 K-L 散度自适应地确定 VMD 的参数,并对误差序列进行分解,结合改进食肉植物算法(ICPA)优化 LSSVM 的参数预测分解子序列,最后对原始预测风速序列进行误差修正,得到最终风速预测值。在仿真实验中,此方法较好地提高了风速预测的精度。

1 主要模型算法

1.1 K-L 散度

K-L 散度又称相对熵,可衡量两个分布间的信息损失程度^[22]。若两者差异越小,K-L 散度则越小,若两者差异越大,K-L 散度也越大。当两者分布一致时,K-L 散度为 0。其公式如下:

$$\text{KL}(P\|Q) = \sum P(x) \log \frac{P(x)}{Q(x)}. \quad (1)$$

其中, $P(x)$, $Q(x)$ 是随机变量 X 上的概率分布。

1.2 变分模态分解

变分模态分解的核心是构建和求解变分问题^[23]。具体步骤如下:

1) 构造变分问题。在模态之和与原始信号相等的约束条件下表达式^[24]为

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}, \\ \text{s.t. } \sum_{k=1}^K u_k = S. \end{cases} \quad (2)$$

其中, $S(t)$ 为原始信号, $\{u_k\}$ 为分解的 K 个模态分量, $\{\omega_k\}$ 为各分量的中心频率, ∂_t 为偏导数计算, $\delta(t)$ 为单位脉冲函数, j 为虚数单位, $*$ 为卷积运算^[22]。

2) 求解约束变分表达式。引入惩罚因子 α 、Lagrange 乘法算子 λ , 得增广 Lagrange 表达式为

$$L[\{u_k(t)\}, \{\omega_k\}, \lambda(t)] = \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 +$$

$$\left\| S(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \langle \lambda(t), S(t) - \sum_{k=1}^K u_k(t) \rangle. \quad (3)$$

3) 求解增广 Lagrange 表达式鞍点。初始化参数 $u_1, \omega_1, \lambda_1, n$ 并设置 $n+1$ 循环, u_k, ω_k, λ 根据式(4)、(5)、(6)交替更新,当分量满足(7)时,求解结束。

$$\hat{S}(\omega) = \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2$$

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{S}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2}, \quad (4)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega}, \quad (5)$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left(\hat{S}(\omega) - \sum_{k=1}^K \hat{u}_k^{n+1}(\omega) \right), \quad (6)$$

$$\sum_{k=1}^K \frac{\|u_k^{n+1} - u_k^n\|_2^2}{\|u_k^n\|_2^2} < \varepsilon. \quad (7)$$

其中, \cdot 为傅里叶变换, n 为迭代次数, τ 为保真系数^[22]。

1.3 食肉植物算法

食肉植物算法(CPA)可通过模拟其生存技能来达到优化的效果^[25],步骤如下:

- 1) 随机初始化个体种群 POP 及参数。
- 2) 将训练集的均方根误差作为适应度函数计算个体适应度值。
- 3) 按适应度值升序对个体排序(考虑最小化问题)。

4) 个体分组时将最佳适应值的猎物分配给位居首位的食肉植物,以此类推。

5) 随机选择各组中的猎物,若吸引率 attraction_rate 高于随机值,则食肉植物生长。新食肉植物生长模型为

$$\begin{cases} \text{NewCP}_{i,j} = \text{growth} \times \text{CP}_{i,j} + (1 - \text{growth}) \times \text{Prey}_{v,j}, \\ \text{growth} = \text{growth_rate} \times \text{rand}_{i,j}. \end{cases} \quad (8)$$

其中, $\text{CP}_{i,j}$ 为第 i 个食肉植物, j 为维度, $\text{Prey}_{v,j}$ 为随机选择的猎物, growth_rate 为预定值, rand 为 $[0, 1]$ 上的随机值。若吸引率 attraction_rate 低于随机值,则猎物生长,其模型为

$$\begin{cases} \text{NewPrey}_{i,j} = \text{growth} \times \text{Prey}_{u,j} + (1 - \text{growth}) \times \text{Prey}_{v,j}, u \neq v, \\ \text{growth} = \begin{cases} \text{growth_rate} \times \text{rand}_{i,j}, f(\text{Prey}_v) > f(\text{Prey}_u), \\ 1 - \text{growth_rate} \times \text{rand}_{i,j}, f(\text{Prey}_v) < f(\text{Prey}_u). \end{cases} \end{cases} \quad (9)$$

其中, $\text{Prey}_{i,j}$ 为第 i 组中随机选择的猎物, $f(\cdot)$ 为求适应度值. 重复食肉植物和猎物的生长过程以达到组迭代值 group_iteration .

6) 排名第一的食肉植物可生长和繁殖. 其过程可表示为

$$\begin{cases} \text{NewCP}_{i,j} = \text{CP}_{1,j} + \text{Reproduction_rate} \times \text{rand}_{i,j} \times \text{mate}_{i,j}, \\ \text{mate}_{i,j} = \begin{cases} \text{CP}_{v,j} - \text{CP}_{i,j}, & f(\text{CP}_i) > f(\text{CP}_v), \\ \text{CP}_{i,j} - \text{CP}_{v,j}, & f(\text{CP}_i) < f(\text{CP}_v), \end{cases} & i \neq v \neq 1. \end{cases} \quad (10)$$

其中, Reproduction_rate 为预定值, $\text{CP}_{1,j}$ 为当前最优的食肉植物, $\text{CP}_{v,j}$ 为随机选择的第 v 个食肉植物.

7) 将新产生的个体与原种群重组排序, 选择前 N 的个体作为新的候选解.

8) 判断是否满足迭代停止条件. 若满足, 迭代停止; 否则, 重复步骤 3) — 7).

1.4 最小二乘支持向量机

最小二乘支持向量机是 SVM 算法的改进. 对于给定训练集的输入 $X = \{x_1, x_2, \dots, x_n\}$, 输出 $Y = \{y_1, y_2, \dots, y_n\}$, LSSVM 算法可表示为

$$y(\mathbf{x}) = \boldsymbol{\omega}^T \boldsymbol{\theta} + \mathbf{b}. \quad (11)$$

其中, $\boldsymbol{\theta}$ 为非线性映射, \mathbf{b} 为偏差向量, $\boldsymbol{\omega}^T$ 为权重向量. 基于风险最小化原则, 可转化为约束问题:

$$\begin{cases} \min_{\boldsymbol{\omega}, \mathbf{b}, \mathbf{e}} J(\boldsymbol{\omega}, \mathbf{e}) = \frac{1}{2} \|\boldsymbol{\omega}\|^2 + \frac{1}{2} \gamma \sum_{i=1}^n \mathbf{e}_i^2, \\ \text{s.t. } y_i(\boldsymbol{\omega}^T \boldsymbol{\varphi}(\mathbf{x}_i) + \mathbf{b}) = 1 - \mathbf{e}_i, i = 1, 2, \dots, n. \end{cases} \quad (12)$$

其中, \mathbf{e}_i 为误差, γ 为正则化因数, $\boldsymbol{\varphi}(\cdot)$ 为非线性映射函数.

引入 Lagrange 乘子 $\boldsymbol{\alpha}$, 可得:

$$L(\boldsymbol{\omega}, \mathbf{b}, \mathbf{e}, \boldsymbol{\alpha}) = J(\boldsymbol{\omega}, \mathbf{e}) - \sum_{i=1}^n \boldsymbol{\alpha}_i (y_i(\boldsymbol{\omega}^T \boldsymbol{\varphi}(\mathbf{x}_i) + \mathbf{b}) - 1 + \mathbf{e}_i). \quad (13)$$

对式(13)中 $\boldsymbol{\omega}, \mathbf{b}, \mathbf{e}, \boldsymbol{\alpha}$ 求偏导, 得到回归函数:

$$y(\mathbf{x}) = \sum_{i=1}^n \boldsymbol{\alpha}_i K(\mathbf{x}, \mathbf{x}_i) + \mathbf{b}. \quad (14)$$

式中, $K(\mathbf{x}, \mathbf{x}_i)$ 为核函数. 径向基核函数可表示为式(15), σ 为核宽度:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right). \quad (15)$$

2 改进食肉植物算法

为了更好地搜索最优解, 以达到预测精度和收敛速度提高的效果, 提出改进食肉植物算法 (Improve Carnivorous Plant Algorithm, ICPA). 对食肉植物算法的改进, 主要分为以下 3 方面:

1) Logistic 混沌映射策略优化初始种群

混沌映射初始化种群可有效解决解聚集、解空间覆盖率不高以及个体间差异小的问题. Logistic 混沌映射是混沌映射中的典型, 其数学表达式为

$$X_{i+1} = \mu X_i (1 - X_i). \quad (16)$$

其中, $X_i \in [0, 1]$, $\mu \in [0, 4]$ 为 Logistic 参数. μ 越接近 4, X 的取值范围则越均匀地分布到 $[0, 1]$ 区域; μ 取 4 时, 系统处于完全混沌状态, 映射分布均匀性达到极值.

部分初始种群如图 1 所示.

2) 自适应策略下的参数调整

在食肉植物算法的生长探索过程中, 算法的探

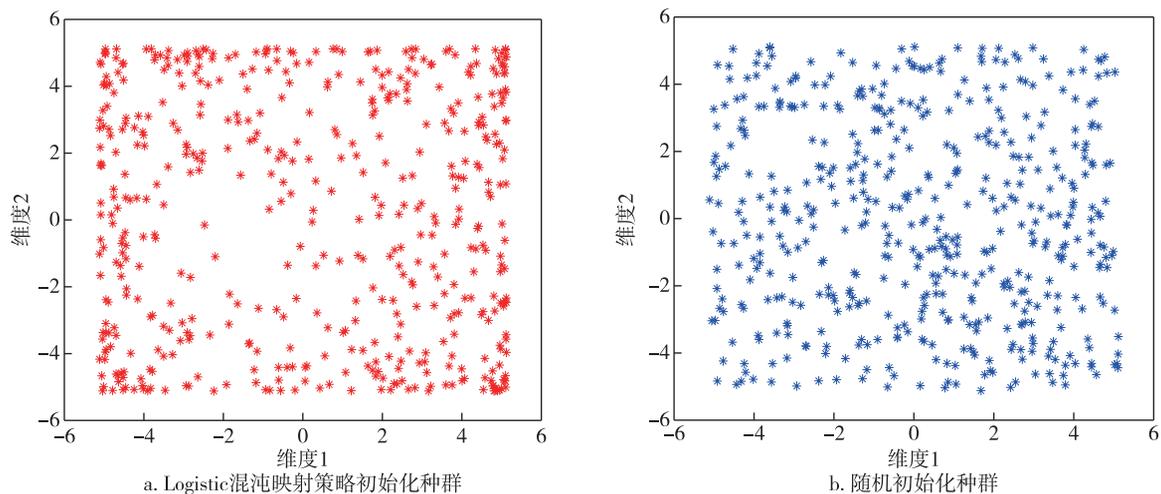


图 1 部分初始种群

Fig. 1 Logistic chaotic mapping strategy population (a) and random initialization population (b)

素受生长率 a 的影响.生长率 a 越大,搜索范围就越广,但错过全局最优解的可能性也越大.因此,需选择合适的生长率来调整算法的搜索精度.具体的自适应策略参数更新表达式如下:

$$a = \begin{cases} a_{\min} - \frac{(a_{\max} - a_{\min})(f - f_{\min})}{f_{\text{avg}} - f_{\min}}, & f \leq f_{\text{avg}}; \\ a_{\max}, & f > f_{\text{avg}}. \end{cases} \quad (17)$$

其中, a_{\max} 和 a_{\min} 分别为生长率 a 的最大值以及最小值, f 表示个体当前的适应度值, f_{avg} 和 f_{\min} 分别表示所有个体的平均适应度值和最小适应度值.通过加入自适应策略,能够较好地控制个体的搜索范围.

3) 个体自适应变异策略

食肉植物算法分为初始化种群、分类分组、生长、繁殖、重组 5 个过程,实际上在植物生长过程中也存在变异.为解决算法易陷入局部最优的问题,提出自适应个体变异策略,从而增加种群的多样性,同时引导种群向最优个体位置进化.

在算法的第 n 次迭代中,随机选取 3 个不同个体 $X_{r_1}(n), X_{r_2}(n), X_{r_3}(n)$ 进行变异操作,生成的变异个体为

$$N_i(n) = X_{r_1}(n) + F(X_{r_1}(n) - X_{r_3}(n)). \quad (18)$$

其中, F 为变异算子,其可通过以下方式进行自适应调整:

$$F = F_l + (F_u - F_l) \frac{f_{r_2} - f_{r_1}}{f_{r_3} - f_{r_1}}. \quad (19)$$

式中, $F_l = 0.1, F_u = 0.9, f_{r_1}, f_{r_2}, f_{r_3}$ 分别为随机个体的适应度值.

ICPA 流程如图 2 所示.

3 实验模型的建立

为增强风速序列的可预测性,实现风速的准确预测,建立以下风速预测模型.

3.1 实验数据

原始数据选自科罗拉多山脉的美国国家风速观

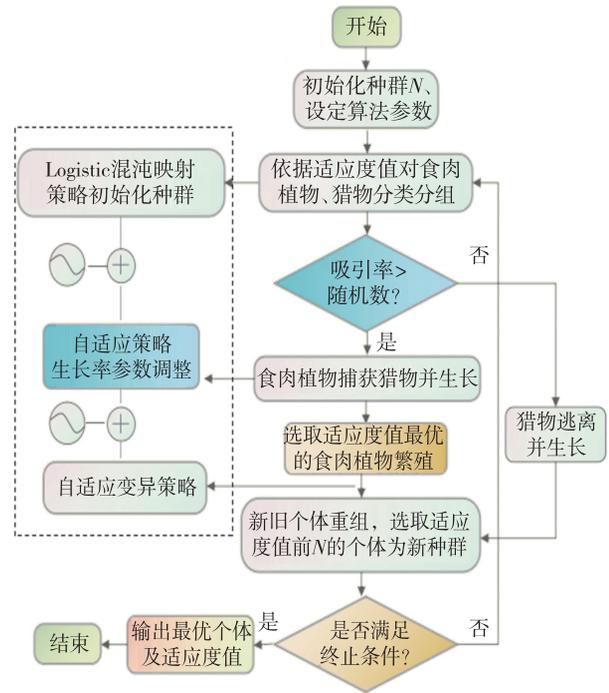


图 2 改进食肉植物算法

Fig. 2 Flow chart of improved carnivorous plant algorithm

测站 (<http://dx.doi.org/10.5439/1052222>), 包含风速数据以及温度、相对湿度、海平面压力、大气电场、平均风切变等气象因素,其中平均风切变为风矢量(风向、风速)在空中水平或垂直距离上的变化情况.采样间隔为 1 h,共采取 1 578 个样本点,部分数据如表 1 所示.

由原始风速序列图 3 可知,最大风速高达 23.268 3 m/s,最小风速仅为 0.372 3 m/s,风速变化范围较大,其无规律的变化使得风速的可预测性较低.

由表 2 中原始风速序列的统计指标值可知,平均风速为 5.153 1 m/s,从风速的方差和标准差来看,数据离散程度较大,风速的波动性与不稳定性较强.

表 1 原始风速数据

Table 1 Original wind speed data

日期	MST	风速/ (m/s)	温度/ ℃	相对湿度/ %	海平面压力/ kPa	大气电场/ (kV/m)	平均风切变/ s ⁻¹
2022-10-01	00	1.145 7	10.269 7	79.510 5	102.610 2	-0.047 4	0.005 4
2022-10-01	01	2.035 4	10.144 0	80.257 2	102.539 4	-0.056 1	0.003 8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2022-12-04	23	5.811 0	14.177 3	13.080 7	101.435 2	-0.070 6	0.064 9

注:MST为美国山地时间

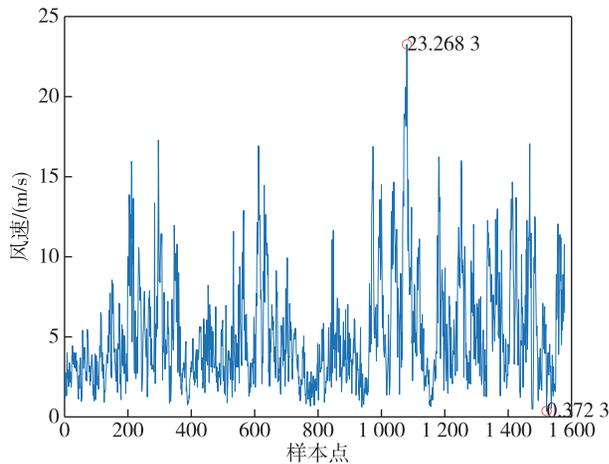


图3 原始风速序列

Fig. 3 Original wind speed sequence

表2 原始风速序列统计值

Table 2 Statistical values of original wind speed series

统计指标	值
最大值/(m/s)	23.268 3
最小值/(m/s)	0.372 3
均值/(m/s)	5.153 1
方差/(m/s) ²	11.781 4
标准差/(m/s)	3.432 4

3.2 实验模型

风速的随机性、不稳定性等特点使得风速预测难度增加,同时,风速还受气象因素的影响,因此在考虑风速本身时序特性时不能忽视气象因素的作用。

为综合考虑气象因素和时序特性的影响,本文采用特征与时间融合的预测模型对风速进行预测.具体建模过程分为以下4个部分。

1) 在第I部分中,针对气象因素的影响,首先运用交叉分组对数据集进行划分,由于温度、相对湿度、海平面压力、大气电场、平均风切变5个气象因

素与风速有着显著相关性(表3),故选取其作为LSSVM的输入(维度为5)逐一对分组后的各测试集提前1h的短期预测,再将每组预测值合并得预测风速序列。

交叉分组是指将1578个样本均分为6组,每组263个样本点,轮流将每组中263个样本作为测试集,其余部分为训练集,具体如图4所示。

2) 在第II部分中,误差序列由原始风速序列和第I部分得到的预测风速序列之差构成.第I部分和第II部分建模流程如图5所示。

3) 在第III部分中,由于误差序列是一组复杂度较高的时间序列数据,故针对时序特性的影响,对误差序列VMD分解,同时将分解后的分量作为ICPA-LSSVM模型的输入进行预测(维度为1),再对预测后序列相加重构即为误差预测值.其中,VMD的模式分量K和惩罚因子 α 由K-L散度确定。

4) 在第IV部分中,预测的误差序列用于修正原始序列的预测结果,得到最终预测的风速序列.第III部分和第IV部分建模流程如图6所示。

3.3 评价指标与统计指标

为了评价模型的预测精度,选取均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)及拟合优度决定系数(R^2)作为评价指标,公式如下:

$$\sigma_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2}, \quad (20)$$

$$\sigma_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N |y'_i - y_i|, \quad (21)$$

$$\sigma_{\text{MAPE}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - y'_i}{y_i} \right| \times 100\%, \quad (22)$$

$$R^2 = \frac{\left[\sum_{i=1}^N (y_i - \bar{y})(y'_i - \bar{y}') \right]^2}{\left[\sum_{i=1}^N (y_i - \bar{y})^2 \right] \left[\sum_{i=1}^N (y'_i - \bar{y}')^2 \right]}. \quad (23)$$

表3 风速与各气象变量间的皮尔逊相关性

Table 3 Correlation between wind speed and meteorological variables

	风速	温度	相对湿度	海平面压力	大气电场	平均风切变
风速	1	0.027 *	-0.326 **	-0.271 **	0.093 **	0.878 **
温度	0.027 *	1	0.017	-0.010	-0.008	0.141 **
相对湿度	-0.326 **	0.017	1	0.138 **	-0.228 **	-0.245 **
海平面压力	-0.271 **	-0.010	0.138 **	1	0.117 **	-0.211 **
大气电场	0.093 **	-0.008	-0.228 **	0.117 **	1	0.135 **
平均风切变	0.878 **	0.141 **	-0.245 **	-0.211 **	0.135 **	1

注: *表示显著性水平0.05(双尾)下相关性显著; **表示显著性水平0.01(双尾)下相关性显著。

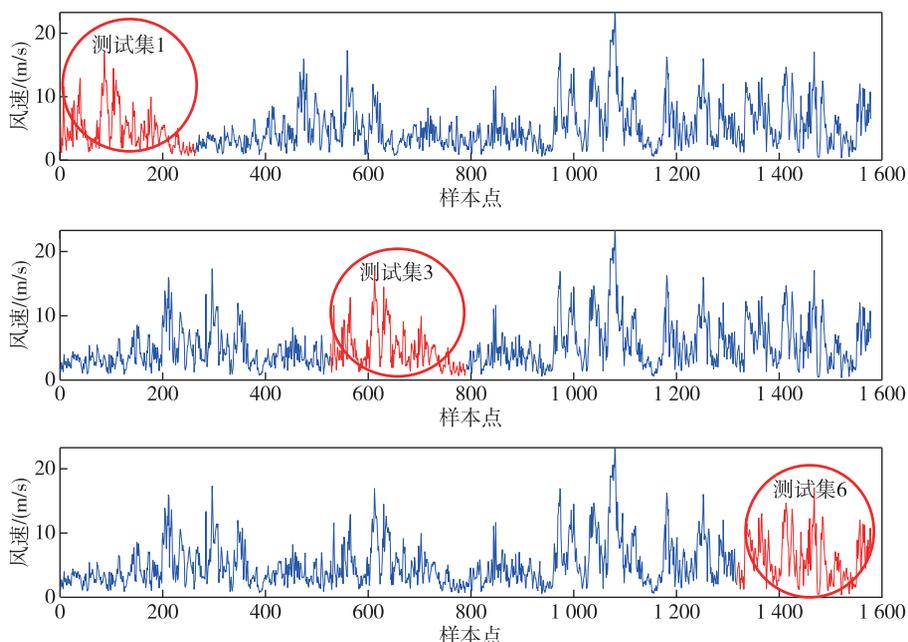


图 4 交叉分组

Fig. 4 Cross grouping diagram

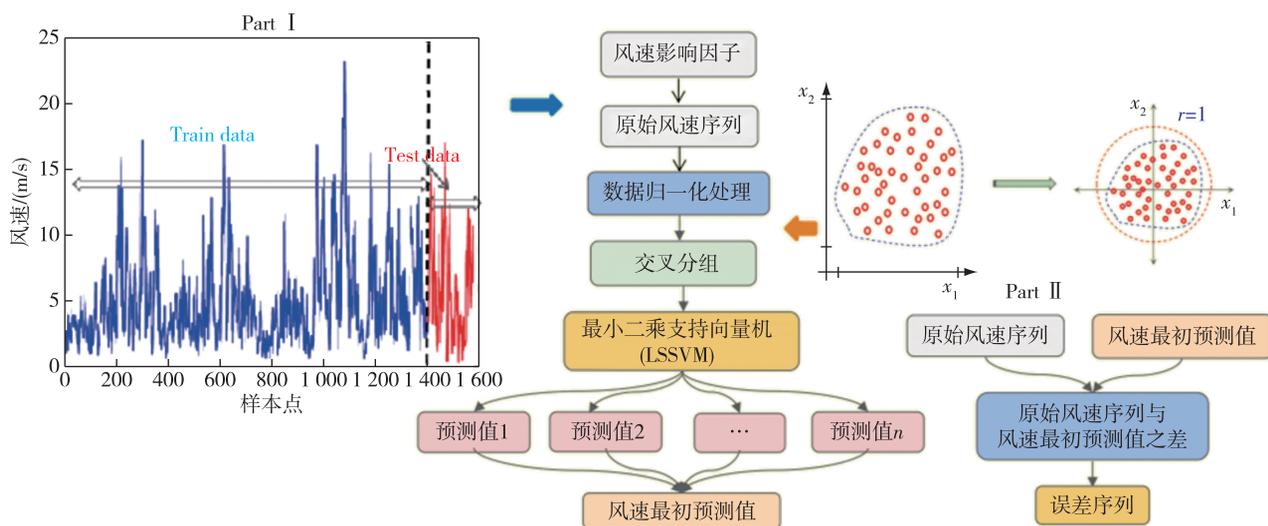


图 5 第 I 部分和第 II 部分建模流程

Fig. 5 Part I and Part II modeling flow chart

其中, N 为样本数, y_i 为实际值, y'_i 为预测值, \bar{y} 为实际平均值, \bar{y}' 为预测平均值. RMSE、MAPE、MAE 的值越小, 预测精度越好, R^2 值越大, 模型的拟合程度越高.

4 仿真结果与对比分析

4.1 基于误差修正和 VMD-ICPA-LSSVM 模型的风速预测结果分析

在 MATLAB R2021a 的环境下, 基于采集的风速实测数据, 对特征与时间融合的风速预测模型的精

度进行测试, 并分析其预测效果.

在模型第 I 部分中, 将气象因子作为 LSSVM 的输入并对各组测试集的风速预测, 其中 LSSVM 的核函数为径向基核函数, 随机设置参数 γ 和 σ .

由预测结果(图 7)可知, 将气象变量作为输入, 运用 LSSVM 预测各组风速, 预测曲线与原始风速曲线趋势大致相同, 但预测精度不高, 部分预测数据偏差大, 需进一步提高预测精度.

在模型第 I 部分只考虑了气象因素对风速预测

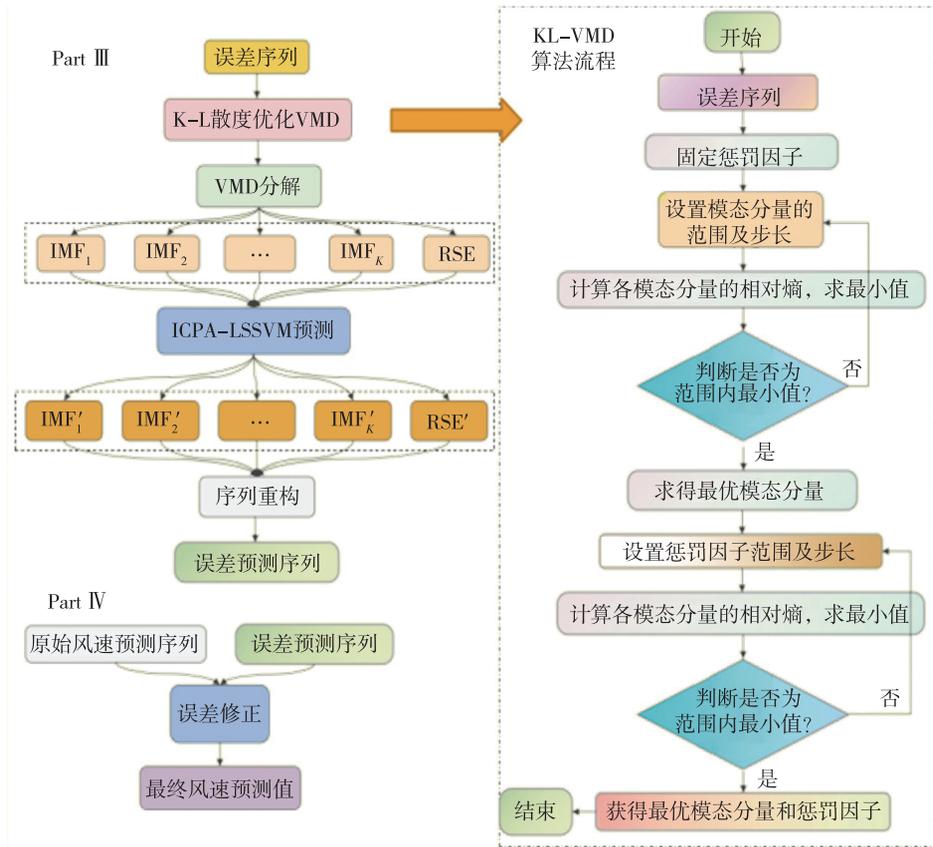


图6 第三部分和第四部分建模流程

Fig. 6 Part III and Part IV modeling flow chart

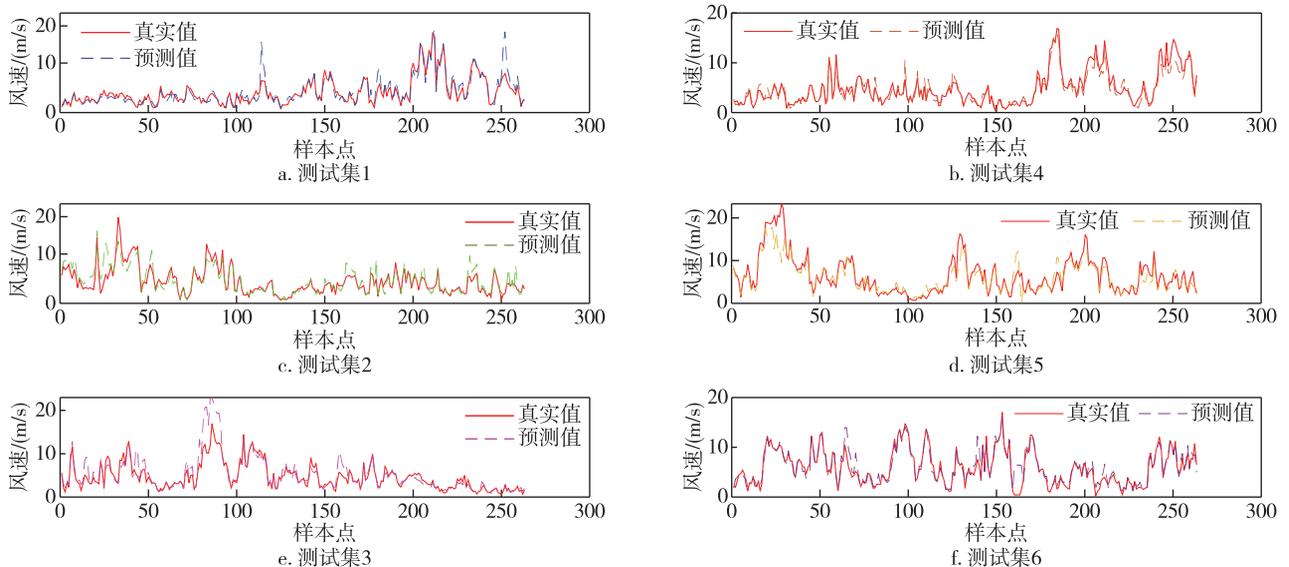


图7 交叉分组中各测试集风速LSSVM预测结果

Fig. 7 Results of wind speed predicted by LSSVM for each test set in cross-group

的影响,忽略了风速自身的随机性、波动性等特点.在模型第II部分,将综合考虑气象因素和风速自身时序特性对风速预测的影响,进一步对风速进行预

测.将原始风速序列与预测序列作差得到一组复杂度较高的误差序列(图8),可以综合体现气象因素和时序特性两方面特点.

由误差直方图 8b 可知,误差值多集中在 0 附近,但部分误差值偏离 0 较远,也可看出 LSSVM 总体预测精度不高.故可对误差序列进行预测,进而对 LSSVM 的预测值修正,提高预测精度.

在模型第 III 部分中,利用 K-L 散度对 VMD 的模态分量 K 和惩罚因子 α 进行择优选取.首先固定 $\alpha = 1\ 000$,设 K 的寻优范围为 $(2, 10)$,步长设为 1,由图 9a 可知,当 $K=3$ 时相对熵最小,故最优模态分量为 3.在 $K=3$ 的情况下,寻找最优惩罚因子 α ,设 α 的寻优范围为 $(1\ 000, 3\ 000)$,步长设为 500,由图 9b 可知,当 $\alpha = 2\ 000$ 时相对熵最小,故最优惩罚因子为 2 000.

在确定最优 K 和 α 后,对误差序列进行 VMD 分解,分解得到 3 个 IMF,1 个残差序列,分解 IMF 曲线如图 10 所示.可以看到 VMD 能准确地分离出误差序列中位于不同频段内的误差特征信号分量.

为提高模型的预测效果,选取 ICPA 算法搜索 LSSVM 模型中的参数 γ 和 σ ,分别设置 2 个参数的寻优范围为 $(0.1, 500)$ 和 $(0.01, 10)$,前 1 401 条数据为训练集,其余为测试集,将分解后每个子序列和残差序列训练集的 RMSE 作为优化目标函数进行 ICPA-LSSVM 预测(图 11),各分量 ICPA 算法优化 LSSVM 模型参数的值如表 4 所示,其中,ICPA 算法参数设置如表 5 所示.

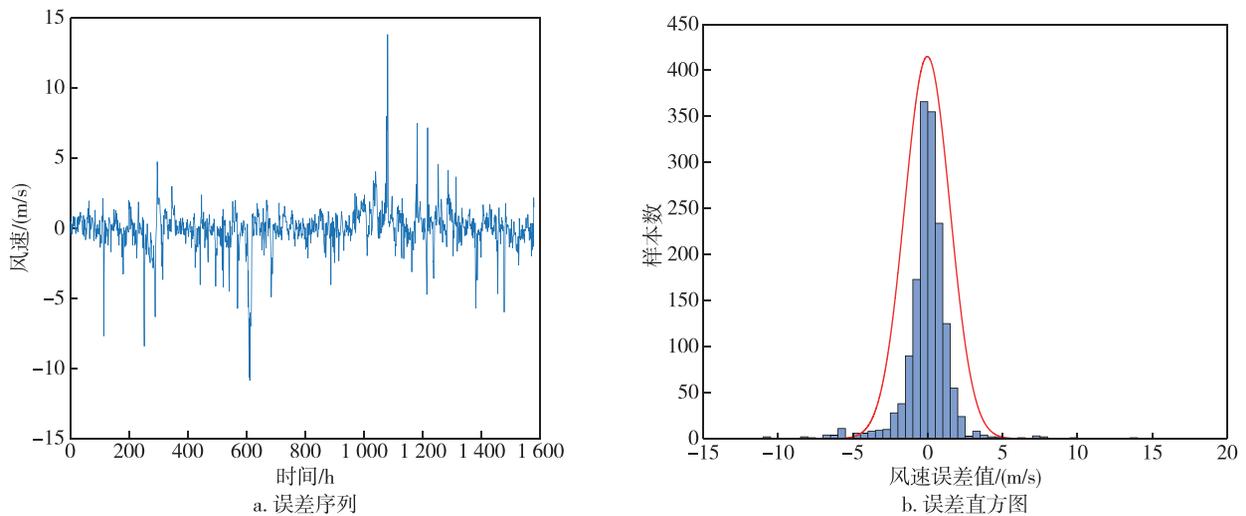


图 8 误差序列及其直方图
Fig. 8 Error sequence and its histogram

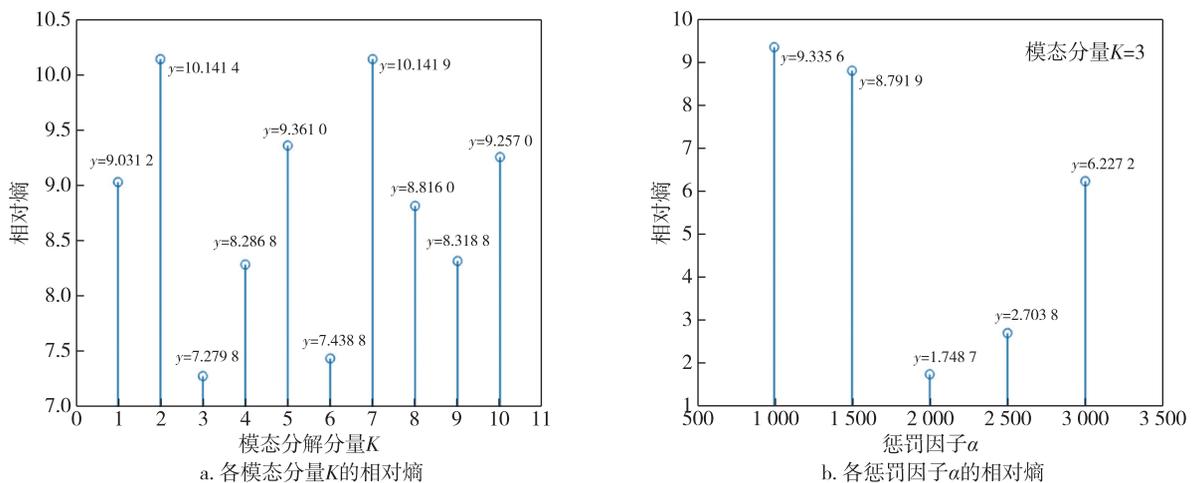


图 9 各模态分量 K 和惩罚因子 α 的相对熵
Fig. 9 Relative entropies of each modal component K and penalty factor α

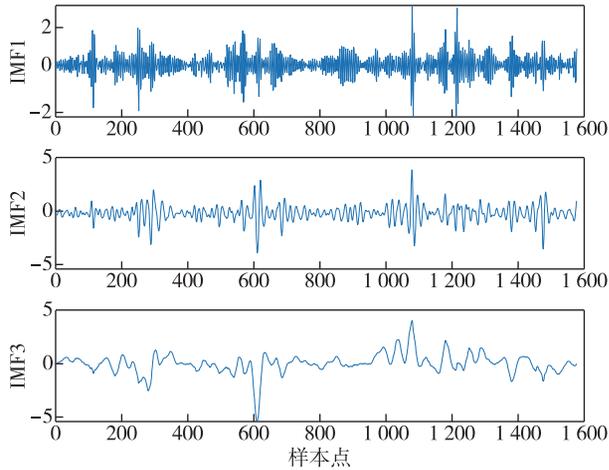


图 10 变分模态分解

Fig. 10 Variational mode decomposition results

表 4 参数寻优结果

Table 4 Parameter optimization results

分量	γ	σ
IMF1	442.652 7	0.001 010 52
IMF2	378.921 2	0.001 000 38
IMF3	185.879 5	0.001 065 29
残差	352.235 3	0.001 031 59

表 5 ICPA 算法初始参数

Table 5 Initial parameters of ICPA

参数	参数值	参数	参数值
group_iteration	2	reproduction_rate	1.8
attraction_rate	0.8	nCPlant	10
growth_rate		nPrey	20

采用改进后的 CPA 优化算法对 LSSVM 参数寻优, 其针对分解后每个子序列和残差序列的预测结果如图 11 所示, 可知改进后的 ICPA 算法在对 LSSVM 寻优后预测效果较佳, 各子序列的预测曲线与真实曲线较为吻合。

在模型第 IV 部分中, 将 ICPA-LSSVM 预测后的误差子序列相加重构, 并对第 I 部分中 LSSVM 预测的测试集数据误差修正, 得到最终风速预测值(图 12)。由图 12 可知, 基于误差修正和 VMD-ICPA-LSSVM 模型 (Error-VMD-ICPA-LSSVM) 预测曲线与原始风速曲线吻合程度较高, 预测精度较好, 误差修正也有较好的效果。

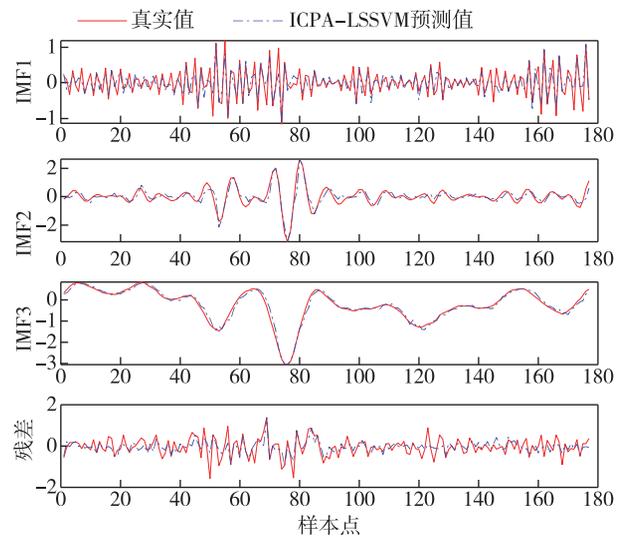
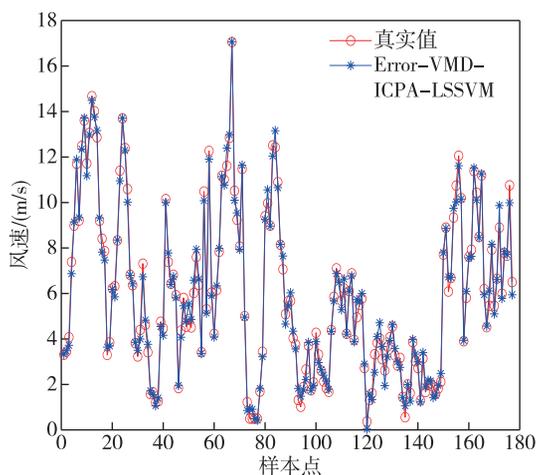
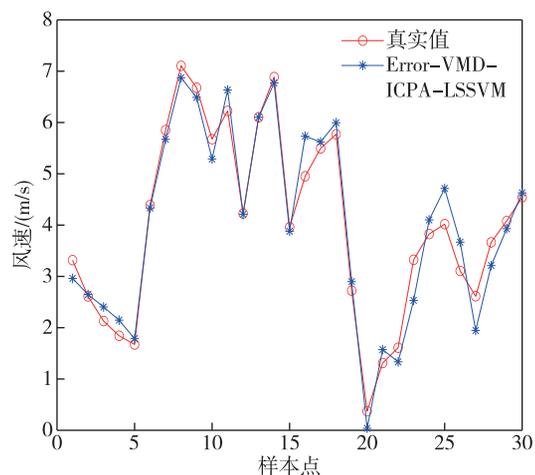


图 11 各分解序列的 ICPA-LSSVM 预测结果

Fig. 11 ICPA-LSSVM prediction results for each decomposition sequence



a. 整体预测结果



b. 部分样本点(101~130)预测结果

图 12 Error-VMD-ICPA-LSSVM 风速预测

Fig. 12 Wind speed predicted by Error-VMD-ICPA-LSSVM

4.2 模型对比分析

4.2.1 各模型风速预测曲线对比

为更好地验证模型的预测精度,将所提出的基于误差修正和 VMD-ICPA-LSSVM 风速预测模型 (Error-VMD-ICPA-LSSVM) 与 ELM、SVM、LSSVM、误差修正和 VMD-LSSVM 模型 (Error-VMD-LSSVM)、误差修正和 VMD-DE-LSSVM 模型 (Error-VMD-DE-LSSVM)、误差修正和 VMD-PSO-LSSVM 模型 (Error-VMD-PSO-LSSVM)、误差修正和 VMD-CPA-LSSVM 模型 (Error-VMD-CPA-LSSVM) 预测的精度进行比较,预测结果如图 13 所示。

由图 13 可知,相比 ELM、SVM 预测模型,LSSVM

模型预测的精度较高,预测曲线与实际风速曲线较为吻合,其均方根误差较 ELM、SVM 模型小,预测效果较好,因此,在模型实验过程中选取 LSSVM 作为预测主模型。

由图 14 可知,利用 LSSVM 直接对风速进行预测时,由于风速的随机性,预测误差相对较大,而通过 VMD 对误差进行分解后,综合气象因素和时域特征的 Error-VMD-LSSVM 模型则可增加风速预测的精度。由于 LSSVM 的参数 γ 和 σ 影响模型的精度,故未经优化的 Error-VMD-LSSVM 虽效果比 LSSVM 好,但对于部分局部风速预测效果不明显,仍存在一定的误差。

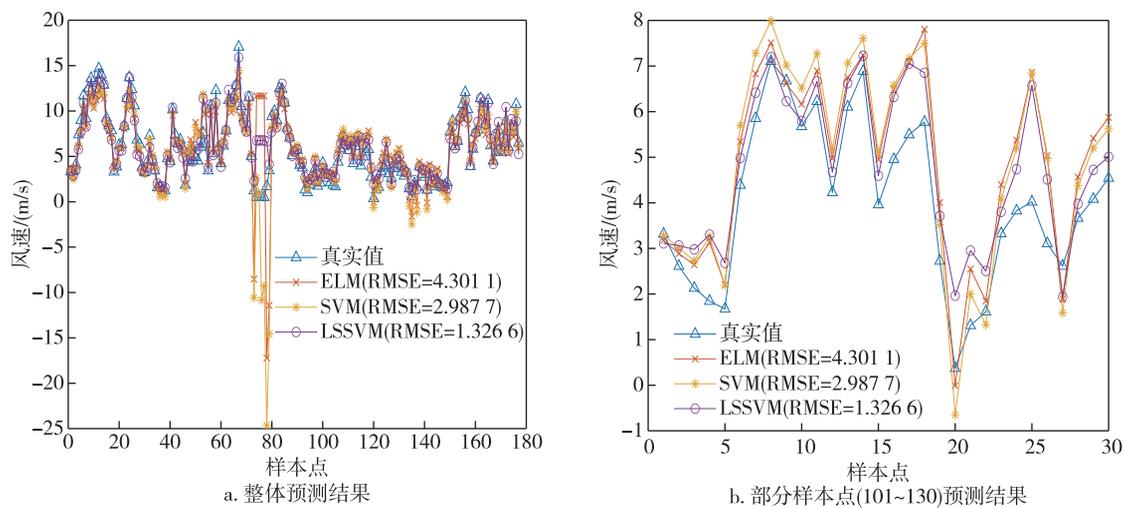


图 13 ELM、SVM、LSSVM 风速预测结果

Fig. 13 Wind speed predicted by ELM, SVM and LSSVM

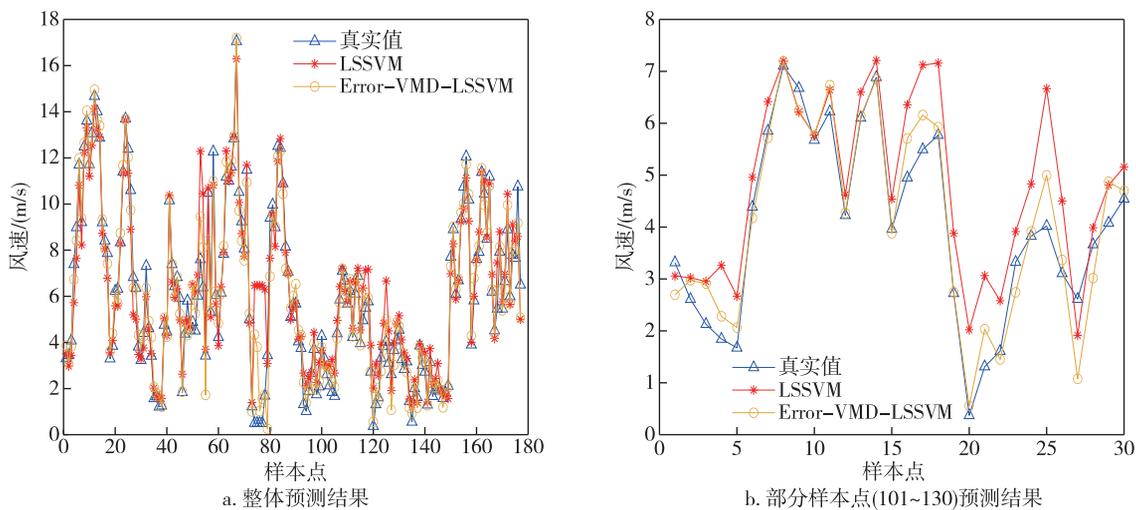


图 14 LSSVM 与 Error-VMD-LSSVM 风速预测结果

Fig. 14 Wind speed predicted by LSSVM and Error-VMD-LSSVM

由图 15 所示,通过 VMD 对误差分解后,利用 PSO、DE、CPA 算法对 LSSVM 参数优化后的模型则在预测精度上较 Error-VMD-LSSVM 有着明显的提升,其不仅可以较为准确地预测出整体风速序列的值,对于局部的风速预测也较为准确,能够有效地对原始风速序列曲线进行拟合,预测效果较好。

由图 16 可知,利用 ICPA 优化的模型在风速预测上与 CPA 优化的预测模型整体上有着相似的预测精度,预测曲线与风速实际曲线在整体上均较为吻合,但在局部风速预测上,利用 ICPA 优化的模型较 CPA 优化的模型预测精度稍高,与实际曲线更为贴近,由此可知改进后的食肉植物算法在模型优化上有一定的效果。

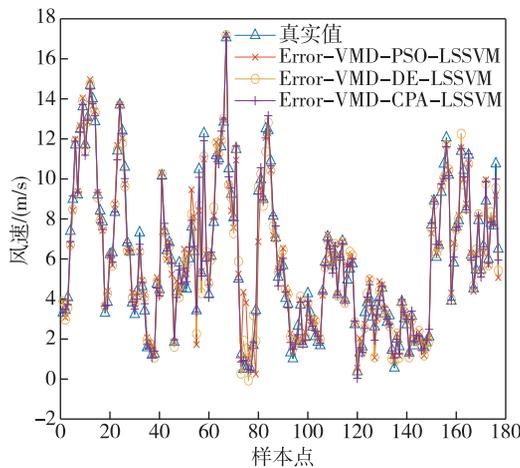
4.2.2 各模型风速预测误差分析

为更直观地比较预测效果,对各预测模型进行误差分析,评价指标值如表 6 所示。

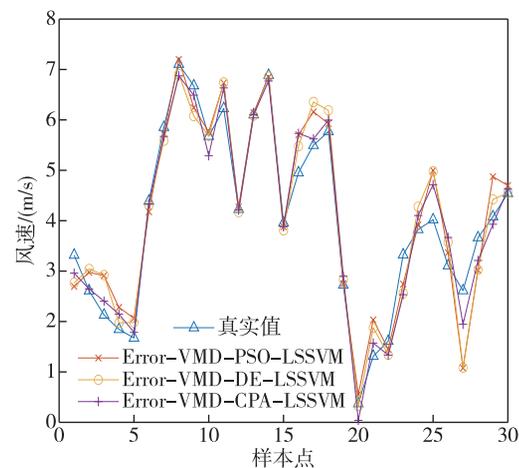
表 6 模型评价指标值

Table 6 Values of model evaluation indexes

预测模型	RMSE/ (m/s)	MAPE/ %	MAE/ (m/s)	R^2
Model1: ELM	3.040 8	93.52	1.434 6	0.599 7
Model2: SVM	2.987 7	66.76	1.222 5	0.618 2
Model3: LSSVM	1.326 6	48.09	0.787 4	0.875 8
Model4: Error-VMD-LSSVM	0.752 3	20.58	0.488 0	0.957 8
Model5: Error-VMD-DE-LSSVM	0.467 6	11.76	0.364 2	0.983 6
Model6: Error-VMD-PSO-LSSVM	0.572 0	14.29	0.442 0	0.975 5
Model7: Error-VMD-CPA-LSSVM	0.329 4	7.99	0.259 8	0.991 8
Model8: Error-VMD-ICPA-LSSVM	0.321 0	7.78	0.252 3	0.992 3



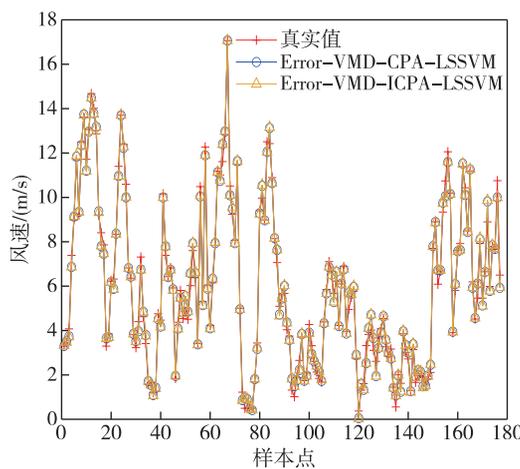
a. 整体预测结果



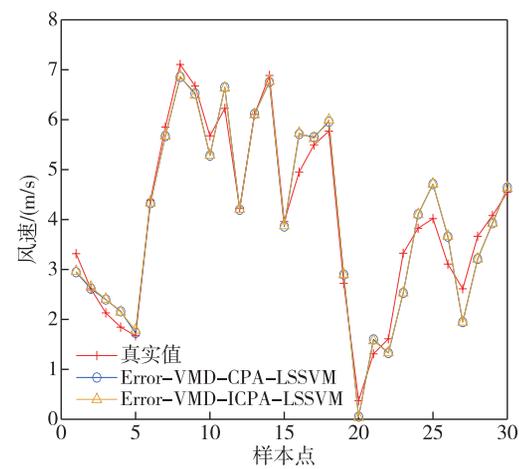
b. 部分样本点(101~130)预测结果

图 15 算法优化的风速预测结果

Fig. 15 Wind speed predicted by optimized algorithms



a. 整体预测结果



b. 部分样本点(101~130)预测结果

图 16 ICPA 与 CPA 算法优化后的风速预测结果

Fig. 16 Wind speed predicted by optimized ICPA and CPA

由表 6 可知, Error-VMD-ICPA-LSSVM 模型 (Model8) 的 RMSE、MAPE、MAE、 R^2 值分别为 0.321 0、7.78%、0.252 3 和 0.992 3,在各模型中均为最优,由此说明在综合气象因素和时序特征进行误差修正以及运用 ICPA 算法优化 LSSVM 模型的参数有着较好的效果。

从各预测环节上来说,加入了误差修正和 VMD 算法的 Error-VMD-LSSVM 模型 (Model4) 的 RMSE 和 MAPE 较 LSSVM (Model3) 分别下降了 43.29% 和 57.21%,由此也可看出综合气象因素和时序特征进行误差修正后比直接对风速进行 LSSVM 预测效果要好。加入了优化算法的 Model5、Model6、Model7 在精度上较未经过优化的 Model4 有一定的提升,说明优化算法对模型参数寻优有效,而 Model7 的 RMSE 和 MAPE 相比 Model6 下降了 42.41% 和 44.09%,则说明在风速预测时 CPA 算法相比 PSO 算法优化效果相对较好,能更大程度的提升模型预测效果。Model8 相比 Model7 的各项指标值进一步优化,由此可说明 ICPA 算法有一定的改进效果。

图 17 可更为直观地观察到各模型的标准差、均方根误差以及预测值与真实值相关系数间的关系。ELM、SVM、LSSVM 模型在泰勒图中距离参考点较远,而进行了误差修正和算法优化的模型与参考点的距离较近,预测误差也逐渐减小。

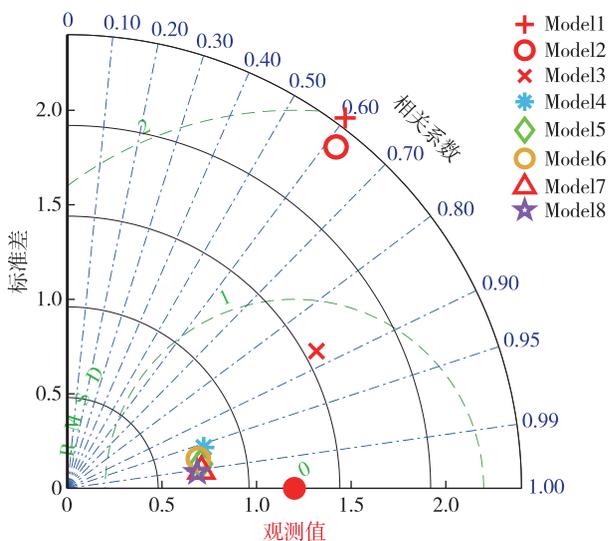


图 17 各模型标准差、均方根误差、相关系数泰勒图

Fig. 17 Taylor chart of standard deviation, root mean square error and correlation coefficient

5 结论

风速的准确预测对人们的生产生活有着重要意

义。为提高短期风速的预测精度,本文采用 Logistic 混沌映射策略、自适应参数调整策略以及引入变异策略对 CPA 优化算法进行改进,以增加其优化的效果,再综合气象因素和风速的时序特性,利用交叉分组、VMD 分解算法以及 ICPA 算法对 LSSVM 参数寻优,结合误差修正,提出一种基于误差修正和 VMD-ICPA-LSSVM 的短期风速预测模型。通过仿真实验结果和对比分析,可得出以下结论:

1) 针对气象因素对风速的影响,利用交叉分组,将气象特征作为 LSSVM 模型的输入,得到的误差时间序列有效地包含了气象因素和时序特性两方面特点,为风速的准确预测提供了基础。

2) 针对时间序列的不稳定性特点,利用 KL-VMD 对误差序列进行分解,消除模态混叠现象,可有效减小误差序列的复杂性和随机性,对分解后的序列进行预测有效地提高了模型的预测效果。

3) 寻优结果表明,在本文进行短期风速预测时,ICPA 算法的全局寻优能力优于 PSO、DE、CPA 算法。

4) 对于只考虑了气象因素影响的 LSSVM 风速预测值进行误差修正,预测效果有着明显的提高。

5) 通过各模型的对比分析,本文所提出的误差修正和 VMD-ICPA-LSSVM 风速预测模型效果较好,预测曲线的拟合度也较高。

参考文献

References

- [1] Jung C, Taubert D, Schindler D. The temporal variability of global wind energy: long-term trends and inter-annual variability [J]. Energy Conversion and Management, 2019, 188: 462-472
- [2] 王月普. 风力发电现状与发展趋势分析 [J]. 电力设备管理, 2020(11): 21-22
WANG Yuepu. Analysis of current situation and development trend of wind power generation [J]. Electric Power Equipment Management, 2020(11): 21-22
- [3] 赵林果. 大规模风电并网条件下的电力系统调度探析 [J]. 中国设备工程, 2021(4): 233-234
ZHAO Linguo. Analysis of power system scheduling under the condition of large-scale wind power grid connection [J]. China Plant Engineering, 2021(4): 233-234
- [4] Liu H, Chen C, Lv X, et al. Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods [J]. Energy Conversion and Management, 2019, 195: 328-345
- [5] Hoolohan V, Tomlin A S, Cockerill T. Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data [J]. Renewable Energy, 2018, 126: 1043-1054

- [6] Ahmed A, Khalid M. An intelligent framework for short-term multi-step wind speed forecasting based on functional networks [J]. Applied Energy, 2018, 225: 902-911
- [7] 赵征, 南宏钢, 乔锦涛. 基于二次分解的改进时间序列超短期风速预测研究 [J]. 华北电力大学学报(自然科学版), 2020, 47(4): 53-60
ZHAO Zheng, NAN Honggang, QIAO Jintao. Research on improved time series ultra-short-term wind speed prediction based on quadratic decomposition [J]. Journal of North China Electric Power University, 2020, 47(4): 53-60
- [8] 凡航, 张雪敏, 梅生伟, 等. 基于时空神经网络的风电场超短期风速预测模型 [J]. 电力系统自动化, 2021, 45(1): 28-35
FAN Hang, ZHANG Xuemin, MEI Shengwei, et al. Ultra-short-term wind speed prediction model for wind farms based on spatiotemporal neural network [J]. Automation of Electric Power Systems, 2021, 45(1): 28-35
- [9] 王顺江, 范永鑫, 潘超, 等. 基于主成分约简聚类的优化 ELM 短期风速组合预测 [J]. 太阳能学报, 2021, 42(8): 368-373
WANG Shunjiang, FAN Yongxin, PAN Chao, et al. Short-term wind speed combined forecasting based on optimized ELM of principal component reduction clustering [J]. Acta Energeiae Solaris Sinica, 2021, 42(8): 368-373
- [10] 王岩, 陈耀然, 韩兆龙, 等. 基于互信息理论与递归神经网络的短期风速预测模型 [J]. 上海交通大学学报, 2021, 55(9): 1080-1086
WANG Yan, CHEN Yaoran, HAN Zhaolong, et al. Short-term wind speed forecasting model based on mutual information and recursive neural network [J]. Journal of Shanghai Jiao Tong University, 2021, 55(9): 1080-1086
- [11] He J, Xu J L. Ultra-short-term wind speed forecasting based on support vector machine with combined kernel function and similar data [J]. EURASIP Journal on Wireless Communications and Networking, 2019, 2019(1): 1-7
- [12] 李忠, 刘景霞. 基于遗传算法和最小二乘支持向量机的风电场超短期风速预测 [J]. 电工技术, 2021(13): 56-59, 66
LI Zhong, LIU Jingxia. Ultra-short-term wind speed prediction based on GA and LSSVM in wind farm [J]. Electric Engineering, 2021(13): 56-59, 66
- [13] Tian Z D. Short-term wind speed prediction based on LMD and improved FA optimized combined kernel function LSSVM [J]. Engineering Applications of Artificial Intelligence, 2020, 91: 103573
- [14] 王玲, 程耕国, 袁志强, 等. 关于风力发电系统短期风速预测仿真 [J]. 计算机仿真, 2018, 35(11): 98-102
WANG Ling, CHENG Gengguo, YUAN Zhiqiang, et al. Simulation of short-term wind speed forecasting for wind power generation system [J]. Computer Simulation, 2018, 35(11): 98-102
- [15] Liu H, Duan Z, Han F Z, et al. Big multi-step wind speed forecasting model based on secondary decomposition, ensemble method and error correction algorithm [J]. Energy Conversion and Management, 2018, 156: 525-541
- [16] 赵征, 汪向硕, 乔锦涛. 基于 VMD 和改进 ARIMA 模型的超短期风速预测 [J]. 华北电力大学学报(自然科学版), 2019, 46(1): 54-59
ZHAO Zheng, WANG Xiangshuo, QIAO Jintao. Ultra-short-term wind speed prediction based on VMD and improved ARIMA model [J]. Journal of North China Electric Power University (Natural Science Edition), 2019, 46(1): 54-59
- [17] Xu H Y, Chang Y Q, Zhao Y, et al. A hybrid model for multi-step wind speed forecasting based on secondary decomposition, deep learning, and error correction algorithms [J]. Journal of Intelligent & Fuzzy Systems, 2021, 41(2): 3443-3462
- [18] 桑茂景, 谢丽蓉, 李进卫, 等. 基于 LMD-IMVO-LSSVM 的短期风速预测 [J]. 可再生能源, 2021, 39(9): 1197-1203
SANG Maojing, XIE Lirong, LI Jinwei, et al. Short-term wind speed prediction based on LMD-IMVO-LSSVM [J]. Renewable Energy Resources, 2021, 39(9): 1197-1203
- [19] Ma Z R, Chen H W, Wang J J, et al. Application of hybrid model based on double decomposition, error correction and deep learning in short-term wind speed prediction [J]. Energy Conversion and Management, 2020, 205: 112345
- [20] 摆玉龙, 路亚妮, 刘名得. 基于变分模态分解的机器学习模型择优风速预测系统 [J]. 地球科学进展, 2021, 36(9): 937-949
BAI Yulong, LU Yani, LIU Mingde. Wind speed forecasting system based on variational mode decomposition and the optimal machine learning models [J]. Advances in Earth Science, 2021, 36(9): 937-949
- [21] 尹佳璠, 陈小奇, 李世林, 等. 优化 VMD 与 NLM 结合的信号去噪 [J]. 计算机工程与设计, 2021, 42(4): 1135-1142
YIN Jiafan, CHEN Xiaoqi, LI Shilin, et al. Signal denoising based on optimized VMD and NLM [J]. Computer Engineering and Design, 2021, 42(4): 1135-1142
- [22] 张炎亮, 毛贺年, 赵华东. 基于 VMD 与 IFWA-SVM 的滚动轴承故障诊断研究 [J]. 机床与液压, 2022, 50(6): 180-185
ZHANG Yanliang, MAO Henian, ZHAO Huadong. Research on rolling bearing fault diagnosis based on VMD and IFWA-SVM [J]. Machine Tool & Hydraulics, 2022, 50(6): 180-185
- [23] 李文武, 石强, 王凯, 等. 基于变分模态分解和深度门控网络的径流预测 [J]. 水力发电学报, 2020, 39(3): 34-44
LI Wenwu, SHI Qiang, WANG Kai, et al. Runoff prediction based on variational mode decomposition and deep gated network [J]. Journal of Hydroelectric Engineering, 2020, 39(3): 34-44
- [24] Ong K M, Ong P, Sia C K. A carnivorous plant algorithm for solving global optimization problems [J]. Applied Soft Computing, 2021, 98: 106833

Short term wind speed prediction modeling based on error correction and VMD-ICPA-LSSVM

ZHONG Lin¹ YAN Qisheng¹

¹ School of Science, East China University of Technology, Nanchang 330013, China

Abstract Accurate wind speed prediction is the key to large-scale application of wind energy in power system, but the randomness and volatility of wind speed sequence make it difficult to predict. Herein, strategies of Logistic chaotic mapping, adaptive parameter adjustment, and the introduction of mutation are used to improve the Carnivorous Plant Algorithm (CPA), and a short-term wind speed prediction model based on error correction and VMD-ICPA-LSSVM is proposed. First, meteorological factors are used as inputs for Least Squares Support Vector Machine (LSSVM) to predict wind speed and obtain an error sequence. Then, K-L divergence is used to adaptively determine the parameters of Variational Mode Decomposition (VMD) and decompose the error sequence. Then the Improved Carnivorous Plant Algorithm (ICPA) is combined to optimize the adjustable parameters of LSSVM to predict the decomposed subsequences. The prediction results of each subsequence are stacked and error correction is performed on the original prediction sequence to obtain the final wind speed prediction values. The experimental results show that the proposed model has excellent prediction accuracy and generalization performance.

Key words variational mode decomposition (VMD); carnivorous plant algorithms (CPA); least squares support vector machine (LSSVM); error correction; wind speed prediction