

基于 EEMD 和 LS-SVM 模型的风电功率短期预测方法

程启明¹, 陈路¹, 程尹曼², 张强¹, 高杰¹

(1. 上海电力学院 自动化工程学院 上海市电站自动化技术重点实验室, 上海 200090;

2. 上海电力公司 市北供电分公司, 上海 200041)

摘要:原始风速信号具有的间歇波动性特征给风电场的功率预测带来了挑战,采用集合经验模态分解(EEMD)法将原始风速信号分解为频域稳定的子序列,有效地提高了预测精度,避免了传统经验模态分解(EMD)存在的模态混叠现象。提出一种改进型果蝇优化算法(FOA),将风速子序列重构参数和最小二乘支持向量机(LS-SVM)参数作为优化目标建立风速预测模型,扩大了参数搜索范围,提高了优化收敛速度;通过风速-风功率转化关系可以求得风电场的功率值。实验结果验证了所提方法相比于 EMD 和 LS-SVM 预测方法具有更高的预测精度。

关键词:微电网;功率预测;风电场;模态分解;支持向量机;相空间重构;果蝇优化算法

中图分类号:TM 615

文献标识码:A

DOI:10.16081/j.issn.1006-6047.2018.05.004

0 引言

风功率预测是评估风电场运行状态的基础,其随机波动特性给电网安全运行带来了挑战^[1]。为了保证电力系统安全稳定可靠运行,缓解电力系统调峰、调频压力,必须提高风电功率预测的精度^[2]。对于风电功率预测,主要是基于风电场的历史数据^[3]、数值天气预报 NWP (Numerical Weather Prediction)^[4]、地理位置和气象环境因素^[5]、风速-风功率转化特性^[6],结合物理、统计及组合等预测模型,实现多时间尺度的预测^[7]。采用物理模型必须依靠成熟的 NWP 体系,量化分析处理风电场风速、风向、温度、大气湿度和气压等历史气象数据,并将其代入功率曲线求出实际功率^[8],此方法受限于 NWP 更新速度,适合风电功率的中短期预测,因此在风机检修、调试等场合较为常见。统计回归和学习理论均基于历史统计数据、实时监测值和输出功率之间的映射关系,具有较强的泛化能力,无需考虑风机周边的具象特性等优点,广泛应用于风电功率的短期预测。然而,由于统计回归法对历史数据的依赖性,该方法不适合小数据样本的情况,面对复杂多变的气象环境,其适用范围有待提高^[9-10]。统计学习法基于机器学习的思想,弥补了统计法的不足,提高了模型的

灵活性^[11]。目前研究的重点是风电功率对风速、风向的波动响应,对于风速的研究是今后发展的重点^[12-13]。文献[14]将风序列相空间重构参数和支持向量机参数整合为一个隶属度指标,提高了风速的回归精度,对风功率预测模型的参数优化进行了理论补充;文献[15]基于相关性分析修正了经验变异函数,得到整个风电场群内的风速分布;文献[16]考虑爬坡特性,根据不同功率分区内风电爬坡率和误差二维核密度估计概率预测模型,基于小波神经网络建立风电功率预测模型,有效提高了精度;文献[17]提出区间预测的概念,采用粒子群-极限学习机(PSO-KELM)建立风功率预测模型,实现了对风电功率的快速区间预测,改善了学习速度和泛化能力,但粒子群优化 PSO (Particle Swarm Optimization) 算法的迭代收敛速度仍有待提高;文献[18]通过采用经验模态分解 EMD (Empirical Mode Decomposition) 实现原始信号的分解,然而,该方法在波动范围较大时可能会出现模态混叠,影响预测的精度。

风力发电短期功率预测主要针对预测时间在 72 h 以内的功率预测,本文关于风功率预测的研究限定在未来 24 h,基于 NWP 和气象监测站提供的数据,建立风电功率短期预测模型,提出了改进型果蝇优化算法 FOA (Fruit fly Optimization Algorithm),有效确定可调参数,提高了收敛速度,扩大了搜索范围;采用集合经验模态分解 EEMD (Ensemble EMD)^[19],避免了波形分解过程中可能出现的模态混叠现象,提高了模型预测的健壮性,改善了风电功率预测的精度。

1 历史气象数据采集与优化算法的改进

本文实验数据取自微网实验平台,以数据采集与监控 SCADA (Supervisory Control And Data Acquisition) 系统提供的风机在 2016 年 1 月 27 日至 3 月 9

收稿日期:2017-09-04;**修回日期:**2018-02-19

基金项目:国家自然科学基金资助项目(61573239);上海市重点科技攻关计划项目(14110500700);上海市电站自动化技术重点实验室项目(13DZ2273800);上海市自然科学基金资助项目(15ZR1418600)

Project supported by the National Natural Science Foundation of China (61573239), Key Science and Technology Plan of Shanghai Science and Technology Commission (14110500700), Shanghai Power Station Automation Technology Key Laboratory Program (13DZ2273800) and the Natural Science Foundation of Shanghai (15ZR1418600)

日之间运行的历史数据作为实验数据。该平台风力发电系统模拟风机容量为 1.5 MW, 提供每个预测日 288 个预测数据点, 包括风速、风向、温度、大气压力、相对湿度等特征数据, 采样数据与预测数据的分辨率均为 5 min。

1.1 风电场历史数据处理

实际运行过程中会存在一些异常数据, 严重影响风电预测模型的精度。因此, 需要对原始数据进行预处理, 将缺失数据补充完整, 异常数据合理替换, 以保证能够符合下列预测模型的运行要求:

a. 当输出功率值过大, 超过了风电机组的装机容量时, 可用装机容量额定值来取代;

b. 当风速值小于切入风速时, 现实情况下风机输出功率应为 0, 但实际记录数据不为 0, 设输出功率为 0, 将所有风速值为负的值置为 0;

c. 当存在缺失数据时, 采用线性插值法补充完整。

1.2 FOA

FOA^[20] 是一种全局搜索学习算法, 与遗传算法 GA (Genetic Algorithm) 和 PSO 算法相比, FOA 具有可调参数少、搜索时间短、学习速度快等优点。FOA 是根据果蝇觅食行为推演出的一种全局寻优算法, 其算法寻优过程如图 1 所示。

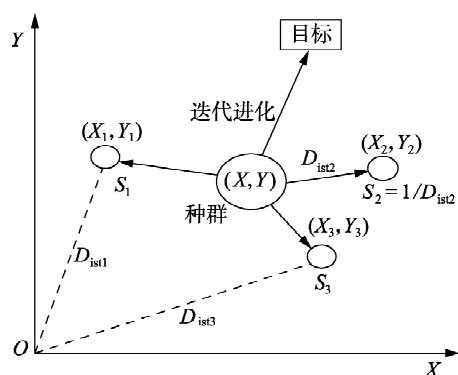


图 1 FOA 迭代觅食过程

Fig.1 Iterative foraging process of FOA

由图 1 可见, 首先在二维空间内定义果蝇种群搜寻目标的随机方向:

$$X_i = x_i + \Delta\delta_x, Y_i = y_i + \Delta\delta_y \quad i = 1, 2, \dots, n \quad (1)$$

其中, $\Delta\delta_x, \Delta\delta_y \in \mathbf{R}$, 为 $-1 \sim 1$ 之间任意随机数; (x_i, y_i) 为随机初始果蝇群体的坐标; n 为种群规模。味道浓度判断值 S_i 为:

$$S_i = 1/D_{isti} = 1/\sqrt{X_i^2 + Y_i^2} \quad i = 1, 2, \dots, n \quad (2)$$

其中, $D_{isti} = \sqrt{X_i^2 + Y_i^2}$ 为与原点的距离。

将 S_i 代入适应度函数得到该果蝇个体坐标的味道浓度值 S_{melli} 为:

$$S_{melli} = f(S_i) \quad i = 1, 2, \dots, n \quad (3)$$

最后求出种群中最大值 $S_{mellmax}$ 和该值所对应的个体 $(x_{smellmax}, y_{smellmax})$, 该个体即为所求最佳路线目标值, 将其代入式 (2) 中确定搜索方向, 最终实现寻优迭代的学习过程。

1.3 改进 FOA

由 FOA 原理可知, 由于 $S_i > 0$, 导致算法只能在坐标的第 1、4 象限搜寻最优值, 若所求问题的适应度函数极值在第 2、3 象限, 就会陷入局部极值, 因此需对果蝇个体坐标的味道浓度判定值 S_i 进行调整, 以提高搜索范围。

a. 首先在计算适应度值时, 可将 D_{isti} 改为:

$$D_{isti} = \text{sign}(\Delta\delta) \cdot \sqrt{X_i^2 + Y_i^2} \quad i = 1, 2, \dots, n \quad (4)$$

其中, $\Delta\delta \in \mathbf{R}$, 为 $-1 \sim 1$ 之间任意随机数。

$\text{sign}(\Delta\delta)$ 确保了 D_{isti} 符号的随机多变性, 避免了陷入第 1、4 象限寻优的可能, 使得算法在寻优过程中跳出极值。

b. 然后在迭代寻优时, 将式 (2) 中的 S_i 更新迭代为:

$$S'_i = \text{sign}(D_{isti}) / \sqrt{X_i^2 + Y_i^2} \quad i = 1, 2, \dots, n \quad (5)$$

其中, S'_i 为每次迭代更新后的浓度判定值。这样就保证了更新后 S'_i 的符号 (果蝇搜寻目标的方向) 与原浓度值 S_i 保持一致, 使得个体在寻优过程中缩短搜索路径, 加快搜索时间, 达到全局搜索的目的。

2 EMD 方法原理及其改进

风速、风功率信号具有不平稳、波动性大的特点, 在风电功率预测时往往会造成无规律变动, 导致出现功率序列响应能力不足的问题。为此, 本文基于信号尺度分解降维的思想, 采用 EEMD 的方法建立预测模型, 有效地处理了非线性的时序风电信号, 提高了功率预测的精度。

2.1 EMD 原理及依据

EMD 是由 Huang 提出的一种针对信号分析的自适应数据挖掘方法^[15]。它通过将非线性序列分解为若干不同尺度的本征模态函数 IMF (Intrinsic Mode Function) 分量和 1 个剩余分量, 获得平稳序列, 在分解过程中基于信号的自身尺度, 非常适合用来处理波动性较大的风速波形, 理论上可以应用于任何类型的时间序列信号。其中 IMF 分量需符合以下 2 个条件:

a. 信号过零点个数与局部极值点数至多相差 1 个;

b. 整个定义域范围内的序列均值趋于 0。

EMD 具体分解过程如下。

a. 求出原始信号 $X(t)$ 中所有极值点 (包括极大值和极小值), 采用 3 次样条差值函数拟合出原数据的上、下包络线 $l_1(t)$ 、 $l_2(t)$ 。

b. 求出上、下包络线的中位值:

$$m_1(t) = [l_1(t) + l_2(t)] / 2 \quad (6)$$

c. 令 $h_1(t) = X(t) - m_1(t)$, 若 $h_1(t)$ 不满足 IMF 分量的 2 个充分条件, 则继续重复步骤 a、b, 直至 k 次迭代后求出 $h_{1k}(t)$ 满足这 2 个条件, 即可得 $C_1(t) = h_{1k}(t)$ 。

d. 将 IMF1 分离出原始信号, 以剩余分量 $r_1(t) = X(t) - C_1(t)$ 作为原始信号, 再重复上述步骤, 重新分解序列信号, 得到 n 个 IMF 分量, 当剩余分量 $r_n(t)$ 满足单调性时即为最终结果, 分解后信号为:

$$x(t) = \sum_{i=1}^n C_i(t) + r_n(t) \quad (7)$$

其中, $C_i(t)$ 为 IMF 分量; $r_n(t)$ 为剩余分量。

文献[21]采用相空间重构计及风电序列的 Lyapunov 指数, 验证了风电时间序列的混沌特性。通过模态分解原始风速信号, 得到一系列有限频带宽度的子序列, 能够降低波动性带来的预测误差。

2.2 EEMD

风电数据的监测和采集往往会发生信号间断、噪声和设备故障引起异常脉冲干扰等现象, 在模态分解过程中会导致错误的 IMF 分量, 极易出现模态混叠现象, 从而不能实现较为理想的效果。目前解决该现象较好的方法是 EEMD。该方法改进过程如下:

a. 首先在原始风电功率信号中加入服从正态分布的白噪声序列, 构成新的目标序列, 定义 α 为噪声幅值, ε 为标准差;

b. 对该目标序列进行 EMD, 求出 n 个 IMF 分量 $C_i(t)$ 和 1 个剩余的残差分量 $r_n(t)$;

c. 将步骤 a、b 循环迭代 r 次, 每次采用不同幅值的白噪声序列, 将 r 次分解得到的 IMF 分量求整体平均, 作为原始时序信号的 IMF 分量。

3 基于 EEMD 的短期风功率预测模型

风速风功率数据的特征表现为波动范围大, 其非平稳性严重制约功率预测模型的建立, 并对精度的提高有着很大的影响。本文搭建的风电短期功率预测模型基于 EEMD 多尺度分解法, 保留了原始数据的性质, 有效解决预测波形的波动大问题。图 2 为功率预测模型的整体框架图。

具体建模步骤如下:

a. 将从 SCADA 系统中得到的原始风速序列经 EEMD, 得到 n 个 IMF 分量;

b. 利用相空间重构法, 分别确定 IMF 分量的延迟时间 τ 和嵌入维数 m ;

c. 根据不同的重构 IMF 分量分别建立最小二乘支持向量机 LS-SVM (Least Square Support Vector Machine) 预测模型, 预测风速子序列;

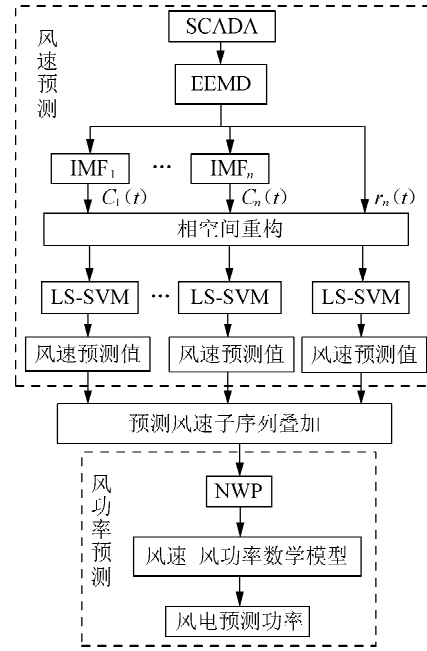


图 2 基于 EEMD 风电短期功率预测模型

Fig.2 Short-term wind power forecasting model based on EEMD

d. 叠加不同尺度下的风速预测分量, 求出原始风速预测值;

e. 结合 NWP, 将风速预测值代入风速-风功率数学模型, 求出最终的风功率预测值。

3.1 EEMD 参数的优化

EEMD 在 EMD 时序信号波形中加入了服从正态分布的白噪声序列, 有效地抑制了模态混淆。但白噪声幅值系数 α 和迭代次数 r 的选取对于模态分解的整体效果具有重要的物理意义。由 $e_r = \alpha / \sqrt{r}$ 可得两参数之间的关系, 其中 e_r 为加入白噪声时造成的总体分解误差, 即为:

$$e_r = \sqrt{\frac{1}{r} \sum_{j=1}^r \left[\left(\sum_{i=1}^n C_i^j(t) + r_n^j(t) \right) - x(t) \right]^2} \quad (8)$$

考虑两参数之间平衡关系: 若 α 值过大, 会导致模态分解后出现较大的误差, 严重时覆盖原始信号的幅频特性, 失去分解的目的; 若 α 过小, 虽然有助于提高分类的精度, 但会得到无法改变原始信号的局部极值点, 从而不能获取原始信号的信息。从运算时间成本考虑, r 值并非越大越好, 会增加计算成本。

由于在 $e_r \leq 0.01$ 时, 残留噪声引起的分解误差能够达到较为理想的效果, 序列重构处于稳定的状态, 故本文可采用改进 FOA 优化算法确定 EEMD 参数, 其适应度函数由参数 α 和 r 确定, 评价指标为总体分解误差 e_r 。对两参数进行优化选取时, 首先限制定义域, 设噪声幅值 $\alpha \in [0.1, 0.3]$, 总体平均次数 $r \geq 100$, 然后由 FOA 寻优最终得两参数优化值分别为 $\alpha = 0.18$ 、 $r = 200$ 。

3.2 风速序列模态分解

本文选取2016年2月的数据,结合NWP发布的气象预报数据搭建功率预测模型,时间跨度为10日,采样分辨率为5 min,共计2 880个采样点。前9日样本作为历史数据用来训练,第10日功率样本值作为预测日测试数据,图3为系统记录的原始风速序列曲线图。

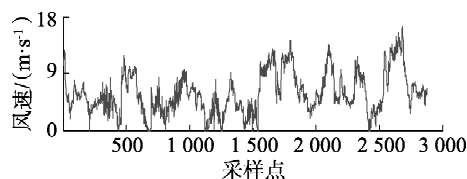


图3 风电场10日风速时间序列

Fig.3 Wind speed time series in wind farm for 10 days

首先预处理风速数据,风速值为负的数据设为0,大于切出风速值设为等于切出风速。规定前2 592个数据样本点为训练集,后288个数据样本点为测试集,采用EEMD将SCADA系统记录的原始风速时间序列进行模态分解,代入第3.1节优化得到的参数值 $\alpha=0.18$ 、 $r=200$ 。图4为最终得到的9个

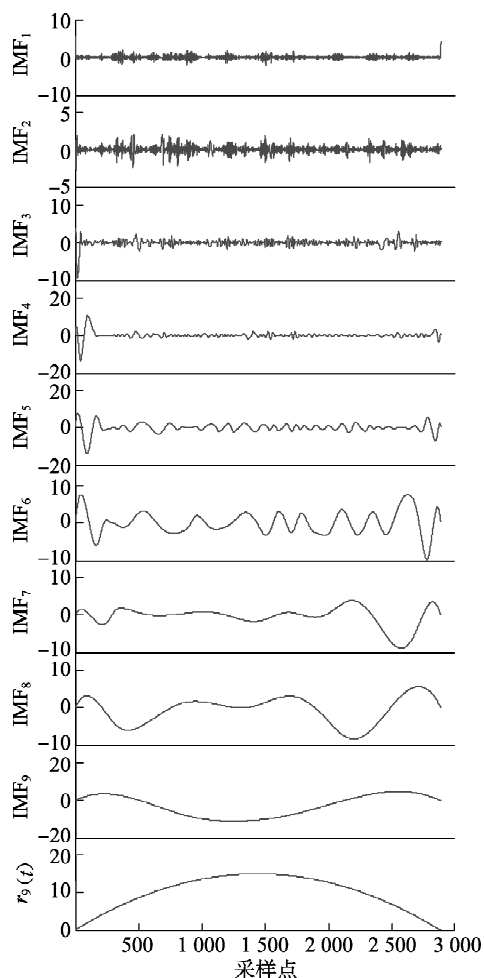


图4 风速时间序列模态分解结果

Fig.4 Results of modal decomposition for wind speed time series

本征模态函数分量 IMF_1 — IMF_9 和1个剩余分量 $r_9(t)$ 的波形图。

由图4可见,分解后的IMF分量与图3所示的原始风速序列相比,其波动变化较为平稳,频谱特征也由IMF分量从高频到低频依次表征出来。

3.3 风速序列的相空间重构

经EEMD得到较为平稳的子序列后,其稳定性得到增强。由于风速信号不仅对于原始条件具有敏感性,而且存在非周期的运动,故可采用混沌理论来分析风速序列的内在属性,即延迟坐标状态空间重构法^[14]。

定义时间序列 $\{x(i):1,\dots,N\}$,设嵌入维度为 $m(m \geq 2d+1, d$ 为系统动力学维数),并定义延迟时间 τ 为相空间采样间隔,对于 $1 \leq i \leq N+(m-1)\tau$,得到重构的相空间 $\{\mathbf{x}_i\}_{i=1}^N$ 为:

$$\begin{cases} \mathbf{x}_1 = [x(1), x(1+\tau), \dots, x(1+(m-1)\tau)] \\ \mathbf{x}_2 = [x(2), x(2+\tau), \dots, x(2+(m-1)\tau)] \\ \vdots \\ \mathbf{x}_N = [x(N), x(N+\tau), \dots, x(N+(m-1)\tau)] \end{cases} \quad (9)$$

相空间重构将原序列扩展至 N 个样本,而时间延迟 τ 和嵌入维度 m 的确定直接影响重构后信息的获取。若维数 m 过低,则会导致信息重合;若 m 过高则增加计算量。若延迟时间 τ 过短,则空间坐标相关性过于密集,无法形成较高的辨识度,从而不能提供独立的坐标分量;若 τ 太长,混沌吸引子轨迹在两分量所指方向的投影就失去了相关性。因此需要确定合适的参数,从而在独立和相关两者之间达到一种平衡。

3.4 改进FOA的参数优化

由于SCADA采集的风速数据样本时间分辨率为5 min,所以最小计算单位的延迟时间 τ 为5 min。同时考虑LS-SVM模型的2个参数:正则化参数 γ 和核函数参数 σ ,将每个子序列的预测值误差作为最终优化目标函数值。

本文采用LS-SVM建立风速预测模型。LS-SVM将线性最小二乘系统代替二次规划作为损失函数,用等式约束取代标准支持向量机的不等式约束,有效提高了计算时间,减少了计算量。考虑训练样本 $S = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N)\} \in \mathbf{R}^n \times \mathbf{R}$,其中 $\mathbf{x}_i$ 为第 i 个样本, $y_i \in \{-1, +1\}$,支持向量机的不等式约束为:

$$\begin{aligned} \min_{\omega, b, \xi} J(\omega, e) &= \omega^T \omega / 2 + C \sum_{i=1}^N \xi_i \\ \text{s.t. } y_i(\omega^T \varphi(\mathbf{x}_i) + b) &\geq 1 - \xi_i, \xi_i \geq 0 \quad i=1, 2, \dots, N \end{aligned} \quad (10)$$

其中, ω 为超平面的法向量; C 为惩罚参数; ξ_i 为松弛变量; b 为支持向量机模型参数; N 为训练样本长

度; $\varphi(\mathbf{x}_i)$ 为训练样本的非线性变换。

将其变为等式约束,即为:

$$\begin{aligned} \min_{\omega, b, e} J(\omega, e) &= \omega^T \omega / 2 + \frac{C_s}{2} \sum_{i=1}^N e_i^2 \\ \text{s.t. } y_i [\omega^T \varphi(\mathbf{x}_i) + b] &= 1 - e_i \quad i = 1, 2, \dots, N \end{aligned} \quad (11)$$

其中,权重常数 C_s 与惩罚参数 C 意义相同,用于平衡寻找最优超平面,最小化偏差量; $e = [e_1, e_2, \dots, e_N]$ 为误差向量。

采用 Lagrange 乘数法将原问题转化为对参数 a 求极值的问题:

$$L(\omega, b, e, a) = J(\omega, e) - \sum_{i=1}^N a_i \{y_i [\omega^T \varphi(\mathbf{x}_i) + b] - 1 + e_i\} \quad (12)$$

对参数 ω, b, e_i, a_i 求导得:

$$\begin{cases} \partial L / \partial \omega = 0 \Leftrightarrow \omega = \sum_{i=1}^N a_i y_i \varphi(\mathbf{x}_i) \\ \partial L / \partial b = 0 \Leftrightarrow \sum_{i=1}^N a_i y_i = 0 \\ \partial L / \partial e_i = 0 \Leftrightarrow a_i = C_s e_i \\ \partial L / \partial a_i = 0 \Leftrightarrow y_i [\omega^T \varphi(\mathbf{x}_i) + b] - 1 + e_i = 0 \end{cases} \quad (13)$$

其中, $i = 1, 2, \dots, N$ 。

核函数 $K(\mathbf{x}_i, \mathbf{x}_j) = f(\mathbf{x}_i, \mathbf{x}_j)$ 是满足 Mercer 条件的对称函数。由上式可列出线性方程组为:

$$\begin{bmatrix} 0 & \mathbf{y}^T \\ \mathbf{y} & \mathbf{K} + \mathbf{I}/C_s \end{bmatrix} \begin{bmatrix} b \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \quad (14)$$

其中,核矩阵 $\mathbf{K}$ 和 $\mathbf{y}$ 为:

$$\begin{aligned} \mathbf{K} &= [K(\mathbf{x}_i, \mathbf{x}_j)]_{N \times N} \\ \mathbf{y}^T &= [y_1, y_2, \dots, y_N] \end{aligned} \quad (15)$$

最终得到 LS-SVM 的线性分类表达式为:

$$y(x) = \text{sign} \left[\sum_{i=1}^N a_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + b \right] \quad (16)$$

采用基于改进型的 FOA 优化算法确定 LS-SVM 模型中预测精度受到影响的 4 个参数:模型数据嵌入维度 m 、时间延迟 τ 、正则化参数 γ 和核函数参数 σ 。本文选用高斯核函数,即 $K(\mathbf{x}_i, \mathbf{x}_j) = \exp[-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / (2\sigma)]$ 。模型参数的优化过程如下:

a. 初始化模型参数,其中系统动力学维数 d 设为 2,时间延迟 τ 以 5 min 为基准,正则化参数 γ 和核函数参数 σ 均设为随机化初始值;

b. 建立改进型 FOA 优化参数模型,定义适应度函数以训练数据样本均方差为准则;

c. 提取优化后的最佳参数,将预测日测试数据代入模型,最终得到预测结果。

优化模型的参数值见表 1。

表 1 子序列预测模型的 4 个参数优化值

Table 1 Optimization values of four parameters for subsequence forecasting model

子序列	相空间重构		LS-SVM 模型参数	
	m	τ	γ	σ
IMF ₁	5	2	74.393	353.342
IMF ₂	9	1	128.498	224.430
IMF ₃	6	2	147.584	82.423
IMF ₄	8	1	88.485	143.320
IMF ₅	8	1	45.958	64.394
IMF ₆	10	1	45.485	39.000
IMF ₇	12	1	89.454	93.212
IMF ₈	6	2	89.433	379.482
IMF ₉	8	2	95.493	65.320
$r_9(t)$	8	2	90.340	44.238

注: τ 取值 2 表示 $2 \times 5 = 10(\text{min})$,其他类似。

为了验证本文所用的改进型 FOA 对于参数优化的优越性,将 GA、PSO 算法和改进前 FOA 分别优化上述参数。图 5 为 4 种优化算法的适应度值迭代曲线。

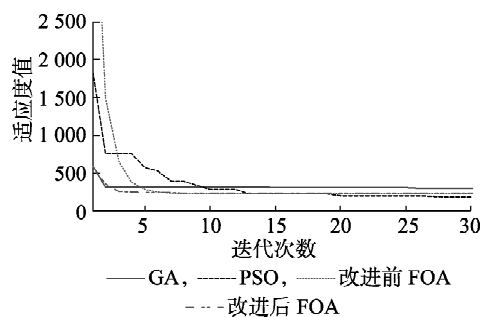


图 5 不同学习算法优化参数的适应度曲线

Fig.5 Fitness curves of different learning algorithms for optimizing parameter

由图 5 可见,采用改进型 FOA 优化模型参数时,其迭代次数最少,4 次就能达到收敛;而未经改进的 FOA 实现收敛,需迭代 8 次;GA 在优化参数时适应度值较大;采用 PSO 算法优化参数最终能够得到更小的结果,但达到稳态收敛需要迭代 14 次左右。显然在模型参数的优化方面,改进型的 FOA 整体效果更为理想。改进前后 FOA 优化算法的种群搜索寻优路径如图 6 所示。

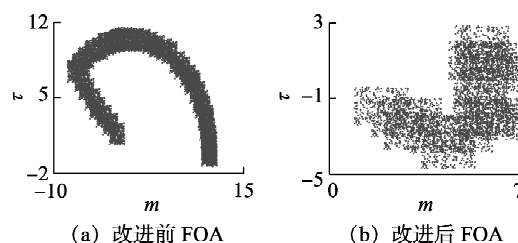


图 6 改进前后 FOA 搜索寻优路径图

Fig.6 Optimal searching path for FOA before and after improvement

由图 6 可见,改进后 FOA 种群搜索范围(宽度)

更广,搜索路径更短,有利于优化参数的选取,避免陷入局部最优,缩短了计算时间。

3.5 风速短期预测结果

由改进的 FOA 算法优化参数后,得到原始风速时序信号模态分解的相空间重构参数和 LS-SVM 模型参数,结合这些参数值建立基于 LS-SVM 的风速预测模型。风速分解序列预测结果如图 7 所示(采样分辨率为 5 min)。由图可见,原始序列波形 IMF_1 — IMF_3 由于频率特性波动较大,其预测波形存在较大的误差,而 IMF_4 — IMF_9 以及剩余分量 $r_9(t)$ 比较平缓,频率较为稳定,有利于曲线的拟合,预测波形与真实风速序列的变化规律基本一致。

风速预测子序列叠加后即可得到最终风速的预测值。本文分别基于 EEMD 模型、EMD 模型、LS-SVM 模型以及 ARMA 模型对风速进行短期预测,通过对比证明本文方法的有效性。4 种方法对比波形图如图 8 所示(采样分辨率为 10 min),由图可见,基于 EEMD 模型的风速预测模型精度更高。

为了能够更加直观地评估这 4 种模型的整体预测性能,本文分别将 4 种模型的波形误差量化分析,

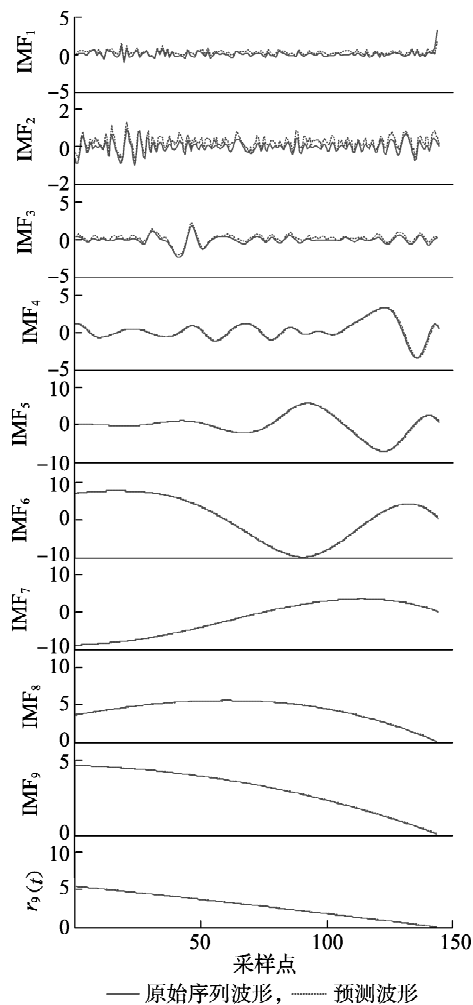
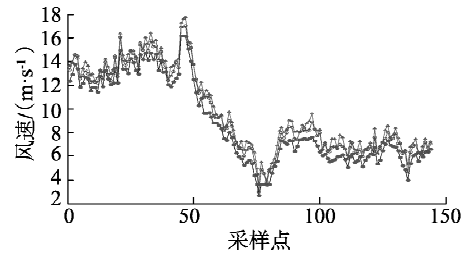
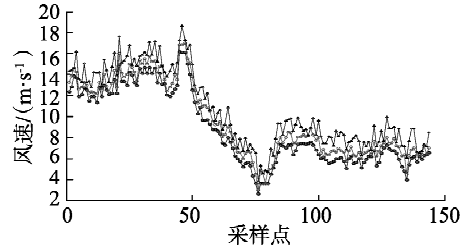


图 7 风速序列波形

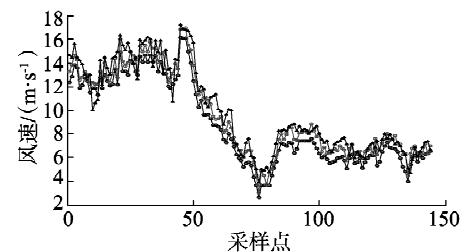
Fig.7 Waveform of wind speed sequence



(a) EEMD 与 EMD 的风速预测结果



(b) EEMD 与 LS-SVM 的风速预测结果



(c) EEMD 与 ARMA 的风速预测结果

—实际风速值, —EEMD 预测风速值, —ARMA 预测风速值

图 8 4 种风速预测方法的结果对比

Fig.8 Comparison of results among four wind speed forecasting methods

采用平均绝对误差百分比 MAPE (Mean Absolute Percentage Error) 和均方根误差 RMSE (Root Mean Square Error) 进行对比。这 2 种误差指标表达式为:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{H_{fct}(i) - H_{true}(i)}{H_{true}(i)} \right| \times 100\% \quad (17)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (H_{fct}(i) - H_{true}(i))^2} \quad (18)$$

其中, $H_{fct}(i)$ 为预测值; $H_{true}(i)$ 为实测值。MAPE 可评估模型的预测误差程度, 而 RMSE 可衡量预测结果的整体精度。预测误差统计指标见表 2。

表 2 风速误差评价指标

Table 2 Evaluation indexes for wind speed error

预测模型	评价指标	
	MAPE/%	RMSE
EEMD	9.15	0.71
EMD	14.41	1.10
LS-SVM	22.99	1.80
ARMA	20.12	1.54

从表 2 风速预测误差指标值和图 8 风速预测波形对比可见:

a. 采用基于 EEMD 和 EMD 的方法对风速序列进行模态分解后建立的预测模型预测精度高于传统

ARMA 模型、LS-SVM 模型,这说明对于风速这种具有不平稳、波动性大的时间序列,考虑波形的模态分解能够有效改善预测的精度;

b. 由 EEMD 模型与 EMD 模型的结果对比可以发现,改进后的模型在一定程度上提高了精度,有效避免了模态混叠现象对波形分解的干扰,验证了本文所建模型的合理性。

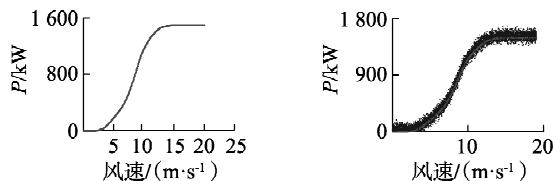
3.6 风功率短期预测结果

考虑风功率相关气象特征因素,利用 NWP 提供的天气预报数据结合 EEMD 短期风速预测结果,考虑风电机组可获风轮输出功率为:

$$P = \rho A C_p v^3 / 2 \quad (19)$$

其中, P 为风轮输出功率(kW); v 为风速(m/s); ρ 为空气密度(kg/m^3); A 为风轮扫掠面积(m^2); C_p 为风能利用系数,取最大利用系数 $C_{p,\max} = 16/27$ 。

图9为风速-风功率拟合曲线。对于单台风机而言, A 为常数,通常认为空气密度也保持在恒定值,同时考虑风机的最大输出功率、风机的切入风速 v_{in} 和切出风速 v_{off} ,可以得到微网能量管理系统模拟 1.5 MW 风机的理想输出功率拟合曲线如图 9(a) 所示。由于风机运行期间,实际出力不可能严格服从理想功率曲线,风机受外界随机因素的影响会出现一些浮动,如图 9(b) 所示。



(a) 风机理论输出功率曲线 (b) 风速-风功率拟合曲线散点图

图9 风速-风功率拟合曲线

Fig.9 Fitting curve of wind speed vs. wind power

将风速-风功率拟合曲线线性化,划分为4个阶段:

a. 初始低风速阶段($v < v_{\text{in}}$),风速小于切入风速,不足以带动风机输出功率;

b. 中风速上升阶段($v_{\text{in}} \leq v < v_N$),风速不小于切入风速且不大于额定风速 v_N 时,较小的风速变化会产生明显的功率输出;

c. 高风速饱和阶段($v_N \leq v < v_{\text{off}}$),风速不小于于额定风速但不大于切出风速时,风机输出功率为恒定值即额定功率 P_N ,风速的变化不会造成功率的变化;

d. 风机停止运行阶段($v \geq v_{\text{off}}$),风速大于切出风速时,为保护风机此时风机应停止工作,输出功率为0。

可得风速-风功率之间函数关系式为:

$$P(v) = \begin{cases} 0 & 0 < v < v_{\text{in}} \\ P_N \frac{v - v_{\text{in}}}{v_N - v_{\text{off}}} & v_{\text{in}} \leq v < v_N \\ P_N & v_N \leq v < v_{\text{off}} \\ 0 & v \geq v_{\text{off}} \end{cases} \quad (20)$$

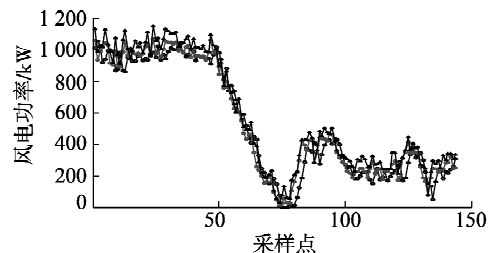
最终,风速-风功率转化曲线建立数学模型后,结合风速预测值便可求得风电功率的预测值。

预测值的误差评价指标和风电功率短期预测曲线分别见表3和图10(采样分辨率为10 min)。

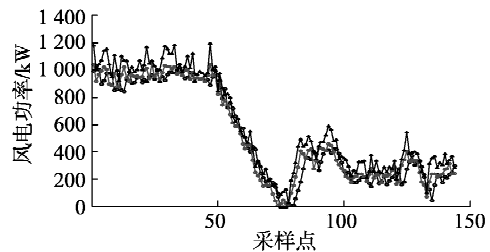
表3 功率误差评价指标

Table 3 Evaluation indexes of power error

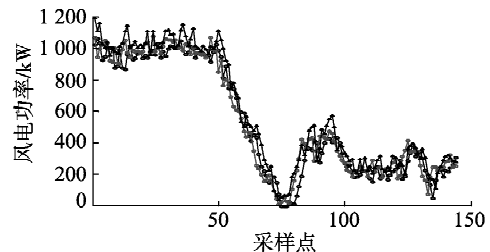
预测模型	评价指标	
	MAPE/%	RMSE
EEMD	19.73	29.36
EMD	34.32	58.04
LS-SVM	55.38	121.91
ARMA	46.24	89.45



(a) EEMD 与 EMD 的功率预测结果



(b) EEMD 与 LS-SVM 的功率预测结果



(c) EEMD 与 ARMA 的功率预测结果

— 实际功率值, — EEMD 预测值, — ARMA 预测值

图10 4种风电功率预测模型的结果对比

Fig.10 Comparison of results obtained by four wind power forecasting models

由表4和图10可得,采用 EEMD 建立的短期风功率预测模型精度高于其他3种模型,这是由于 EEMD 模型将风速序列信号分解为较为平稳的波形,有利于波形的回归,并且避免了分解后子序列之

间产生的模态混叠。在风机处于风速波动大的外界环境下,仅针对历史风速数据无法有效改善功率预测的精度,本文建立的风电预测模型通过间接法改进预测风速的模型,最终提高了风电功率预测的精度。

4 结论

风电功率由于间歇性和随机波动性,在预测时会得到较高的误差。因此,本文基于 EEMD 和 LS-SVM 建立了风电功率短期预测模型,采用改进的 FOA 优化算法确定了 LS-SVM 和风速分解子序列空间重构的参数,并与 EMD、LS-SVM 以及传统 ARMA 模型分别进行对比,结果表明:

a. 与 GA、PSO 算法相比,本文所采用的改进 FOA 能够提高收敛和迭代速度,扩大搜索路径范围,并能够有效避免陷入局部极值;

b. 与传统的 EMD 相比,本文采用改进的 EEMD 算法能够消除模态混叠现象对于风速预测的不利影响,得到了相对平稳的 IMF 分量,降低了波动性对原始预测值的影响;

c. 为了体现本文所建模型的优点,采用 EMD 和 LS-SVM 这 2 种模型进行对比分析,结果证明采用 EEMD 模型能够更好地实现风电功率的预测,有效地提高了预测精度。同时与 ARMA 模型进行对比,更充分地证明了本文所建模型的优势。

参考文献:

- [1] 叶瑞丽,郭志忠,刘瑞叶,等. 基于置信区间估计及储能装置优化配置的风电场可靠出力研究[J]. 电力自动化设备,2017,37(5):85-91.
YE Ruili, GUO Zhizhong, LIU Ruiye, et al. Reliable power output based on confidence interval estimation and optimal ESS configuration of wind farm[J]. Electric Power Automation Equipment, 2017, 37(5):85-91.
- [2] 李茜,刘天琪,何川,等. 含风电系统的有功和备用协调优化方法[J]. 电力自动化设备,2016,36(7):7-14.
LI Qian, LIU Tianqi, HE Chuan, et al. Coordinated optimization of active power and reserve capacity for power grid with wind farm[J]. Electric Power Automation Equipment, 2016, 36(7):7-14.
- [3] 刘燕华,刘冲,李伟花,等. 基于出力模式匹配的风电集群点多时间尺度功率预测[J]. 中国电机工程学报,2014,34(25):4350-4358.
LIU Yanhua, LIU Chong, LI Weihua, et al. Multi-time scale power prediction of wind farm cluster based on profile pattern matching[J]. Proceedings of the CSEE, 2014, 34(25):4350-4358.
- [4] 彭小圣,熊磊,文劲宇,等. 风电集群短期及超短期功率预测精度改进方法综述[J]. 中国电机工程学报,2016,36(23):6315-6327.
PENG Xiaosheng, XIONG Lei, WEN Jinyu, et al. A summary of the state of the art for short-term and ultra-short-term wind power prediction of regions[J]. Proceedings of the CSEE, 2016, 36(23):6315-6327.
- [5] 王铮,王伟胜,刘纯,等. 基于风过程方法的风电功率预测结果不确定性估计[J]. 电网技术,2013,37(1):242-247.
WANG Zheng, WANG Weisheng, LIU Chun, et al. Uncertainty estimation of wind power prediction result based on wind process method[J]. Power System Technology, 2013, 37(1):242-247.
- [6] 李湃,管晓宏,吴江,等. 基于天气分类的风电场群总体出力特性分析[J]. 电网技术,2015,39(7):1866-1872.
LI Pai, GUAN Xiaohong, WU Jiang, et al. Analyzing characteristics of aggregated wind power generation based on weather regime classification[J]. Power System Technology, 2015, 39(7):1866-1872.
- [7] 洪翠,温步瀛,林维明. 基于改进 OLS-RBF 神经网络模型的短期风电场出力预测[J]. 电力自动化设备,2012,32(9):40-43,59.
HONG Cui, WEN Buying, LIN Weiming. Short-term forecasting of wind power output based on improved OLS-RBF ANN model[J]. Electric Power Automation Equipment, 2012, 32(9):40-43, 59.
- [8] 钱政,裴岩,曹利霄,等. 风电功率预测方法综述[J]. 高电压技术,2016,42(4):1047-1060.
QIAN Zheng, PEI Yan, CAO Lixiao, et al. Review of wind power forecasting method[J]. High Voltage Engineering, 2016, 42(4):1047-1060.
- [9] 薛禹胜,雷兴,薛峰,等. 关于风电不确定性对电力系统影响的评述[J]. 中国电机工程学报,2014,34(29):5029-5040.
XUE Yusheng, LEI Xing, XUE Feng, et al. A review on impacts of wind power uncertainties on power systems[J]. Proceedings of the CSEE, 2014, 34(29):5029-5040.
- [10] 欧阳庭辉,查晓明,秦亮,等. 含核函数切换的风电功率短期预测新方法[J]. 电力自动化设备,2016,36(9):80-86.
OUYANG Tinghui, ZHA Xiaoming, QIN Liang, et al. Short-term wind power prediction based on kernel function switching[J]. Electric Power Automation Equipment, 2016, 36(9):80-86.
- [11] CHOWDHURY M A, SHEN W X, HIJIAZIN I, et al. Impact of DFIG wind turbines on transient stability of power systems—a review [C]//IEEE Conference on Industrial Electronics and Applications, 2013. ICIEA'08. Melbourne, Australia: [s.n.], 2013:73-78.
- [12] 叶燕飞,王琦,陈宁,等. 考虑时空分布特性的风速预测模型[J]. 电力系统保护与控制,2017,45(4):114-120.
YE Yanfei, WANG Qi, CHEN Ning, et al. Wind forecast model considering the characteristics of temporal and spatial distribution[J]. Power System Protection and Control, 2017, 45(4):114-120.
- [13] 刘立阳,孟邵良,吴军基,等. 基于风电预测误差区间的动态经济调度[J]. 电力自动化设备,2016,36(9):87-93.
LIU Liyang, MENG Shaoliang, WU Junji, et al. Dynamic economic dispatch based on wind power forecast error interval[J]. Electric Power Automation Equipment, 2016, 36(9):87-93.
- [14] 朱霄珣,韩中合. 基于 PSO 参数优化的 LS-SVM 风速预测方法研究[J]. 中国电机工程学报,2016,36(23):6337-6343.
ZHU Xiaoxun, HAN Zhonghe. Research on LS-SVM wind speed prediction method based on PSO[J]. Proceedings of the CSEE, 2016, 36(23):6337-6343.
- [15] 雷若冰,徐箭,孙辉,等. 基于相关性分析的风电场群风速分布预测方法[J]. 电力自动化设备,2016,36(5):134-140.
LEI Ruobing, XU Jian, SUN Hui, et al. Wind speed distribution forecasting based on correlation analysis for wind farm group[J]. Electric Power Automation Equipment, 2016, 36(5):134-140.
- [16] 甘迪,柯德平,孙元章,等. 考虑爬坡特性的短期风电功率概率预测[J]. 电力自动化设备,2016,36(4):145-150.
GAN Di, KE Deping, SUN Yuanzhang, et al. Short-term probabilistic wind power forecast considering ramp characteristics[J]. Electric Power Automation Equipment, 2016, 36(5):134-150.

- [17] 杨锡运,关文渊,刘玉奇,等. 基于粒子群优化的核极限学习机模型的风电功率区间预测方法[J]. 中国电机工程学报,2015,35(增刊1):146-153.
YANG Xiyun, GUAN Wenyuan, LIU Yuqi, et al. Prediction intervals forecasts of wind power based on PSO-KELM[J]. Proceedings of the CSEE, 2015, 35(Supplement 1): 146-153.
- [18] 阳霜,罗滇生,何洪英,等. 基于 EMD-LSSVM 的光伏发电系统功率预测方法研究[J]. 太阳能学报,2016,37(6):1387-1395.
YANG Shuang, LUO Diansheng, HE Hongying, et al. Output power forecast of PV power system based on EMD-LSSVM model[J]. Acta Energaie Solaris Sinica, 2016, 37(6): 1387-1395.
- [19] 王贺,胡志坚,陈珍,等. 基于集合经验模态分解和小波神经网络的短期风功率组合预测[J]. 电工技术学报,2013,28(9):137-144.
WANG He, HU Zhijian, CHEN Zhen, et al. A hybrid model for wind power forecasting based on ensemble empirical mode decomposition and wavelet neural networks[J]. Transactions of China Electrotechnical Society, 2013, 28(9): 137-144.
- [20] SHAN D, CAO G H, DONG H J, et al. LGMS-FOA: an improved fruit fly optimization algorithm for solving optimization problems[J]. Mathematical Problems in Engineering, 2013, 2013(7): 1256-1271.
- [21] 张学清,梁军. 风电功率时间序列混沌特性分析及预测模型研究[J]. 物理学报,2012,61(19):70-81.

ZHANG Xueqing, LIANG Jun. Chaotic characteristics analysis and prediction model study on wind power time series[J]. Acta Physica Sinica, 2012, 61(19): 70-81.

作者简介:



程启明

程启明(1965—),男,江苏盐城人,教授,主要研究方向为电力系统自动化、发电过程控制、先进控制及应用等(**E-mail**: chengqiming@sina.com);

陈路(1990—),男,安徽淮南人,硕士研究生,主要研究方向为电力系统自动化、电机控制(**E-mail**: 644401269@qq.com);

程尹曼(1990—),女,上海人,硕士研究生,主要研究方向为电力系统自动化、新能源发电控制等(**E-mail**: chengyinman@hotmail.com);

张强(1989—),男,河南焦作人,硕士研究生,主要研究方向为电力系统自动化、微网优化与运行等(**E-mail**: 276727579@qq.com);

高杰(1993—),男,浙江温州人,硕士研究生,研究方向为电力系统自动化、继电保护等(**E-mail**: iamgaojie1993@163.com)。

Short-term wind power forecasting method based on EEMD and LS-SVM model

CHENG Qiming¹, CHEN Lu¹, CHENG Yinman², ZHANG Qiang¹, GAO Jie¹

(1. Shanghai Key Laboratory Power Station Automation Technology Laboratory, College of Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China;

2. North Power Supply Branch, Shanghai Electric Power Company, Shanghai 200041, China)

Abstract: Since the intermittent fluctuation characteristics of original wind speed signal have brought challenges to wind power forecasting, the EEMD(Ensemble Empirical Mode Decomposition) method is adopted to decompose the wind speed signal into stable sub-sequence signals on the frequency domain, which effectively improves the forecasting accuracy and avoids the modal aliasing phenomenon of the traditional EMD(Empirical Mode Decomposition) method. An improved FOA(Fruit fly Optimization Algorithm) is proposed, which takes the parameters of wind speed sub-sequence reconstruction and LS-SVM(Least Squares Support Vector Machine) as the optimal objective to establish the wind speed forecasting model, expanding the search range of the parameters and improving the convergence speed. The wind power can be obtained according to the transformation relationship between wind speed and wind power. Experimental results show that the proposed method has higher forecasting accuracy than the EMD or LS-SVM method.

Key words: microgrid; power forecasting; wind farms; mode decomposition; support vector machines; phase space reconstruction; fruit fly optimization algorithm