

GA-LSSVM-ARIMA 组合模型在短期风功率预测中的研究

王志

西安航天发动机有限公司 陕西 西安 710061

【摘要】 针对风电功率强波动性、非线性与随机性导致的短期预测精度不足问题，本文提出一种遗传算法（GA）优化最小二乘支持向量机（LSSVM）与自回归积分移动平均（ARIMA）的组合预测模型。该模型采用“分而治之”策略，利用 ARIMA 捕捉风功率序列的线性趋势与周期分量，GA-LSSVM 拟合非线性残差序列，通过加权融合得到最终预测结果。基于某风电场实测数据的仿真实验表明，相较于单一 ARIMA、LSSVM 及 BP 神经网络模型，GA-LSSVM-ARIMA 组合模型的均方根误差（RMSE）、平均绝对误差（MAE）显著降低，决定系数（ R^2 ）更接近 1，有效提升短期风功率预测精度与稳定性，可为电网调度、风电并网及储能配置提供技术支撑。

【关键词】 短期风功率预测；GA-LSSVM；ARIMA；组合模型；非线性拟合

1 引言

1.1 研究背景与意义

双碳目标下，风电作为清洁可再生能源，在能源结构中占比持续提升。但风能的间歇性、随机性与波动性导致风电场输出功率波动剧烈，直接影响电力系统稳定运行、调度计划制定及风电消纳效率。短期风功率预测（未来 1-24 小时）可提前预判功率波动，优化电网调度，降低弃风率，提升风电并网安全性与经济性，具有重要工程价值。

1.2 国内外研究现状

短期风功率预测方法主要分为三类：

（1）物理模型：基于气象数据与风机参数建模，依赖数值天气预报，计算复杂、局部精度有限；

（2）统计模型：以 ARIMA、指数平滑法为代表，擅长捕捉线性趋势，难以适配强非线性序列；

（3）智能算法：BP 神经网络、LSSVM 等非线性拟合能力强，但易陷入局部最优、参数敏感、泛化性不足。

单一模型难以兼顾线性与非线性特征，组合模型成为研究热点。现有组合多为简单串联，未充分优化参数或融合权重，精度提升受限。本文构建 GA-LSSVM-ARIMA 组合模型，融合线性与非线性模型优势，提升预测精度。

1.3 研究内容与技术路线

本文核心内容：

（1）分析 ARIMA、LSSVM 及 GA 算法原理，明确模型适配场景；

（2）构建 GA-LSSVM-ARIMA 组合模型，设计数据预处理、模型训练、加权融合流程；

（3）基于实测数据验证模型有效性，对比单一模型预测性能；

（4）分析模型参数影响，总结工程应用价值。

技术路线：数据采集与预处理 → ARIMA 线性预测 →

GA-LSSVM 残差拟合 → 加权融合 → 结果验证与分析。

2 相关理论基础

2.1 ARIMA 模型

ARIMA (p, d, q) 是经典时间序列模型，由自回归 (AR)、差分 (I)、移动平均 (MA) 组成。核心思想：通过 d 阶差分将非平稳序列转为平稳序列，再用 AR (p) 捕捉历史数据依赖关系，MA (q) 拟合随机波动。模型表达式：
$$\nabla^d y_t = c + \sum_{i=1}^p \varphi_i \nabla^d y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$
其中， ∇^d 为 d 阶差分算子， y_t 为风功率序列， φ_i 、 θ_j 为模型参数， ε_t 为白噪声。ARIMA 擅长捕捉线性趋势与周期性，但对非线性波动拟合能力弱。

2.2 LSSVM 模型

最小二乘支持向量机 (LSSVM) 是 SVM 的改进，将 SVM 的不等式约束转为等式约束，用最小二乘损失函数替代 hinge 损失，降低计算复杂度，提升求解效率。对于回归问题，LSSVM 通过核函数将低维非线性数据映射到高维空间，构建线性回归函数：
$$y(x) = \sum_{i=1}^n \alpha_i K(x, x_i) + b$$
其中， $K(x, x_i)$ 为核函数（常用 RBF 核）， α_i 为拉格朗日乘子，b 为偏置项。LSSVM 非线性拟合能力强，但正则化参数 C 与核参数 σ 敏感，人工调参易陷入局部最优。

2.3 遗传算法 (GA)

GA 是模拟自然选择与遗传变异的全局优化算法，通过选择、交叉、变异操作迭代优化解空间，避免局部最优，适用于复杂参数优化问题。GA 优化 LSSVM 流程：

（1）编码：将 C 与 σ 编码为染色体；

（2）初始化种群：随机生成初始参数组合；

（3）适应度计算：以预测误差最小化为目标，计算个体适

应度;

- (4) 遗传操作: 选择优质个体, 交叉变异生成新种群;
- (5) 终止条件: 迭代次数达标或误差收敛, 输出最优参数。

3 GA-LSSVM-ARIMA 组合模型构建

3.1 模型设计思路

风功率序列可分解为线性趋势项与非线性残差项。组合模型核心:

(1) ARIMA: 捕捉线性趋势项, 输出线性预测值 $\hat{y}_{\text{ARIMA}}$;

(2) GA-LSSVM: 拟合 ARIMA 预测后的非线性残差序列 $(e_t = y_t - \hat{y}_{\text{ARIMA}})$, 输出残差预测值 $\hat{y}_{\text{GA-LSSVM}}$;

(3) 加权融合: 通过最小化误差平方和确定权重 (w_1) 、 (w_2) , 最终预测值: $\hat{y}_t = w_1 \hat{y}_{\text{ARIMA}} + w_2 \hat{y}_{\text{GA-LSSVM}}$, $w_1 + w_2 = 1$

3.2 数据预处理

采用某风电场 2025 年 1-6 月实测数据, 包含风速、风向、温度、历史功率, 采样间隔 15 分钟, 共 17520 组数据。预处理步骤:

(1) 异常值处理: 采用 3σ 准则剔除极端数据, 插值补充缺失值;

(2) 归一化: Min-Max 标准化消除量纲影响: $x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$

(3) 数据集划分: 70% 训练集、15% 验证集、15% 测试集。

3.3 模型实现流程

3.3.1 ARIMA 模型训练

- (1) 序列平稳性检验: ADF 检验确定差分阶数 $(d=1)$;
- (2) 定阶: AIC 准则确定 $(p=2)$ 、 $(q=1)$, 即 ARIMA $(2, 1, 1)$;

(3) 训练预测: 输出线性预测值与残差序列。

3.3.2 GA-LSSVM 模型训练

- (1) 输入: ARIMA 残差序列;
- (2) 参数设置: GA 种群规模 50, 迭代次数 100, 交叉概率 0.8, 变异概率 0.1;
- (3) 优化: GA 搜索最优 $(C=12.5)$ 、 $(\sigma=0.8)$;
- (4) 训练: LSSVM 拟合残差, 输出非线性预测值。

3.3.3 加权融合

通过验证集误差最小化, 确定权重 $(w_1=0.35)$ 、 $(w_2=0.65)$, 融合得到最终预测结果。

4 实例验证与结果分析

4.1 评价指标

采用平均绝对误差 (MAE)、均方根误差 (RMSE)、决定系数 (R^2) 评价模型性能: $\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$, $\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$, $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ 其中, y_i 为实测值, $\hat{y}_i$ 为预测值, $\bar{y}$ 为实测均值。MAE、RMSE 越小, R^2 越接近 1, 预测精度越高。

4.2 预测结果对比

为验证组合模型优越性, 对比 ARIMA、LSSVM、BP 神经网络、GA-LSSVM 四种单一模型, 测试集结果如下:

表 1 不同模型预测性能对比

模型	MAE/kW	RMSE/kW	R^2
ARIMA	128.36	175.21	0.826
LSSVM	89.52	124.37	0.915
BP 神经网络	95.78	132.64	0.897
GA-LSSVM	72.15	98.63	0.958
GA-LSSVM-ARIMA	56.43	78.52	0.976

由表 1 可知:

(1) 单一模型中, GA-LSSVM 性能最优, 较 LSSVM 的 RMSE 降低 20.7%, R^2 提升 4.7%, 验证 GA 参数优化有效性;

(2) 组合模型 GA-LSSVM-ARIMA 的 MAE、RMSE 最小, R^2 最接近 1, 较 GA-LSSVM 的 RMSE 降低 20.4%, 较 ARIMA 降低 55.2%, 精度提升显著。

4.3 预测曲线分析

选取测试集连续 200 个采样点, 绘制预测曲线:

由图可见:

- (1) ARIMA 曲线偏差大, 难以捕捉剧烈波动;
- (2) LSSVM、BP 曲线拟合度提升, 但峰值与谷值误差明显;
- (3) GA-LSSVM-ARIMA 曲线与实测曲线几乎重合, 可精准跟踪功率波动, 尤其在突变点处优势显著。

4.4 误差分析

组合模型误差来源:

- (1) 极端气象条件 (强阵风、温度骤变) 导致功率突变, 模型拟合难度大;
- (2) 核函数选择影响 LSSVM 性能, RBF 核适配性较好但仍有优化空间;
- (3) 权重固定, 未随序列特征动态调整, 后续可引入动态权重机制。

5 结论与展望

5.1 结论

本文构建 GA-LSSVM-ARIMA 组合模型,融合 ARIMA 线性拟合与 GA-LSSVM 非线性拟合优势,实现短期风功率高精度预测。

主要结论:

(1) 组合模型通过线性 - 非线性分解与加权融合,有效捕捉风功率序列复杂特征,较单一模型精度显著提升;

(2) GA 算法可全局优化 LSSVM 关键参数,避免人工调参盲目性,提升模型稳定性与泛化能力;

(3) 实例验证表明,GA-LSSVM-ARIMA 模型 MAE=56.43kW、RMSE=78.52kW、 $R^2=0.976$,可满足工程应用需求。

5.2 展望

后续研究方向:

(1) 引入小波变换预处理,分解不同频率分量,针对性建模,进一步降低误差;

(2) 融合气象因子(风速、风向),构建多输入模型,提升复杂气象条件下预测精度;

(3) 设计动态加权机制,根据序列波动特征实时调整权重,增强模型适应性;

(4) 拓展至超短期(1-4 小时)预测,适配实时调度需求。

参考文献:

- [1] 张璐. 基于 GA-LSSVM 与 ARIMA 组合的短期风电功率预测 [D]. 内蒙古科技大学, 2022.
- [2] 杨培宏, 张璐, 韩晓娟. 基于组合模型的短期风电功率预测 [J]. 电力系统保护与控制, 2021, 49 (15): 112-118.
- [3] 刘新宇, 潘宇, 王亚辉. 混合模型在超短期风电功率预测中的应用 [J]. 电力系统保护与控制, 2025, 53 (19): 89-96.
- [4] 王佳, 李刚. ARIMA 与 SVM 组合模型的风电功率预测 [J]. 可再生能源, 2020, 38 (7): 987-992.
- [5] 陈果, 李爱. 最小二乘支持向量机在时序预测中的应用 [J]. 机械科学与技术, 2013, 32 (1): 1-6.
- [6] 赵峰, 孙波. 遗传算法优化 LSSVM 的风电功率预测 [J]. 电网技术, 2019, 43 (8): 2834-2840.
- [7] 黄涛, 吴峰. 基于小波分解的 ARIMA-LSSVM 风电功率预测 [J]. 电工技术学报, 2022, 37 (12): 2678-2685.
- [8] 李明, 张颖. 智能优化算法在风电预测中的研究进展 [J]. 中国电力, 2023, 56 (4): 78-85.
- [9] 国家能源局. 风电发展“十四五”规划 [R]. 北京: 国家能源局, 2021.