

11-17-2021

Combination Forecasting Model of Photovoltaic Power Based on Empirical Wavelet Transform

Chen Tao

Engineering Research Center for Internet of Things Technology Application Ministry of Education, Jiangnan University, Wuxi 214122, China;

Wang Yan

Engineering Research Center for Internet of Things Technology Application Ministry of Education, Jiangnan University, Wuxi 214122, China;

Zhicheng Ji

Engineering Research Center for Internet of Things Technology Application Ministry of Education, Jiangnan University, Wuxi 214122, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Combination Forecasting Model of Photovoltaic Power Based on Empirical Wavelet Transform

Abstract

Abstract: In order to improve the prediction accuracy of short-term photovoltaic power, a variable weight combined prediction model based on Empirical Wavelet Transform (EWT) and PSO-optimized random forest(RF) is proposed. *Gray correlation analysis is used to select similar days, EWT is used to decompose the power time series into sub-modes of different frequencies, and three modes of high, medium, and low frequency are reconstructed according to the frequency, PSO-RF and PSO-BP and PSO-LSSVM prediction models are established to dynamically calculate their respective weights for reconstruction, and error correction is performed to output the prediction results.* By predicting the output power of Australian photovoltaic power stations, the results verify the effectiveness of the EWT-PSO-RF combined model, which effectively improves the accuracy of ultra-short-term photovoltaic power prediction.

Keywords

PV power prediction, empirical wavelet transform, random forest, grey correlation degree, model parameter optimization

Recommended Citation

Chen Tao, Wang Yan, Ji Zhicheng. Combination Forecasting Model of Photovoltaic Power Based on Empirical Wavelet Transform[J]. Journal of System Simulation, 2021, 33(11): 2627-2635.

基于经验小波变换的光伏功率组合预测模型

陈涛, 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: 为提高短期光伏功率的预测精度, 提出一种基于经验小波变换(Empirical Wavelet Transform, EWT)和粒子群算法(Particle Swarm Optimization, PSO)优化随机森林(Random Forest, RF)的变权组合预测模型。利用灰色关联分析选出相似日, 使用EWT将功率时间序列分解成不同频率的子模态, 根据频率将其重构为高、中、低频3个模态, 建立PSO-RF, PSO-BP, PSO-LSSVM预测模型动态计算各自权值进行重构, 进行误差校正输出预测结果。通过对澳大利亚光伏电站输出功率进行预测, 结果验证了EWT-PSO-RF组合模型的有效性。

关键词: 光伏功率预测; 经验小波变换; 随机森林; 灰色关联度; 误差校正

中图分类号: TP391.9 文献标志码: A 文章编号: 1004-731X(2021)11-2627-09

DOI: 10.16182/j.issn1004731x.joss.21-FZ0709

Combination Forecasting Model of Photovoltaic Power Based on Empirical Wavelet Transform

Chen Tao, Wang Yan, Ji Zhicheng

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

Abstract: In order to improve the prediction accuracy of short-term photovoltaic power, a variable weight combined prediction model based on Empirical Wavelet Transform (EWT) and PSO-optimized random forest(RF) is proposed. *Gray correlation analysis is used to select similar days, EWT is used to decompose the power time series into sub-modes of different frequencies, and three modes of high, medium, and low frequency are reconstructed according to the frequency, PSO-RF and PSO-BP and PSO-LSSVM prediction models are established to dynamically calculate their respective weights for reconstruction, and error correction is performed to output the prediction results.* By predicting the output power of Australian photovoltaic power stations, the results verify the effectiveness of the EWT-PSO-RF combined model, which effectively improves the accuracy of ultra-short-term photovoltaic power prediction.

Keywords: PV power prediction; empirical wavelet transform; random forest; grey correlation degree; model parameter optimization

引言

伴随着经济的快速发展, 世界各国对于能源的消耗日益增加, 开发新能源已经成为各国可持续发展的必经之路。光伏发电具有无噪声、安全可靠等

优点, 但光伏发电难以大规模存储^[1-2]。随着光伏发电装机容量的迅速增长, 接入电网光伏发电的比重日益提高, 光伏发电的大规模并网会对电网的安全稳定和电能质量产生极大影响。由此可见, 光伏功率预测对电网规划和调度具有重大意义^[3-4]。

收稿日期: 2021-04-20 修回日期: 2021-07-30

基金项目: 国家重点研发计划(2018YFB170903); 国家自然科学基金(61973138)

第一作者: 陈涛(1994-), 男, 硕士生, 研究方向为智能算法与预测控制。E-mail: 15850696098@163.com

通讯作者: 王艳(1978-), 女, 博士, 教授, 研究方向为制造系统能效优化。E-mail: 15850696098@163.com

目前常见的光伏发电功率预测方法主要有 2 类^[5-6]: ①直接使用解析方程对光伏发电系统进行建模, 根据数值天气预报模型 NWP(Numerical Weather Predication)和光伏发电设备状况直接预测光伏电站输出功率, 无需历史数据, 但该类方法过于依赖 NWP 模型, 而 NWP 模型误差较大; ②利用统计学和机器学习方法对光伏功率进行预测, 称为“黑盒”方法, 主要是通过提取光伏电站历史发电数据来预测未来输出功率。这种方法对历史数据的质量要求较高, 实验证明其预测精度优于第一类方法^[7-8]。此外, 由于每种预测方法都有其局限性, 光伏功率预测也会使用组合模型, 组合模型能够结合不同预测模型的优点, 发挥各自优势提高模型预测的准确性。文献[9]将相似日理论和蝙蝠算法优化最小二乘支持向量机 LSSVM(Least Squares Support Vector Machines)相结合, 避免 LSSVM 陷入局部最优从而提高预测精度; 文献[10]提出将变分模态分解、深度回声网络 DESN(Deep Echo State Network)和稀疏高斯混合过程专家模型 MSGP(Multi-Sector General Permit)相结合, 采用 MSGP 对误差进行校正, 能够提高光伏功率的预测误差。除此之外, 还有一些组合模型会考虑将影响光伏发电系统输出功率的相关因素(如天气类型、设备安装位置, 角度)和机器学习模型相结合, 通过物理性能来选取合适的组合预测模型, 文献[11]利用 SOM(Self-Organizing Map)模型将在线的气象数据划分成为晴天、多云和雨天, 然后使用 ANN-RBFN(Artificial Neural Network-Artificial Neural Network)模型训练; Mora-Lopez 等^[12]通过收集的气象信息, 利用 FPA(Function Point Analyse)和多元回归模型预测输出功率。科研人员在光伏功率预测的精度、效率上取得了一些进步, 但每种单一的预测模型都有其固有的优点和缺陷, 人工神经网络模型预测效果较好, 收敛速度较快, 但容易陷入局部最优而无法找到最优解; 支持向量机模型泛化能力强, 易得全局最优解, 对神经网络模型做了补充, 缺点是核函数参数和折中系数的选取没有通

用规则, 模型参数选取对模型预测结果影响较大; 随机森林算法预测精度高、训练速度快, 但对高维度的稀疏数据, 效果不是特别好^[13]。综合上述分析, 不同的模型间互为补充, 如果构建组合模型就能结合各个模型的优点, 不同模型间相互补充克服各自的缺点, 组合模型的预测精度能够提高, 此外如果动态的分配每个模型的权重从而使组合模型适用不同环境的光伏电站, 就能够大幅提高组合模型的可靠性和稳定性。

基于上述分析, 本文提出基于经验小波变换的光伏功率组合预测模型。首先利用灰色关联度选出相似日, 使用经验小波变换 EWT(Empirical Wavelet Transform)将原始的光伏功率时间序列分解成一系列频率互异的子模态, 然后根据频率的大小对子模态进行重构, 建立 PSO-RF, PSO-BP, PSO-LSSVM 预测模型, 利用灰色关联分析动态计算各个模型之间的权重, 按照各自权重对 3 种预测模型重构, 再使用 PSO-LSSVM 方法对组合模型预测结果进行误差校正, 从而得到最终的预测结果。

1 经验小波分解

经验小波分解 EWT (Empirical Wavelet Transform)是一种将经验模态分解 EMD (Empirical Mode Decomposition)自适应性和小波变换的理论框架结合起来而提出的一种信号自适应分析方法, 解决了 EMD 模态混叠问题, 基于信号频谱的自适应分割, 在分割区间内构造具有紧凑支撑特性小波函数, 组成带通滤波器组, 提取信号不同频率部分对应的 IMF(Intrinsic Mode Functions)模态分量, 既能解决经验模态分解 EMD 存在的模态混叠问题, 同时分解得到的子模态数量更少^[14]。EWT 通过对信号频谱进行自适应分割, 将信号 $f(t)$ 分解成 N 个本征模态函数, 每个 IMF 分量都可以看成是一组调幅-调频(AM-FM)信号, 分解结果为^[15]:

$$f(t) = \sum_{k=0}^N f_k(t) \quad (1)$$

$$f_k(t) = \sum_{k=0}^N F_k(t) \cos \varphi_k(t) \quad (2)$$

式中: $F_k(t)$ 为第 k 个 IMF 分量 $f_k(t)$ 的幅值; $\varphi_k(t)$ 为相角。

假设将傅里叶支撑区间 $[0, \Pi]$ 分割成 N 个连续的部分, 各分割片段可表示为 $\Lambda_n = [w_{n-1}, w_n]$, 其中 $n=1, 2, \dots, N, w_0=0, w_n=\pi, w_n$ 为各分段的边界, 显而易见 $\cup_{n=1}^N \Lambda_n = [0, \pi]$ ($n=1, 2, \dots, N$), 其中, $\cup_{n=1}^N$ 表示 N 分割片段 Λ_n 的并集。在傅里叶分解后, 定义经验小波为每个 Λ_n 上的带通滤波器, 并根据 Meyer 小波^[16]的构造方法构造经验小波函数 $\hat{\Psi}_n(w)$ 和经验尺度函数 $\hat{\phi}_n(w)$:

$$\hat{\Psi}_n(w) = \begin{cases} 1, & w_n + \tau_n \leq |w| \leq w_{n+1} + \tau_{n+1} \\ \cos \left\{ \frac{\pi}{2} \beta \left[\frac{1}{2\tau_{n+1}} (|w| - w_{n+1} + \tau_{n+1}) \right] \right\}, & w_{n+1} - \tau_{n+1} \leq |w| \leq w_{n+1} + \tau_{n+1} \\ \sin \left\{ \frac{\pi}{2} \beta \left[\frac{1}{2\tau_n} (|w| - w_n + \tau_n) \right] \right\}, & w_n - \tau_n \leq |w| \leq w_n + \tau_n \\ 0, & \text{other} \end{cases} \quad (3)$$

$$\hat{\phi}_n(w) = \begin{cases} 1, & |w| \leq w_n - \tau_n \\ \cos \left\{ \frac{\pi}{2} \beta \left[\frac{1}{2\tau_{n+1}} (|w| - w_{n+1} + \tau_{n+1}) \right] \right\}, & w_{n+1} - \tau_{n+1} \leq |w| \leq w_{n+1} + \tau_{n+1} \\ 0, & \text{other} \end{cases} \quad (4)$$

式中: $\beta(x)$ 是位于 $[0, 1]$ 区间的任意函数, 一般取 $\beta(x) = x^4(35 - 84x + 70x^2 - 20x^3), \forall x \in [0, 1]; \gamma$ 为位于 $(0, 1)$ 的系数。原始信号 $f(t)$ 可被重构为

$$x(t) = W_x^\varepsilon(0, t) \otimes \phi_1(t) + \sum_{n=1}^N W_x^\varepsilon(n, t) \otimes \phi_n(t) \quad (5)$$

式中: \otimes 为卷积运算; $W_x^\varepsilon(0, t)$ 为逼近系数; $W_x^\varepsilon(n, t)$ 为 $x(t)$ 的经验小波变换。

根据经验小波重构理论, 信号 $x_k(t)$ 被分解后的模态分量为

$$\begin{cases} x_0(t) = W_x^\varepsilon(0, t) \otimes \phi_1(t) \\ x_k(t) = W_x^\varepsilon(n, t) \otimes \phi_n(t) \end{cases} \quad (6)$$

式中: IMF 分量 $x_0(t)$ 由近似系数 $W_x^\delta(0, t)$ 和经验尺度函数 $\phi_1(t)$ 的卷积构成, 代表原始信号的整体变

化趋势部分, 可以看作经验尺度分量; IMF 分量 $x_k(t)$ 由细节系数 $W_x^\delta(n, t)$ 和经验小波函数 $\phi_n(t)$ 的卷积构成, 代表原始信号中不同频域的具有明显差异的特征分量。因此, EWT 能够对信号中具有特征差异的各个分量进行有效分离。

2 基于 PSO-RF 算法的预测模型

2.1 算法流程

随机森林回归算法是采用 CART(Classification And Regression Tree)决策树作为弱学习器、采用并行结构组合加强的集成学习算法, 通过 bootstrap 算法抽取得到训练集及其特征后, 选用 CART 算法作为回归决策树的生成算法^[17]。选取特征中最小均方差的特征点为分切点, 将节点处数据集二分裂, 生成 2 个子节点, 重复该步骤直到分裂完成生成分裂回归树^[18]。节点的分裂规则为最小均方差:

$$m = \min_{A, s} \left[\min_{c_1} \sum_{x_i \in D_1(A, s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in D_2(A, s)} (y_i - c_2)^2 \right] \quad (7)$$

式中: s 为当前节点的所有训练集; A 为特征集, 特征集是从当前节点抽取出来。训练集 s 根据特征集 A 划分成子集 D_1, D_2 , 通过遍历 A 的取值, 算得子集 D_1, D_2 输出 y_i 的最小均方差之和, 最后遍历每个属性得到最小均方差的取值 m , 以及对应的属性和属性值, 从而得到该节点的生长信息。而后对生成的每一个子节点重复上述过程, 直到满足终止条件。

随机森林回归模型由多棵“弱”因子决策树构成, 且每棵决策树之间没有关联, 通过聚合它们的输出来提高预测的准确性和稳定性。在实际的应用中随机森林决策树棵树 n_{tree} 和分裂特征数 m_{try} 对于算法的预测精度影响较大, 通常这些参数普遍依靠经验选取, 针对上述问题, 本文提出使用粒子群算法对随机森林算法中的分裂特征数和决策树的棵树选取进行优化, 从而提高随机森林回归算法的预测精度。算法的流程如下:

(1) 初始化粒子群优化算法和随机森林回归算法的相关参数, 同时设置随机森林算法中的决策树棵数 n_{tree} 和分裂特征数 m_{try} 初值和取值范围;

(2) 设置当前粒子位置为个体最优位置 P_{ibest} , 然后重新计算所有粒子的适应度值, 将适应度值最优的粒子作为全局最优值 g_{best} ;

(3) 迭代计算粒子的位置和速度;

(4) 将得到的决策树的棵数和分裂特征数代入随机森林回归算法中, 计算平均误差, 并更新适应度值;

(5) 将每个粒子的适应度值和 P_{ibest} 相比较, 保留更优者; 再将更新后粒子的 P_{ibest} 和全局最优值 g_{best} 比较, 保留更优者;

(6) 判断是否满足迭代终止条件, 若满足则输出 m_{try} 和 n_{tree} , 若不满足, 则继续第(3)步;

(7) 在得到决策树棵数 n_{tree} 和分裂特征数 m_{try} 后, 在原始样本集中通过 bootstrap 方法随机选取 n_{best} 个训练集, 从而形成 n_{best} 棵决策树;

(8) 随机选取 m_{best} 个特征形成分裂特征集, 然后每棵决策树选择最优分裂特征进行分裂直到叶子节点;

(9) 最后将所有决策树的结果求均值即可得到模型最终的输出结果。

2.2 相似日选取原理

光伏电站的输出功率受众多因素的影响, 不同的天气情况下光伏输出功率差别很大^[19]。选用澳大利亚沙漠知识太阳能中心(DKA Solar Centre)艾丝丽泉地区 2016 年 4 月—2017 年 6 月的实测光伏功率数据, 太阳辐照度、温度、湿度、风速和压强作为光伏功率预测模型的输入变量, 通过分析上述气象因素对光伏功率历史数据进行搜索, 选出相似程度更高、更具针对性的训练样本和测试样本提高模型的预测精度。

首先利用极差法对数据进行归一化:

$$x_j(i) = [d_j(i) - m(i)] / [M(i) - m(i)] \quad (8)$$

式中: $d_j(i)$ 是第 j 个向量的第 i 个分量; $M(i)$ 和

$m(i)$ 分别是样本数据中第 i 个分量的最大值和最小值, 经过归一化之后的预测日和训练样本第 j 日的特征向量分别为

$$\mathbf{x}_0 = [\mathbf{x}_0(1), \mathbf{x}_0(2), \mathbf{x}_0(3), \mathbf{x}_0(4), \mathbf{x}_0(5)]$$

$$\mathbf{x}_j = [\mathbf{x}_j(1), \mathbf{x}_j(2), \mathbf{x}_j(3), \mathbf{x}_j(4), \mathbf{x}_j(5)]$$

\mathbf{x}_0 和 \mathbf{x}_j 在第 k 个影响因素的关联系数为:

$$\xi_i(k) = \frac{\min_j \min_k |\Delta_j(k)| + \rho \max_j \max_k |\Delta_j(k)|}{|\Delta_j(k)| + \rho \max_j \max_k |\Delta_j(k)|} \quad (9)$$

式中: $\Delta_j(k) = \mathbf{x}_0(k) - \mathbf{x}_j(k)$; ρ 是分辨系数, 分辨系数越小, 分辨力越大, 本文 ρ 取 0.5。综合各分量的相关系数, 相似度定义为

$$F_j = \prod_{k=1}^5 \xi_j(k) \quad (10)$$

从待测日开始, 逐一计算和待测日的相似度, 按照相似度的大小进行排序, 即可得到相似日。

3 基于 EWT-PSO-RF 的光伏功率组合预测模型

3.1 基于灰色关联度的变权组合预测

灰色关联度可反映 2 个系统或时间序列之间的相关程度, 可以将变化趋势的相似性进行量化, 一般用于动态过程分析。本文利用灰色关联度将 3 种预测模型进行结合, 提高模型的预测精度。

假设 3 种模型的预测数据序列为: $Y_1 = [y_{11}, y_{12}, \dots, y_{1n}]$, $Y_2 = [y_{21}, y_{22}, \dots, y_{2n}]$, $Y_3 = [y_{31}, y_{32}, \dots, y_{3n}]$, 预测日的功率序列的真实值序列为 $Opt = [o_1, o_2, \dots, o_n]$, 计算 3 组预测功率序列和真实序列间的绝对差值 Abs 。

$$Abs(i, j) = |y_{ij} - o_j| \quad i=1,2,3, \quad j=1,2,\dots,n \quad (11)$$

各个模型间关联系数矩阵 RI 为

$$RI = \frac{\min(Abs) + \rho \cdot \max(Abs)}{Abs + \rho \cdot \max(Abs)} \quad (12)$$

式中: ρ 为分辨系数, 一般取 0.5。

在得到关联系数矩阵后, 将 3 组关联系数均归一化到 [0, 1] 之间, 然后将每组预测模型的预测值和对应的关联系数相乘重构输出即可得组合模型光

伏功率预测值。

$$Y = W_1 Y_1 + W_2 Y_2 + W_3 Y_3 \quad (13)$$

$$W_i = [w_{i,1}, w_{i,2}, \dots, w_{i,n}] \quad i = 1, 2, 3 \quad (14)$$

3.2 误差校正

本文采用 PSO-LSSVM 对系统的误差进行校正, 误差校正能够提高光伏功率预测模型的精度。

step 1: 计算 t 时刻的预测误差值 $\varepsilon_t = y_t - \hat{y}_t$;

step 2: 将 t 时刻的预测误差值 ε_t 输入到 PSO-LSSVM 模型中进行同步学习, 能够获得 PSO-LSSVM 模型学习后的 t 时刻预测误差值 $\hat{\varepsilon}_t$;

step 3: 将各时刻分量的预测值和预测误差相加, 即可得到误差校正后的预测值 $y_t = \hat{y}_t + \hat{\varepsilon}_t$ 。

3.3 基于 EWT-PSO-RF 组合模型光伏功率预测

光伏功率时间序列的波动性较大, 选用一般的预测模型很难获得较高的预测精度。本文建立一种基于 EWT-PSO-RF 光伏功率组合预测模型, 流程如图 1 所示。

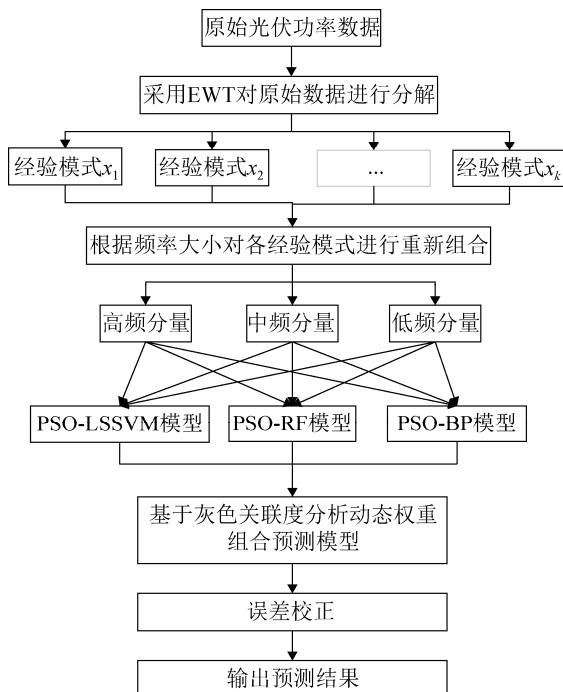


图 1 基于 EWT-PSO-RF 的光伏功率预测流程图
Fig. 1 Photovoltaic power prediction flow chart based on EWT-PSO-RF

步骤如下:

step 1: 对光伏功率原始数据进行预处理, 剔除异常值, 对数据进行归一化处理, 选取相似日;

step 2: 采用 EWT 分解原始光伏功率序列, 得到不同频率 IMF 的分量和一个剩余分量 $res(t)$;

step 3: 按照频率对分解得到的 IMF 分量进行重构, 可得高频、中频和低频分量;

step 4: 对重构后的各个分量分别建立 PSO-RF, PSO-LSSVM 和 PSO-BP 预测模型;

step 5: 对各个模型的预测值进行重构, 对重构后的预测值进行误差校正并输出, 即可得到最终的光伏功率预测结果。

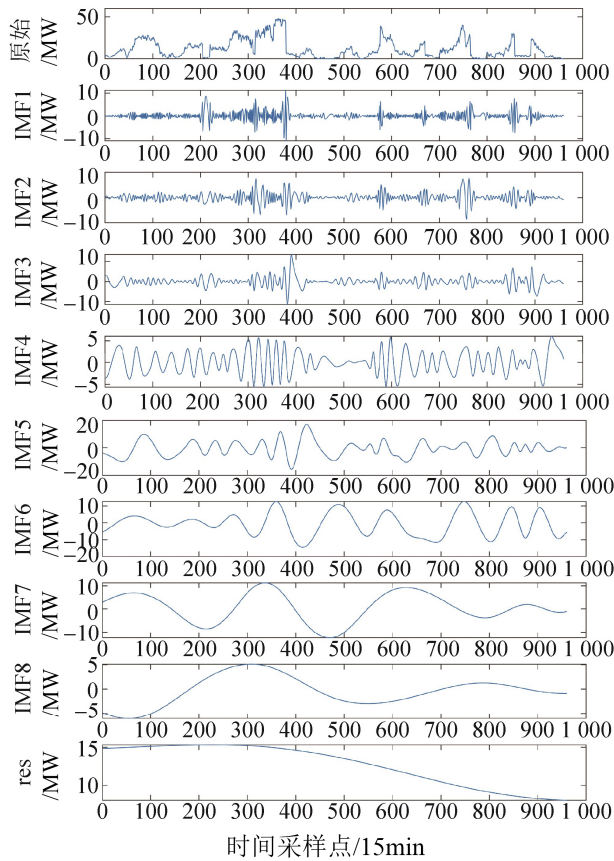
4 算例分析

4.1 EWT 和 EMD 分解结果

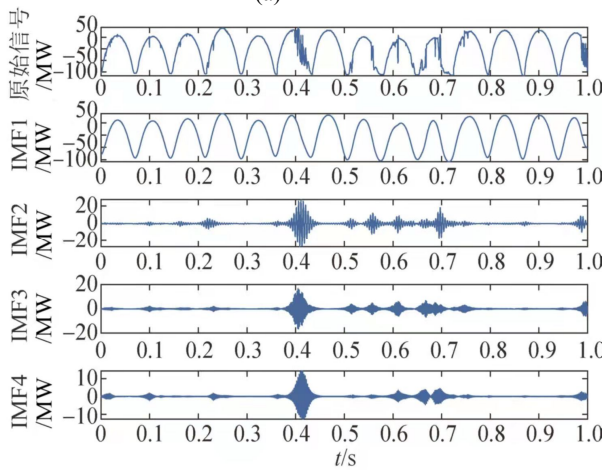
根据光伏功率预测模型, 本文选取澳大利亚沙漠知识太阳能中心(DKA Solar Centre)艾丝丽泉地区 2016 年 4 月—2017 年 6 月的实测光伏功率数据, 原始数据的采样间隔时间为 15 min, 选取 4 月 17 日(晴), 4 月 20 日(阴), 5 月 8 日(小雨), 6 月 1 日(多云)作为预测日, 首先采用灰色关联度法选取相似日, 选取关联度最高的 14 天(共 2 030 个样本点)作为训练样本。对原始的数据集分别采用 EMD 和 EWT 方法进行分解, 分解结果如图 2 所示, 由图 2 可知, EWT 和 EMD 的分解结果有较大差别, EWT 分解可得 5 个模态分量, 而 EMD 分解得到 9 个模态分量, EWT 分解有效地减少了子模态的数量, 降低了组合模型的计算规模; 其次, EMD 分解得到的模态分量是由高频到低频, 并且高频分量 IMF_1 和 IMF_2 幅值的波动较大。而 EWT 分解得到的模态分量是由低频到高频, 低频的波动较大。

4.2 模型评价指标

为了证明模型的有效性, 本文选用平均绝对百分比误差 $MAPE$ (Mean Absolute Percentage Error), 均方根误差 $RMSE$ (Root Mean Square Error)、平均绝对误差 MAE (Mean Absolute Error)和 R^2 作为模型的评价指标。



(a) EMD



(b) EWT

图2 基于EMD和EWT的光伏功率分解结果

Fig. 2 Decomposition results of photovoltaic power based on EEMD and EWT

平均绝对百分比误差:

$$\sigma_{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y(i) - \hat{y}(i)}{y(i)} \right| \times 100\%$$

均方根误差:

$$\sigma_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n [\hat{y}(i) - y(i)]^2}$$

线性回归决定系数:

$$R^2 = 1 - \frac{\sum (y - \hat{y}_i)^2}{\sum (\hat{y}_i - y_i)^2}$$

使用 $Im_{P_{NMAM}}$ 和 $Im_{P_{RMSE}}$ 量化本文所提组合模型与参考模型精度的提高幅度, 即预测误差的减小幅度。

$$Im_{P_{NMAM}} = (MAM_{ref} - MAE_{adv}) / MAM_{ref}$$

$$Im_{P_{RMSE}} = (RMSE_{ref} - RMSE_{adv}) / RMSE_{ref}$$

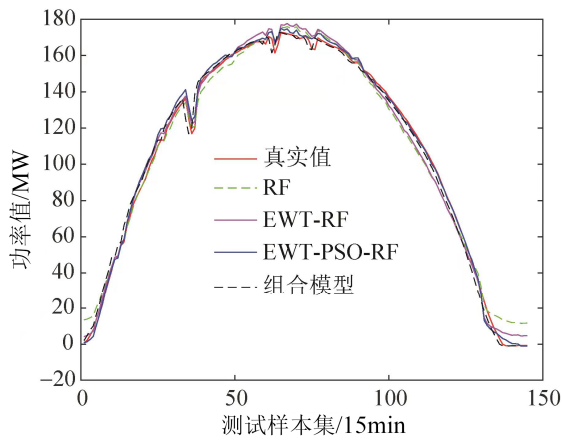
4.3 光伏功率组合预测模型仿真分析

根据光伏功率预测模型, 本文选取澳大利亚沙漠知识太阳能中心(DKA Solar Centre)艾丝丽泉地区2016年4月—2017年6月的实测光伏功率数据, 原始数据的采样间隔时间为15 min, 采样时间为07:00—19:00, 选取4月17日(晴), 4月20日(阴), 5月8日(小雨), 6月1日(多云)作为预测日, 选取相似日, 选取相似程度最高的14天(共2030个样本点)作为训练样本, 选取太阳辐照度、气温、湿度、风速和压强作为光伏功率预测模型的输入变量。首先采用EWT对训练样本进行分解, 按照分解得到子模态的频率大小将其重构为高频、中频、低频3种分量, 然后对重构得到的分量建立PSO-RF, PSO-LSSVM和PSO-BP预测模型, 基于灰色关联度建立变权组合预测模型, 同时采用PSO-LSSVM对预测结果进行误差校正并输出, 从而得到最终预测结果。

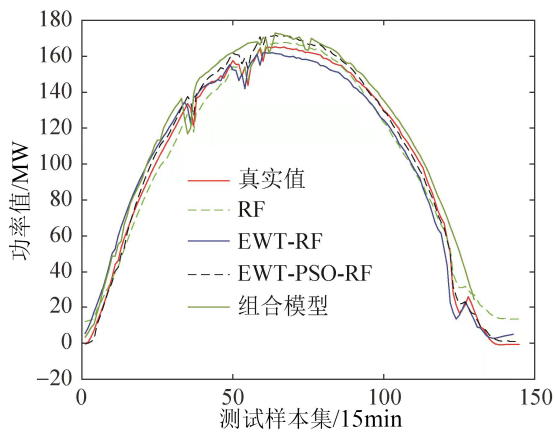
为了验证预测模型的有效性, 分别建立了RF, EWT-RF, EWT-PSO-RF和组合模型对样本进行测试得到的模型仿真结果如图3所示。

为了验证本文所提模型的性能, 分别对4种典型天气类型下光伏发电系统输出功率进行预测, 选用均方根误差、平均绝对误差和决定系数作为模型评价指标。由图3可以看出, 上述4种预测模型中组合模型预测精度最高, 表明训练样本在经过EWT分解后, 降低了IMF分量的非平稳性和随机

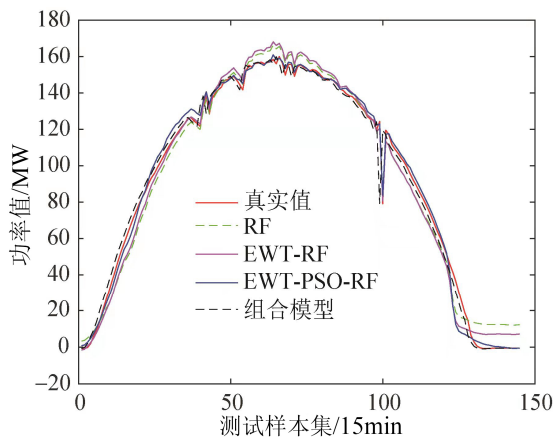
性, 降低了其对预测结果的影响; 同时选用 PSO 算法对 RF 模型进行优化, 模型参数结果如表 1 所示。此外, 从图 3 中可以看出, 07:00–09:00, 16:00–19:00 这 2 个时间段内的预测误差较大, 但是由于早晨和傍晚光伏的发电量较小, 因此其对模型最终的预测结果影响不大。



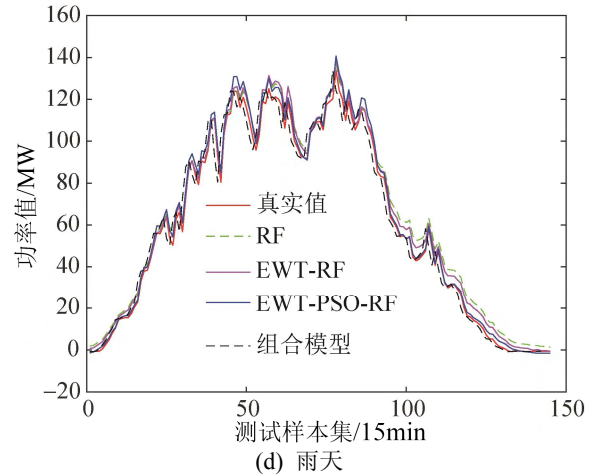
(a) 晴天



(b) 多云



(c) 阴天



(d) 雨天

图 3 4 类气象情况下模型预测结果
Fig. 3 Curves of prediction result in four meteorological conditions

表 1 PSO 优化模型参数结果
Tab. 1 Parameter results of PSO optimization model

气象类型	PSO 算法参数	PSO-RF 参数
晴天	$N=50, ger=20, V_{max}=0.9, V_{min}=0.5$	Mytry=1 Mytree=2
阴天	$N=50, ger=20, V_{max}=0.9, V_{min}=0.5$	Mytry=3 Mytree=3
雨天	$N=50, ger=20, V_{max}=0.9, V_{min}=0.5$	Mytry=6 Mytree=8
多云	$N=50, ger=20; V_{max}=0.9, V_{min}=0.5$	Mytry=2 Mytree=3

从表 2~3 可知: 不同的天气类型对各种模型的影响不同。晴天时, 4 种模型的预测效果都最好, 其中组合模型的预测效果最好, 平均相对误差只有 0.3, 即使是 EWT-RF 模型, 平均相对误差也只有 0.8, 明显低于其他天气类型。而在多云和阴天时, 3 种模型的预测精度明显低于晴天, 这主要是由于多云和阴天时天气变化频繁导致光伏系统输出功率随机性增加, 其中组合模型的预测误差最小, 最根本原因是原始的输出功率序列 EWT 分解后分别建立了 PSO-RF, PSO-LSSVM, PSO-BP 3 种预测模型, 最后通过灰色关联分析建立变权组合预测模型, 降低了不同时间序列间的相互影响, 组合预测模型吸取了 3 种预测模型的优点, 最大限度降低了数据的非平稳性, 模型的泛化能力强, 从而提高了模型的预测精度。而在雨天, 3 种模型的精度均最低, 这是因为雨天天气变化最频繁, 最易发生突变。

表2 4种模型预测结果比较

Tab. 2 Comparison of prediction results of 4 models

气象类型	方法	MAE	RMSE	R^2
晴天	RF	4.863 5	6.564 3	0.815 8
	EWT-RF	4.630 7	6.354 3	0.894 3
	EWT-PSO-RF	4.158 9	5.579 4	0.939 5
	组合模型	3.841 6	5.149 8	0.975 8
阴天	RF	5.438 5	6.435 5	0.802 3
	EWT-RF	5.385 3	6.149 9	0.874 9
	EWT-PSO-RF	4.878 4	5.799 4	0.902 5
	组合模型	4.564 8	5.495 6	0.952 9
雨天	RF	4.304 2	4.656 3	0.794 2
	EWT-RF	3.916	4.449 9	0.861 0
	EWT-PSO-RF	3.579 4	3.984 6	0.913 3
	组合模型	3.266 5	3.782	0.957 6
多云	RF	5.344 3	5.835 4	0.785 3
	EWT-RF	5.015 9	5.368 4	0.886 4
	EWT-PSO-RF	4.674 2	5.148 9	0.904 8
	组合模型	4.268 8	4.879 1	0.965 3

表3 组合模型预测精度提高百分比

Tab. 3 Percentage improvement of prediction accuracy of combined model/%

气象类型	方法	Im_{NAM}	Im_{RMSE}
晴天	EWT-RF	4.784 6	3.199 1
	EWT-PSO-RF	14.487 5	15.003 0
	组合模型	21.011 6	21.548 3
阴天	EWT-RF	3.677 4	4.437 8
	EWT-PSO-RF	10.298 7	9.884 0
	组合模型	17.064 5	14.604 9
雨天	EWT-RF	4.242 3	4.656 3
	EWT-PSO-RF	12.160 2	9.869 5
	组合模型	19.767 4	16.845 7
多云	EWT-RF	3.773 5	3.448 2
	EWT-PSO-RF	11.627 9	10.344 8
	组合模型	18.264 8	17.241 3

5 结论

本文提出了 EWT-PSO-RF 超短期光伏功率组合预测模型, 并采用澳大利亚光伏电站的数据进行仿真验证, 得出结论如下:

(1) 采用 EWT 对光伏功率时间序列进行分解, 通过分解原始光伏功率序列得到不同频率的子模态, 之后对分解得到的子模态进行重构, 降低了不同光伏功率序列间的相互影响。

(2) 本文同时采用 EWT 和 EMD 对光伏功率序列进行分解, 对比二者分解结果可知, EWT 分解得到的子模态的个数明显小于 EMD, 从而显著的降低了模型的计算量。

(3) 选用澳大利亚光伏电场的光伏功率数据进行仿真验证, 选出相似日后, 分别建立 EWT-RF、EWT-PSO-RF 和变权组合模型, 和典型的 RF 预测模型做对比, 组合模型的预测效果明显优于其他模型。组合模型能够结合各个模型的优点, 提高组合模型的预测精度, 从而证明了本文所提出的组合模型的正确性, 为光伏功率预测提供了一种新的思路。

参考文献:

- [1] 龚莺飞, 鲁宗相, 乔颖, 等. 光伏功率预测技术[J]. 电力系统自动化, 2016, 40(4): 140-151.
Gong Yingfei, Lu Zongxiang, Qiao Ying, et al. An Review of Photovoltaic Energy System Output Forecasting Technology[J]. Automation of Electric Power Systems, 2016, 40(4): 140-151.
- [2] 丁明, 王伟胜, 王秀丽, 等. 大规模光伏发电对电力系统影响综述[J]. 中国电机工程学报, 2014, 34(1): 1-14.
Ding Ming, Wang Weisheng, Wang Xiuli, et al. A Review on the Effect of Large-scale PV Generation on Power Systems[J]. Proceedings of the CSEE, 2014, 34(1): 1-14.
- [3] 杨国清, 张凯, 王德意, 等. 基于包络线聚类的多模融合超短期光伏功率预测算法[J]. 电力自动化设备, 2021, 1(2): 1-8.
Yang Guoqing, Zhang Kai, Wang Deyi, et al. Multi-mode Fusion Ultra-short-term Photovoltaic Power Prediction Algorithm Based on Envelop Clustering [J]. Electric Power Automation Equipment, 2021, 1(2): 1-8.
- [4] 许彪, 徐青山, 黄煜, 等. 基于藤 copula 分位数回归的光伏功率日前概率预测[J]. 电网技术 2021, 1(1): 1-13.
Xu Biao, Xu Qingshan, Huang Yu, et al. Prediction of PV Power Day-ahead Probability Based on Vine Copula Quantile Regression[J]. Power Grid Technology, 2021, 1(1): 1-13.
- [5] Antonanzas J, Osorio N, Escobar R, et al. Review of Photovoltaic Power Forecasting[J]. Solar Energy (S0038-092X), 2018, 136(5): 78-111.
- [6] 程泽, 李思宇, 韩丽洁, 等. 基于数据挖掘的光伏阵

- 列发电预测方法研究[J]. 太阳能学报, 2017, 38(3): 726-733.
- Cheng Ze, Li Siyu, Han Lijie, et al. Research on the Prediction Method of Photovoltaic Array Power Generation based on Data Mining [J]. Journal of Solar Energy, 2017, 38(3): 726-733.
- [7] Dolara A, Grimaccia F, Leva S, et al. Aphysically Brid Artificial Neural Network for Short Term Forecasting of PV Plant Power Output[J]. Energies (S1996-1073), 2015, 8(2): 1138-1153.
- [8] Widodo A, Shim M, Caesarendra W, et al. Intelligent Prognostics for Battery Health Monitoring based on Sample Entropy[J]. Expert System With Applications (S0957-4174), 2011, 38(9): 11763-11769.
- [9] 张彩庆, 郑强. SKBA-LSSVM 短期光伏发电功率预测模型[J]. 电力系统及其自动化学报, 2019, 31(8): 86-93. Zhang Caiqing, Zheng Qiang. Short-term PV Power Prediction Model of SKBA-LSSVM[J]. Journal of Power Systems and Automation, 2019, 31(8): 86-93
- [10] 王粟, 江鑫, 曾亮, 等. 基于 VMD-DESN-MSGP 模型的超短期光伏功率预测[J]. 电网技术, 2020, 44(3): 917-926. Wang Su, Jiang Xing, Zeng Liang, et al. Prediction of Ultra-short Term Photovoltaic Power based on VMD-DESN-MSGP Model[J]. Power Grid Technology, 2020, 44(3): 917-926.
- [11] Chen C, Duan S, Cai T, et al. Online24-hsolar Power Forecasting based on Weather Type Classification Using Artificial Neural Network[J]. Solar Energy (S0038-092X), 2011, 85(11): 2856-2870.
- [12] Mora-lopez L, Martinez-marchena I, Ppiliouguine M, et al. Machine Learning Approach for Next Day Energy Production Forecasting Rid Connected Photovoltaic Plants[J]. Environmental Modelling and Software, 2005, 20(6):753-760..
- [13] Shi J, Lee W J, Liu Y, et al. Forecasting Power Output of Photovoltaic System based on Weather Classification and Support Vector Machine[J]. IEEE Transactions on Industry Applications.2012.48(3):1064-1-69.
- [14] 张金金, 张倩, 马愿, 等. 基于改进的随机森林和密度聚类的短期负荷频域预测方法[J]. 控制理论与应用, 2020, 37(10): 2257-2265. Zhang Jinjin, Zhang Qian, Ma Yuan, et al. Short-term Load Frequency Domain Forecasting Method based on Improved Random Forest and Density Clustering[J]. Control Theory and Applications, 2020, 37(10): 2257-2265.
- [15] Daubechies I, Lu J, Wu H T. Synchrosqueezed Wavelet Transforms: An Empirical Mode Decomposition-like Tool[J]. Applied and Computational Harmonic Analysis (S1063-5203), 2011, 30(2): 243-261.
- [16] Daubechies I. Ten Lectures on Wavelets[J].USA: Society for Industrial and Applied Mathematics (S0916-7005), 1992, 21(5): 126-134.
- [17] Breiman L. Random Forests[M]. Statistics Department. University of California: Machine Learning, 2001.
- [18] Everitt B S. Classification and Regression Trees[M] Encyclopedia of Statistics in Behavioral Science. Cornell University, USA: John Wiley & Sons, Ltd, 2005.
- [19] 梁彩霞, 高赵亮. 基于相似日和 GA-DBN 神经网络的光伏发电短期功率预测[J]. 电气应用, 2019, 38(3): 97-102. Liang Caixia, Gao Zhaoliang. Short-term Power Prediction of Photovoltaic Power Generation based on Similar Days and GA-DBN Neural Network[J]. Electrical Application, 2019, 38(3): 97-102.