

基于模态分解和机器学习的水平辐照度预测方法研究

周童¹, 魏涛²

¹贵州大学数学与统计学院, 贵州 贵阳

²贵州省气象台, 贵州 贵阳

收稿日期: 2024年3月21日; 录用日期: 2024年4月19日; 发布日期: 2024年5月16日

摘要

光伏电站日渐发展成为新能源的焦点, 准确、可靠的水平辐照度(GHI)预测是解决光伏电站相关问题的方法之一。因此, 提出了一种有效的混合模型CEEMDAN-RF-LGBM-XGB, 用于预测每小时的水平辐照度。首先, 完全自适应噪声集合经验模态分解(CEEMDAN)将非线性和非平稳性的气象变量序列分解为若干个模态函数, 依据对水平辐照度的影响程度, 随机森林将分解的模态函数进行特征提取, 降低数据复杂度, 达到降维的目的, 完全自适应噪声集合经验模态分解和随机森林相结合得到一组简单且信息量丰富的影响因子。其次, 利用轻量级梯度提升机算法对中国三个不同气候地区的水平辐照度进行预测, 得到初始的水平辐照度预测值。针对预测模型训练中产生的固有误差, 引入极致梯度提升算法进行误差修正, 提升模型的预测性能。针对一些混合模型和独立模型对所提模型的性能进行了测试。测试结果显示, 最小%RMSE (%MAE)提高了23.14% (24.45%)。实验结果证明所提出的模型具有最高的预测精度, 且提出的混合结构模型均提高了单一模型的预测能力。

关键词

太阳辐照度预测, 完全自适应噪声集合经验模态分解(CEEMDAN), 机器学习, 误差修正

Research on Horizontal Irradiance Prediction Method Based on Modal Decomposition and Machine Learning

Tong Zhou¹, Tao Wei²

¹College of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

²Guizhou Provincial Meteorological Observatory, Guiyang Guizhou

Received: Mar. 21st, 2024; accepted: Apr. 19th, 2024; published: May 16th, 2024

Abstract

Photovoltaic power plants are increasingly becoming the focus of new energy, and accurate and reliable horizontal irradiance (GHI) prediction is one of the methods to solve the problems related to photovoltaic power plants. Therefore, an effective hybrid model CEEMDAN-RF-LGBM-XGB was proposed for predicting hourly GHI. Firstly, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) decomposes the nonlinear and non-stationary meteorological variable sequences into several modal functions. According to the degree of influence on horizontal irradiance, the random forest extracts the features of the decomposed modal functions to reduce the data complexity and achieve the goal of dimensionality reduction. The combination of the complete ensemble empirical mode decomposition with adaptive noise and the random forest obtains a set of simple and informative influence factors. Secondly, the lightweight gradient boosting algorithm was used to predict the horizontal irradiance in three different climate regions of China, and the initial horizontal irradiance predicted value was obtained. Aiming at the inherent errors generated during the training of prediction models, the extreme gradient boosting algorithm is introduced for error correction to improve the predictive performance of the model. The performance of the proposed model was tested on some mixed and independent models. The test results show that the minimum %RMSE (%MAE) of the proposed model is increased by 23.14% (24.45%). The experimental results show that the proposed models have the highest prediction accuracy, and the proposed hybrid structure models improve the prediction ability of a single model.

Keywords

Solar Irradiance Prediction, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), Machine Learning, Error Correction

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

太阳能是地球上最重要的可再生能源之一[1], 而太阳辐射是太阳能的主要来源。在“双碳”目标下[2], 为了更好地了解和利用太阳能资源, 研究水平辐照度成为一项重要的任务。水平辐照度(GHI)是指地球表面在水平方向上接收到的太阳辐射总量。研究 GHI 对多个领域都有着重要的意义。在太阳能光伏发电系统中, GHI 是一个关键的参数, 因为它表示在某个地区接收到的可利用太阳能辐射总量。这对于评估太阳能电池板的性能和规划太阳能发电厂非常重要[3]。

太阳辐照度的预测大致分为三类模型: 物理模型、统计模型和机器学习模型。物理模型是基于卫星图像和数值预报来进行预测, 但存在复杂性、高计算成本和不确定性的挑战[4]。传统的方法是将统计模型用于水平辐射照度的预测, 其中自回归(AR)及其变体、自回归移动平均(ARMA)模型等是最普遍的模型[5], 尽管这些模型可以有效地处理平稳时间序列, 但无法考虑到温度、风速和相对湿度等气象因素对辐照度的影响, 因此为了克服该缺陷。季节性 ARIMA 模型(SARIMAX)应运而生, 更加全面的考虑到解释变量对于因变量的影响[6]。随着人工智能的不断发展, 支持向量回归(SVR) [7] [8]、K 近邻(KNN) [9] 和随机森林(RF) [10]等机器学习模型也用于辐照度的预测。比起传统的统计模型和物理模型, 机器学习

模型通常能够更好的拟合数据之间的关系, 具有更好的表现力。

为了从复杂时间序列数据中提取有价值信息, 分解技术应运而生, 其中小波变换是最常用的数据分析技术。该技术的局限性在于预定义基函数的选择, 经验模态分解(EMD) [11]及其变体是自适应方法, EMD的一些流行变体有集成经验模态分解(EEMD) [12]、完全自适应噪声集合经验模态分解(CEEMDAN) [13]和改进的 CEEMDAN (ICEEMDAN) [14]等。

为了进一步增加 GHI 预测的准确性和稳定性, 本文提出了一种基于模态分解和机器学习的混合模型 EMD-RF-LGBM-XGB。该模型首先对收集的 GHI 数据利用完全自适应噪声集合经验模态分解(CEEMDAN)进行模态分解, 提取不同频率模态分量, 以解决数据预处理过程中存在模式混叠问题, 利用随机森林(RF)提取最重要的影响因素。通过模态分解(CEEMDAN)和随机森林(RF)的结合得到了一组简单且信息量丰富的预测因子, 利用轻量级梯度提升机算法(LGBM)对 GHI 进行预测得到最初的预测结果, 最后计算真实值与预测值的误差, 利用极致梯度提升模型(XGB)进行误差修正。实验结果表明, 本文所提出的模型在三个地区的表现更加优越。

2. 算法原理

2.1. CEEMDAN 模态分解

CEEMDAN 是一种改进的经验模态分解方法, 用于信号分解, 特别是用于处理非线性和非平稳信号。主要创新之处在于引入噪声模拟和统计平均的思想, 使得在处理非线性和非平稳信号时具有更好的鲁棒性。CEEMDAN 引入自适应噪声, 通过多次引入不同噪声模拟的信号进行分解。对于第 i 次噪声辅助的信号, 表示为:

$$x_i(t) = x(t) + n_i(t) \quad (1)$$

其中 $n_i(t)$ 是引入的噪声。

对每个噪声辅助的信号进行 CEEMDAN 分解, 得到一组 IMFs。这个过程可以表示为:

$$x_i(t) = \sum_{j=1}^{K_i} C_{ij}(t) + r_{i,K_i}(t) \quad (2)$$

其中, $C_{ij}(t)$ 为第 i 次分解的第 j 个 IMF, $r_{i,K_i}(t)$ 为第 i 次分解的剩余项。

对应位置的 IMFs 进行统计平均, 得到最终的 IMFs。这个过程可以表示为:

$$\text{AverageIMF}_j(t) = \frac{1}{N} \sum_{i=1}^N C_{ij}(t) \quad (3)$$

最终的信号分解为:

$$x(t) = \sum_{j=1}^J \text{AverageIMF}_j(t) + \text{Residue}(t) \quad (4)$$

其中, J 是分解得到的 IMFs 的数量, $\text{Residue}(t)$ 是剩余项的平均。

2.2. 随机森林(RF)

随机森林用于特征选取具有多个优点, 这些优点使得它成为处理高维数据、降低过拟合风险以及提高模型性能的强大工具。构建包含多个决策树的随机森林, 每个决策树使用随机抽样的数据子集和随机选择的特征子集。针对每个决策树, 计算每个特征的重要性。具体的计算方法取决于使用的指标, 如信息增益、基尼不纯度减小等。通常, 特征的重要性是通过评估特征在决策树中用于节点分裂的贡献来计算的。对于每个节点 m , 计算特征 j 的贡献, 对于每个特征 j , 计算其在整个树中的平均贡献。将所有决

策树中每个特征的平均贡献进行累加, 得到该特征的总体重要性分数:

$$\text{Importance}_j = \sum_{\text{trees}} \text{Average Contribution}_j \quad (5)$$

2.3. LGBM 模型

LGBM 是微软于 2017 年发布的梯度增强框架。其核心思想是将弱回归树组合成强回归树。我们的目标是学习一个回归函数, 使得对于输入的特征向量, 能够预测出对应的目标值。在 LGBM 中, 我们的目标是最小化损失函数, 损失函数是由模型的预测值和真实值之间的差异构成的, 通过调整的参数来最小化损失函数。

$$\text{Obj}(t) = \min L(y_i, \hat{y}_i) \quad (6)$$

其中是 y_i 真实值, \hat{y}_i 是模型预测值, $L(y_i, \hat{y}_i)$ 是损失函数。

LGBM 使用梯度提升算法来逐步逼近最优的 $f(x)$, 在第 k 次迭代中, 将当前模型的预测值记作 $f_{k-1}(x)$, 则我们的目标是找到一个回归树模型 $h_k(x)$ 来拟合残差:

$$h_k(x) = \arg \min_h \sum_{i=1}^m L(y_i, f_{k-1}(x_i) + h(x_i)) \quad (7)$$

通过求解这个优化问题, 得到了第 k 棵回归树的参数。然后, 我们通过一个学习率 η 来更新模型的预测值:

$$f_k(x) = f_{k-1}(x) + \eta h_k(x) \quad (8)$$

最终预测值是多棵树的组合

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (9)$$

2.4. XGB 模型

XGB, 全称为 Extreme Gradient Boosting, 是一种基于梯度提升机(Gradient Boosting)的算法, 它在多个方面进行了优化和改进, 使其在并行计算效率和预测性能上都非常出色。基于梯度提升框架, 在目标函数添加正则化项, 控制模型的复杂度, 提高泛化能力。

$D = \{(x_i, y_i)\}$ 表示包含 n 样本, m 个特征的数据集, 其中预测变量是由 k 个基本模型组成的加性模型。样本预测结果如下:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in \phi \quad (10)$$

其中 \hat{y}_i 表示预测标签, x_i 表示其中一个样本, 给定样本的预测值是 $f_k(x_i)$, XGB 的目标函数包括传统的损失函数和模型复杂度。如下所示:

$$\text{Obj} = \sum_{i=1}^m l(y_i, \hat{y}_i^{(t-1)} + f_i(x_i)) + \Omega(f_k) \quad (11)$$

其中 l 是一个损失函数, 它衡量预测值和目标值之间的差异。 Ω 是一个正则化项, 它降低模型的复杂性, 以避免过度拟合。

3. 方法和数据描述

3.1. 数据收集和站点描述

本文数据来源于国家太阳辐射数据库(NSRDB), 采用 PSM V3 模型和 Himawari 卫星, 时间分辨率为

1 h, 空间分辨率为 2×2 km, 选取了来自中国三个代表城市 2019~2020 两年的数据, 由于夜间太阳辐射强度低或为零, 本文只考虑各地区 GHI 非零的时间段, 排除由于太阳辐射强度低或零而导致的剩余时间, 前 80% 的数据用于训练, 剩余的数据用于模型的测试。三个地区的详细情况见表 1:

Table 1. Geographic information

表 1. 地理信息

地点	纬度	经度	气候
北京	39.91	116.39	暖温带半湿润大陆性季风气候
重庆	29.56	106.55	亚热带季风气候
哈尔滨	45.80	126.53	中温带大陆性季风气候

GHI 是考虑温度(T)、相对湿度(RH)、太阳天顶角(SZA)、风速(WS)等其他变量作为预测模型的预测变量进行预测的变量。输入预测因子是根据“Pearson 相关系数”选择的。在现有的气象资料中, 所选变量与目标变量之间存在显著的相关关系。

原始数据中可能存在缺失值和异常值, 采用中位数填充来处理, 将异常值和缺失值替换为整体数据的中位数, 以保持整体趋势。

3.2. 评估指标

在本文中, 使用均方根误差(RMSE)和平均绝对误差(MAE)来评估所提出模型的性能。上述性能指标可表述如下:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

3.3. CEMDAN-RF-LGBM-XG 组合模型

众所周知, GHI 受温度、相对湿度和风速等多种气象因素的影响, 具有明显的间歇性和波动性。同时 GHI 的变化可能受到多种因素的复杂交互影响, 对 GHI 的准确预测具有很大的挑战性。

由于原始数据存在波动性和趋势性, 为了更好地了解数据的特征, 采用 CEMDAN 的方法将原始数据的特征列进行分解, 希望获得不同频率的模态函数[14]。由于分解后的模态函数数目较大, 对于模型的计算量是一个较大的挑战, 基于此, 利用随机森林特征提取的方法对分解后的模态函数进行降维, 得到了一组简单且信息量丰富的预测因子, 同时使得预测模型的学习变得简单有效, 降低计算成本。在此基础上, 对所关注的 GHI 进行预测。在文献中并未找到 RF 与 CEMDAN 的使用。通过 CEMDAN 和 RF 的结合, 使得模型的学习变得简单且有效。为了获得模型更好的预测精度, 考虑到误差序列还存在未被充分提取的信息, 因此将误差序列进行修正, 得到更准确的预测效果。根据 GHI 变化特征, 本文提出的组合模型流程如图 1 所示, 具体步骤表述如下:

(1) 分解。利用 CEMDAN 将原始多元数据的影响因素依次进行模态分解;

(2) 提取影响因素。将随机森林用于特征选取[15], 对分解得到的模态函数利用随机森林进行特征重要性分析, 与目标列重构成一个新的数据后进行归一化处理;

- (3) 预测。在重构的数据集上, 利用网格搜索优化算法确定 LGBM 模型参数, 对 GHI 进行预测[16], 得到初始预测值;
- (4) 误差修正。计算真实值与预测值的误差, 利用 XGB 模型进行误差修正, 得到误差修正值;
- (5) 得到最终结果。将初始预测值与误差修正值结合, 求得最终预测结果。

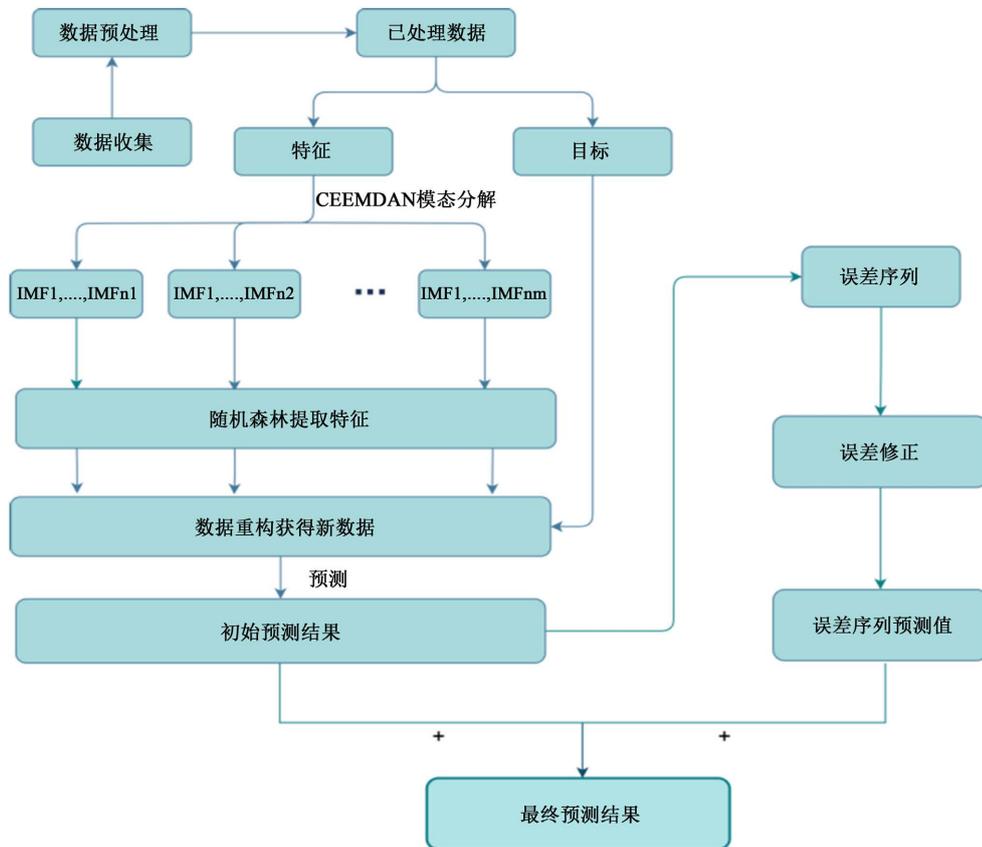


Figure 1. Model flow diagram
图 1. 模型流程图

4. 实证分析

本文所用到的模型命名如下表 2:

Table 2. Naming of models
表 2. 模型命名

命名	模型全称
模型 1	CEEMDAN-RF-LGBM-XGB
模型 2	CEEMDAN-RF-SVR-XGB
模型 3	CEEMDAN-RF-KNN-XGB
模型 4	CEEMDAN-LGBM-XGB
模型 5	CEEMDAN-RF-LGBM

续表

模型 6	LGBM-XG
LGBM	Light Gradient Boosting Machine
SVR	Support Vector Regression
KNN	K-Nearest Neighbors
XGB	Extreme gradient boosting

本文利用随机森林对影响因素列分解的模态函数进行特征选取, 利用网格搜索优化算法确定选取的特征数目为 20, 本文所提出的模型用于预测北京、重庆、哈尔滨三个中国不同气候地区的辐照度预测, 测试集上的误差指标如图, 本文考虑了三种经典的机器学习模型, LGBM 是基于集成学习的模型, 具有较强的泛化能力。本文将三种模型作为独立模型与基于分解和误差修正的混合模型进行比较, 混合模型具有更高的预测精度。

实验结果如下表 3 所示, 在北京地区, 模型 1 相比于 LGBM 的均方根误差和平均绝对误差分别降低了 25.14%、24.45%, 模型 2 相比于 SVR 的均方根误差和平均绝对误差分别降低了 26.09%、35.36%, 模型 3 相比于 KNN 的均方根误差和平均绝对误差分别降低了 18.11%、12.75%; 在重庆地区, 模型 1 相比于 LGBM 的均方根误差和平均绝对误差分别降低了 31.19%、33.12%。模型 2 相比于 SVR 的均方根误差和平均绝对误差分别降低了 28.14%、32.01%, 模型 3 相比于 KNN 的均方根误差和平均绝对误差分别降低了 26.82%、29.42%。在哈尔滨地区, 模型 1 相比于 LGBM 的均方根误差和平均绝对误差分别降低了 23.14%、25.64%, 模型 2 相比于 SVR 的均方根误差和平均绝对误差分别降低了 23.23%、26.30%, 模型 3 相比于 KNN 的均方根误差和平均绝对误差分别降低了 22.32%、17.98%。

Table 3. Performance comparison between mixed model and single model

表 3. 混合模型与单一模型性能比较

地点	评价指标	模型 1	模型 2	模型 3	LGBM	SVR	KNN
北京	RMSE	78.91	84.04	91.18	105.41	113.71	111.35
	MAE	51.76	60.96	64.49	68.51	94.31	73.91
重庆	RMSE	88.17	93.50	96.59	128.15	130.11	131.99
	MAE	62.35	68.75	69.18	93.23	101.12	98.01
哈尔滨	RMSE	92.49	96.05	97.53	120.34	125.12	125.55
	MAE	60.32	70.45	70.19	81.12	95.59	85.58

为了验证模态分解的特征提取的重要性, 将模型 1 与模型 4、模型 6 进行比较, 在三个地区上, 模型的性能表现如表 4 所示。经过模态分解和特征筛选的模型 1 在三个地区的预测精度都得到了提高。CEEMDAN 有助于消除序列中产生的非平稳和非线性缺陷, 且经过特征筛选后, 剔除冗余特征, 降低了数据复杂度。

本文将模型 1 与模型 5、模型 6 和 LGBM 模型比较, 在三个地区上, 模型的性能表现如表 5 所示, 通过对比, 验证利用 XGB 模型进行误差修正的有效性。经过误差修正的模型 1 和模型 6 在三个地区的预测精度都得到了改善, 模型 1 相比于模型 5 最小%RMSE (%MAE)提高了 3.32% (3.47%), 模型 6

相比于 LGBM 最小%RMSE (%MAE)提高了 8.23% (7.06%)。由此可见, 误差修正提高, 了模型的预测性能。

Table 4. The importance of modal decomposition and feature screening

表 4. 模态分解和特征筛选的重要性

地点	评价指标	模型 1	模型 4	模型 6
北京	RMSE	78.91	79.05	96.73
	MAE	51.76	51.95	63.67
重庆	RMSE	88.17	90.18	110.02
	MAE	62.35	64.45	77.25
哈尔滨	RMSE	92.49	95.33	99.49
	MAE	60.32	61.21	70.57

Table 5. The importance of error correction

表 5. 误差修正的重要性

地点	评价指标	模型 1	模型 5	模型 6	LGBM
北京	RMSE	78.91	81.62	96.73	105.41
	MAE	51.76	53.62	63.67	68.51
重庆	RMSE	88.17	95.00	110.02	128.15
	MAE	62.35	68.23	77.25	93.23
哈尔滨	RMSE	92.49	99.36	99.49	120.34
	MAE	60.32	64.34	70.57	81.12

5. 结论

本文提出了一种基于模态分解技术 CEEMDAN、特征提取技术 RF、机器学习模型 LGBM 和误差修正的混合模型。将提出的模型用于中国三个不同气候地区的 GHI 预测, 相比于单独的机器学习模型具有更高的精度, 模态分解将复杂的序列分解为多个模态函数, 随机森林对分解的模态函数进行特征提取, 降低数据的复杂度, 提高运算的效率, LGBM 模型相比于梯度增强决策树(Gradient boosting Decision Trees, GBDT)模型在运行时间加速和内存消耗缓解方面具有更好的表现, 比传统的机器学习模型具有更好的精度。

本研究的主要发现如下:

- 1) 本文所提出的模型有效提高了 GHI 的预测精度, 且在不同气候的三个地区上已进行验证, 相比于其他模型, 本文所提出的模型具有更高的预测精度。
- 2) 本文所提出的基于模态分解和特征提取的混合模型结构是有效的, 与单一模型相比较, 混合结构模型均具有更好的表现能力。混合结构模型的均方根误差(平均绝对误差)最高降低了 18.46% (18.71%)。
- 3) 通过对数据进行特征提取, 筛选重要特征, 达到了过滤冗余特征, 降低数据复杂度的目的。
- 4) 本文利用 XGB 模型进行误差修正, 相比于未进行误差修正的模型, 预测精度均有提高, 证明误差修正的有效性。

基金项目

贵州省数据驱动建模学习与优化创新团队项目(黔科合平台人才[2020] 5016)。

参考文献

- [1] 张金平, 周强, 王定美, 等. 太阳能光热发电技术及其发展综述[J]. 综合智慧能源, 2023, 45(2): 44-52.
- [2] 陈欢, 牟琰. “双碳”目标下重庆市碳排放影响因素及其预测研究[J]. 重庆工商大学学报(自然科学版), 2023, 40(2): 7-14.
- [3] 朱吉庆, 宋雨昂. 太阳能光伏发电技术发展现状与前景[J]. 对外经贸, 2024(1): 31-34+131.
- [4] Dedieu, G., Deschamps, P.Y. and Kerr, Y.H. (1987) Satellite Estimation of Solar Irradiance at the Surface of the Earth and of Surface Albedo Using a Physical Model Applied to Metcosat Data. *Journal of Applied Meteorology and Climatology*, **26**, 79-87. [https://doi.org/10.1175/1520-0450\(1987\)026<0079:SEOSIA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1987)026<0079:SEOSIA>2.0.CO;2)
- [5] Reikard, G., Haupt, S.E. and Jensen, T. (2017) Forecasting Ground-Level Irradiance over Short Horizons: Time Series, Meteorological, and Time-Varying Parameter Models. *Renewable Energy*, **112**, 474-485. <https://doi.org/10.1016/j.renene.2017.05.019>
- [6] 李国栋, 周扬, 李凯. 基于 SARIMAX-XGBoost 模型的区域能耗预测[J]. 电力信息与通信技术, 2022, 20(3): 26-33.
- [7] Ayodele, T.R., Ogunjuyigbe, A.S.O., Amedu, A., *et al.* (2019) Prediction of Global Solar Irradiation Using Hybridized k-Means and Support Vector Regression Algorithms. *Renewable Energy Focus*, No. 29, 78-93. <https://doi.org/10.1016/j.ref.2019.03.003>
- [8] Álvarez-Alvarado, J.M., Ríos-Moreno, J.G., Obregón-Biosca, S.A., *et al.* (2021) Hybrid Techniques to Predict Solar Radiation Using Support Vector Machine and Search Optimization Algorithms: A Review. *Applied Sciences*, **11**, 1044. <https://doi.org/10.3390/app11031044>
- [9] Gupta, P. and Singh, R. (2023) Combining Simple and Less Time Complex ML Models with Multivariate Empirical Mode Decomposition to Obtain Accurate GHI Forecast. *Energy*, **263**, Article ID: 125844. <https://doi.org/10.1016/j.energy.2022.125844>
- [10] Tercha, W., Tadjer, S.A., Chekired, F., *et al.* (2024) Machine Learning-Based Forecasting of Temperature and Solar Irradiance for Photovoltaic Systems. *Energies*, **17**, Article No. 1124. <https://doi.org/10.3390/en17051124>
- [11] Huang, N.E., Shen, Z., Long, S.R., *et al.* (1998) The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **454**, 903-995. <https://doi.org/10.1098/rspa.1998.0193>
- [12] Wu, Z. and Huang, N.E. (2009) Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, **1**, 1-41. <https://doi.org/10.1142/S1793536909000047>
- [13] Yeh, J.R., Shieh, J.S. and Huang, N.E. (2010) Complementary Ensemble Empirical Mode Decomposition: A Novel Noise Enhanced Data Analysis Method. *Advances in Adaptive Data Analysis*, **2**, 135-156. <https://doi.org/10.1142/S1793536910000422>
- [14] Torres, M.E., Colominas, M.A., Schlotthauer, G., *et al.* (2011) A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise. 2011 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, 22-27 May 2011, 4144-4147. <https://doi.org/10.1109/ICASSP.2011.5947265>
- [15] 王瑞, 马祯, 李磊. 基于 CEEMDAN-WOA-SVR 的高铁沿线超短期风速预测方法[J]. 中国铁道科学, 2023, 44(6): 80-86.
- [16] 张程珂, 刘会灯, 朱渝宁, 等. 基于多特征分析提取的随机森林超短期光伏功率预测[J]. 电力需求侧管理, 2023, 25(6): 50-56.