Title:	A hybrid machine learning method with explicit time encoding for improved Malaysian photovoltaic power prediction
Author(s) Name:	Hamza Mubarak, Ahmad Hammoudeh, Shameem Ahmad, Abdallah Abdellatif, Saad Mekhilef, Hazlie Mokhlis, Stéphane Dupont
Contact Email(s):	ahmad.shameem@aiub.edu
Published Journal Name:	Journal of Cleaner Production
Type of Publication:	Journal
Volume:	382 Issue 1
Publisher:	Elsevier
Publication Date:	01/01/2023
ISSN:	1996-1073
DOI:	10.1016/j.jclepro.2022.134979
URL:	https://www.sciencedirect.com/science/article/abs/pii/S0959652622045528
Other Related Info.:	Page 134979, ISI and Scopus indexed

Abstract:

Nowadays, with the growing interest in green energy, further improvements in photovoltaic (PV) power systems are needed. In this regard, the main aim is to find an optimal method to predict the output power of PV systems to maintain a sustainable operation. Hence, this work proposes a hybrid Machine Learning (ML) method LASSO-RFR for an hourly PV power output prediction. The model consists of Least Absolute Shrinkage and Selection Operator (LASSO) and Random Forest Regressor (RFR), where the former model makes a prediction and the latter fine tune it by the addition or subtraction of a relatively small value. The proposed model outperformed other models when tested on real data recorded from 2016 to 2019 for three Malaysian PV systems, namely Thin-Film (TF), Monocrystalline (MC), and Polycrystalline (PC). LASSO-RFR attained the lowest root mean square error (RMSE) of 23.7, 18.2, and 20.8 Wh/m2 for the TF, MC, and PC, respectively. This work also highlights the importance of explicit time encoding in improving PV power prediction. Although it is used to be ignored in the literature when developing ML models, the time feature is the second most influencing factor of PV power prediction after solar irradiance, as shown by the SHAP analysis (shapely additive explanation). For the study implications, the developed prediction model can assist the industry in predicting 1 h ahead of PV power output, demand-side management, and building operations and maintenance.