ECOLE HASSANIA DES TRAVAUX PUBLICS المدرسة الحسنية للأشغال العمومية

Green Energy Park

MOHAMMED VI POLYTECHNIC UNIVERSITY

IRESEN Institut de Recherche en Energie Solaire et en Energies Nouvelles

A Photovoltaic Power Prediction Approach Enhanced by Feature Engineering and Stacked Machine Learning Model

Ibtihal Ait Abdelmoula[a,b*] , Said Elhamaoui[b], Omaima Elalani[b], Abdellatif Ghennioui[b] , Mohamed El Aroussi[a]

[a] SIRC/LAGeS-EHTP, Hassania School of Public Works, Casablanca, Morocco [b] Green Energy Park - UM6P, Benguerir, Morocco

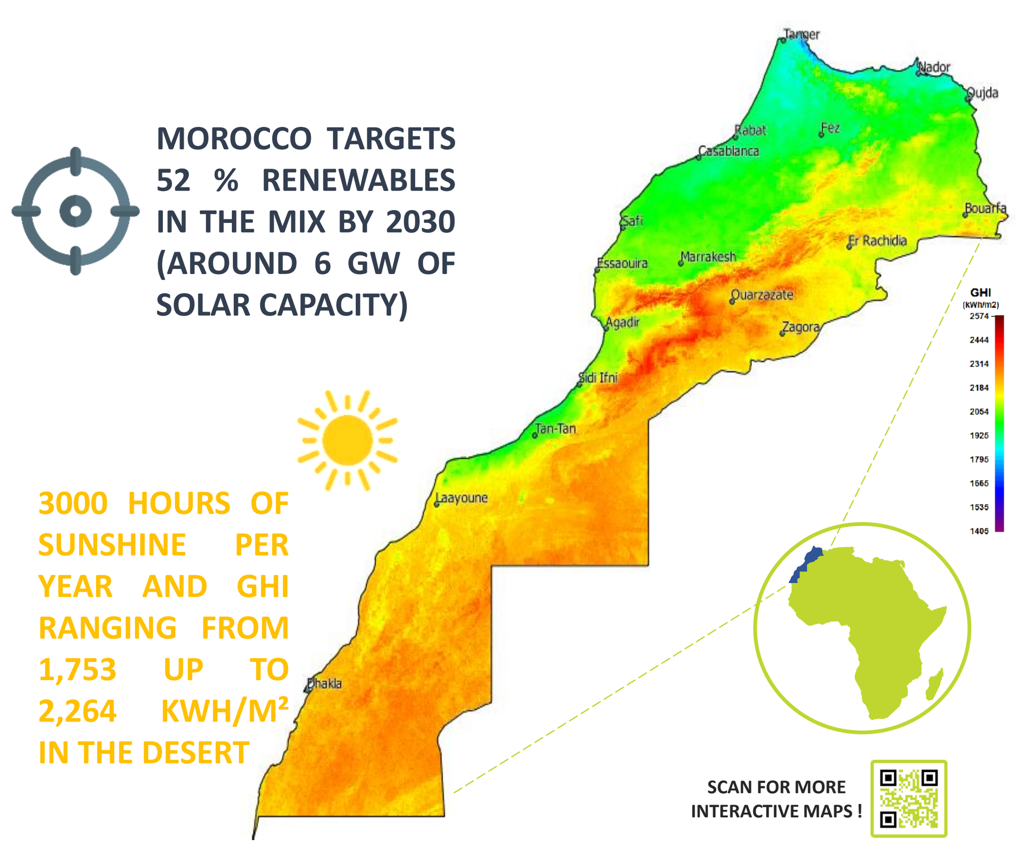
* Corresponding author. Tel.: +212 676170872 e-mail address: aitabdelmoula@greenenergypark.ma

ACCESS FULL PAPER

MOTIVATION

The Moroccan Solar Plan is considered one of the most ambitious renewable programmes in the African continent. Characterized by its high insolation and arid/semi-arid climate, Morocco relies on R&D to efficiently integrate renewables into its national grid. Understanding the distinctive features of each environment as well as how they affect photovoltaic power is the first step towards developing intelligent and data-driven maintenance algorithms suited to the region's environmental context. A Stacked Machine Learning Model is used in this study to make hourly power predictions of two PV systems varying in size, age and technology.

Figure 1 : Solar radiation map of Morocco

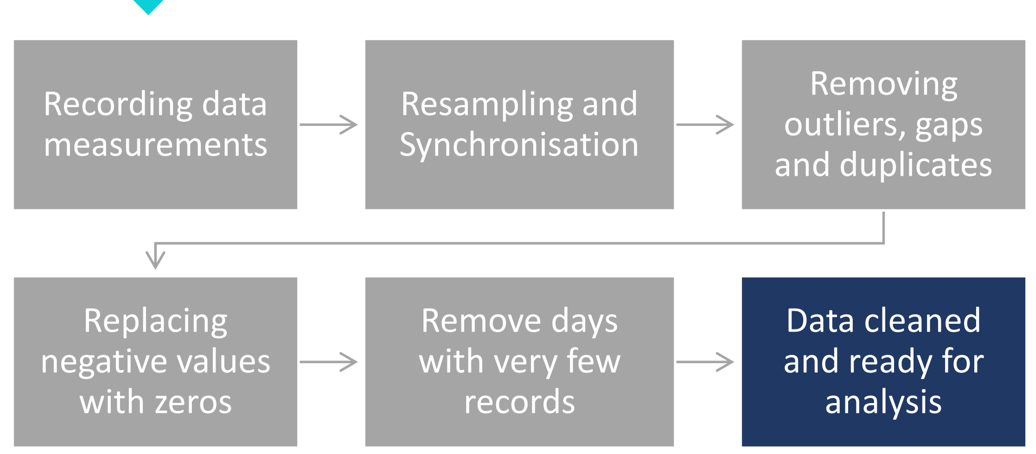


MATERIALS AND METHODS

The case study was conducted within the Green Energy Park (GEP) research and test platform. The studied PV systems have a total capacity of 16.56 kWp (System 1) and 5.34 kWp (System 2). The electrical and meteorological parameters are monitored with a 2-min resolution using adapted dataloggers. The dataset used in the study ranges from October 2021 to March 2022 and represents the winter and autumn seasons known for the high fluctuation in solar irradiance and prediction complexity.

Figure 2 : Representation of the testbed used in the case study

Figure 3 : Data Handling methodology



MODEL IMPLEMENTATION

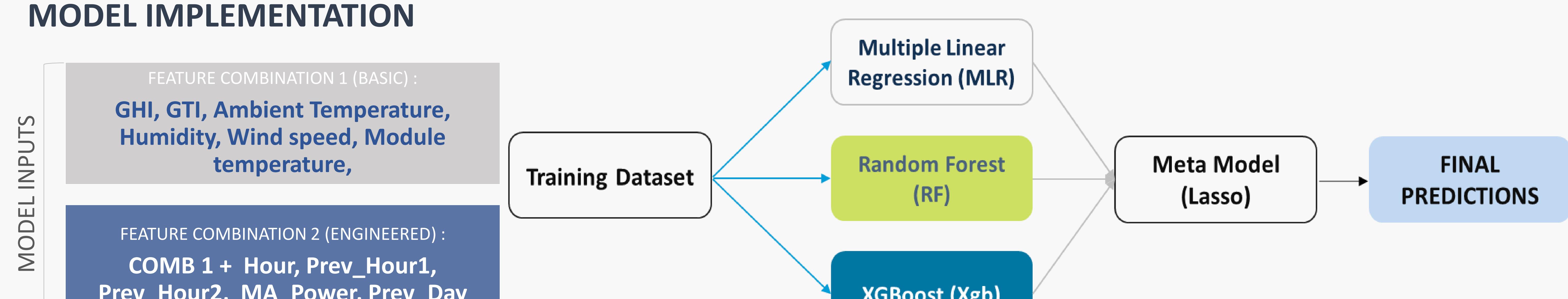
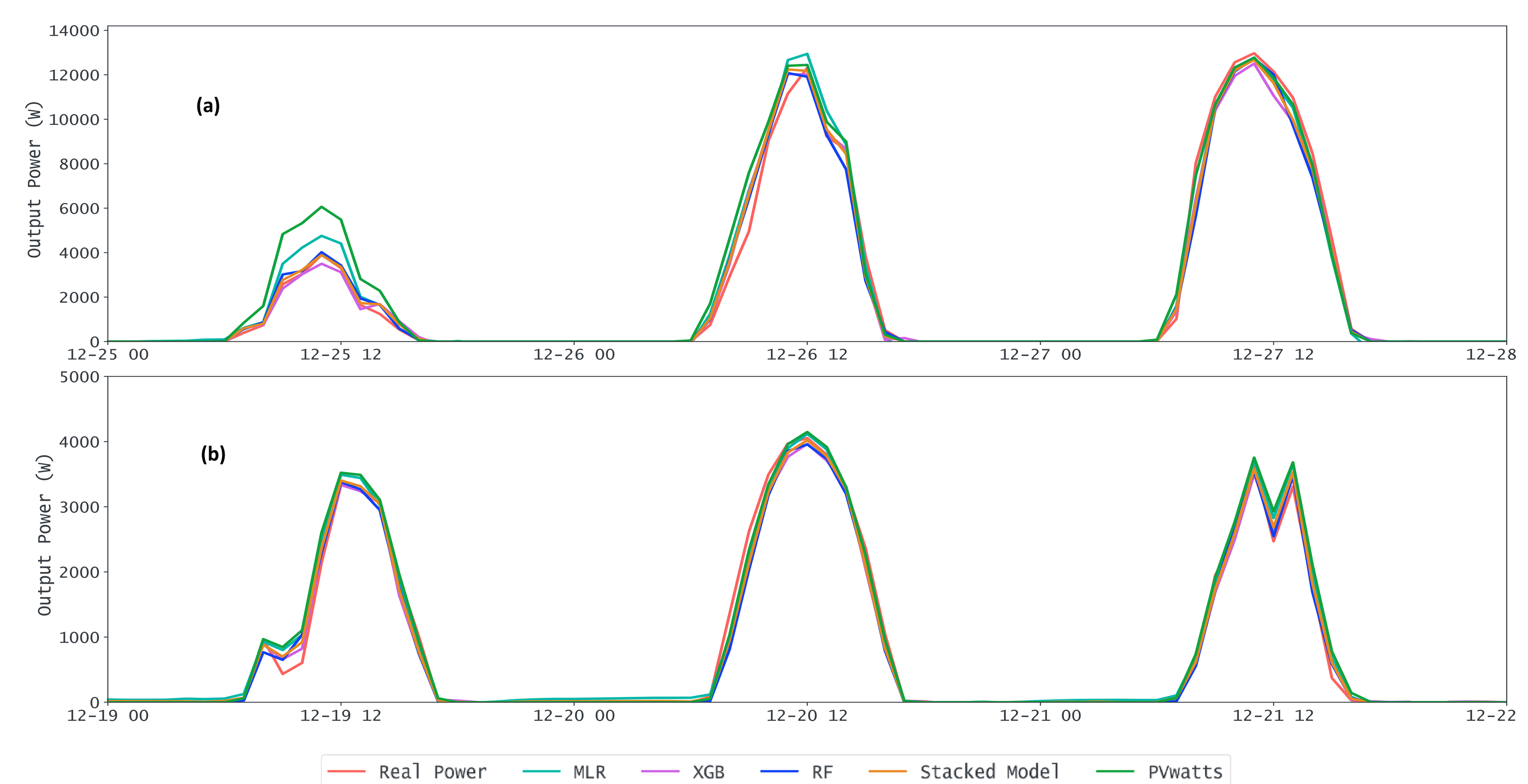
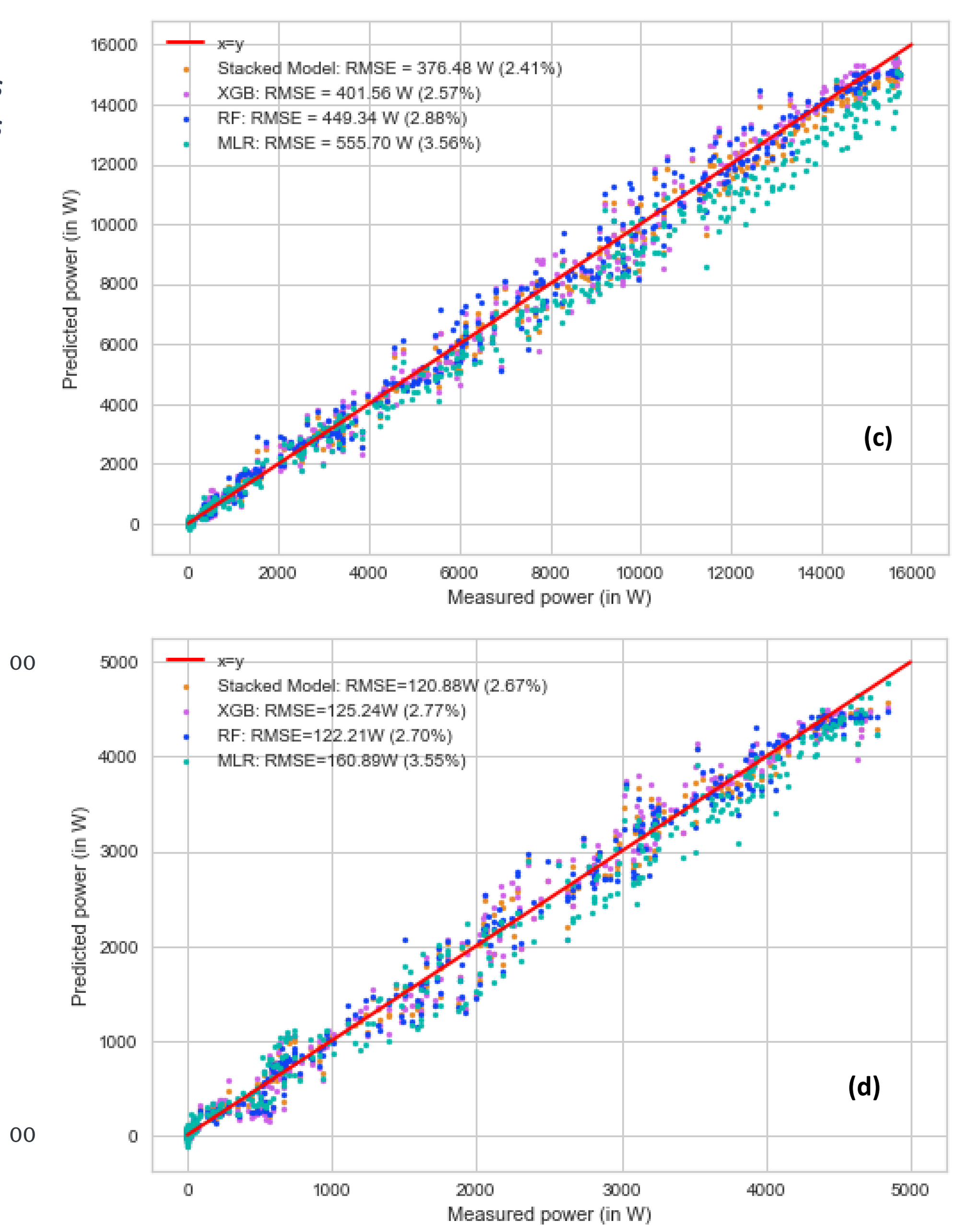


Figure 4 : Architecture of the stacked model proposed in the study

RESULTS

Figure 5 : Performance Comparison for the studied systems – For System 1 (a), the best results were found using the engineered features whereas for System 2 (b) the basic features gave better results. The plots show only the best-case scenario for each system. The scatter plots of measured vs predicted power for each model are shown in figures (c) and (d)





KEY FINDINGS

- Investigation of a model built employing stacked Machine Learning Algorithms (RF, XGBoost, MLR), historical electrical and meteorological data, and feature engineering of input parameters to enhance the model accuracy. **The stacked model gave better results than individual models scoring a RMSE of 386.31 W for system 1 and 131.54 W for system 2.**
- Validation of the model on diverse PV technologies, including differences in module type, plant size and exposure duration, with the aim of examining their impact on the PV power prediction.

ONGOING WORK

- Testing the model on a larger dataset to study the trade-off between the size of the dataset and the prediction performance on both systems.
- Evaluating the complexity and implementation feasibility of the models on an edge computing environment.