PHOTOVOLTAIC PREDICTION IMPROVEMENT VIA MACHINE LEARNING IN DAY-AHEAD AND NOWCASTING APPLICATIONS

B. Nam, M. Saleptsis, R. Ramaschi, A. Gandelli, S. Leva, M. Mussetta, E. Ogliari

Politecnico di Milano, Dipartimento di Energia, Sezione Elettrica
Via Lambruschini, 4, 20156 Milano (MI)

Research on Photovoltaic (PV) forecasting is currently focusing on the improvement of Day-Ahead (DA) and NowCasting (NC) predictions.

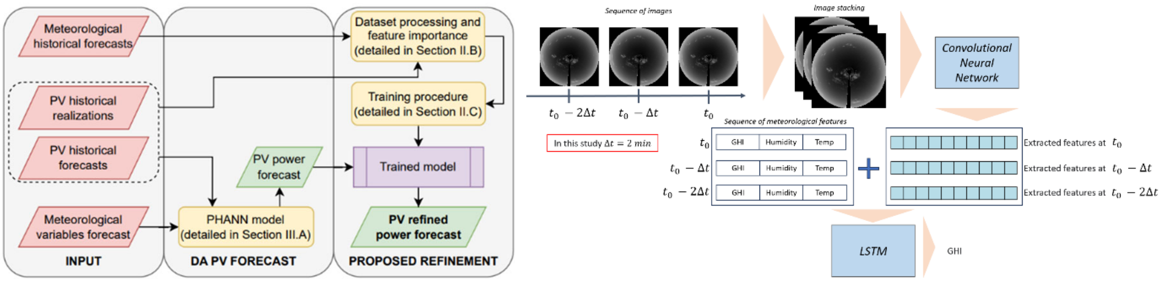

Fig. 1: Proposed methods for DA (left) and NC predictions (right)

PV power production is characterized by its inherent intermittency, primarily due to the strong influence of weather phenomena, which introduce uncertainties. DA 0 forecasters are fundamentals tools of modern energy systems, as they facilitate accurate scheduling and unit commitment procedures, thereby enhancing the efficiency and reliability of electrical grid. While many studies have been explored the development of DA PV forecasting, a novel approach of introducing a two-layer PV forecaster is developed in [2]. The main structure of this work presents a framework to refine PV forecasting by leveraging a data-driven technique to construct feature datasets and train the model. This approach utilizes widely available historical data, enabling the development of an efficient forecasting framework. Specifically, an eXtreme Gradient Boosting (XGBoost) algorithm is developed to be placed as the second layer of the main Physical Hybridized Artificial Neural Network (PHANN) model that is operating as the main DA PV forecasting tool. As it is illustrated in Fig. 1 (left), the high-level analysis of the model is separated into three parts. Firstly, the input is preprocessed. Secondly, the training procedure utilizes the historical PV and weather actual and forecasted data. Finally, a Bayesian optimized XGBoost algorithm is employed to find patterns between the two and exports the trained model, and the featured dataset. This third part utilizes the DA PV forecast data as an input and the trained model. The refinement script employs these three steps to refine the DA PV forecast and to export it to a new dataset. In the evaluation of the refinement yield of the XGBoost-based refinement model we established a comparative framework between the PHANN, and XGBoost, whose results are shown in Table 1 for a 6-months testing period.

Table 1: Model comparison results performed on MG²Lab (Politecnico di Milano) data

Metrics	PHANN	XGBoost Refinement
MAE (kW)	0.94	0.88
RMSE (kW)	1.86	1.86
EMAE (%)	21.63	20.23

These results – lower MAE and EMAE, equal RMSE – show the relevance of implementing a data-driven refinement to a well-established DA PV forecaster. For what concerns NC, accurate and reliable ultra short-term solar radiation forecasting is critical for the efficient operation and management of solar energy systems, improving grid reliability, and enabling real-time energy trading. A hybrid deep learning model combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks for ultra short-term (5-minute horizon) has been developed and tested for PV NC application. The model leverages spatiotemporal data, using a CNN to extract spatial features from sequential infrared All-sky Images (ASI) and an LSTM to capture temporal dependencies from both image sequences and meteorological measurements. ASIs provide high-resolution sky views, capturing cloud dynamics critical for short-term predictions [3]. The data used in this study originates from the SolarTechLAB at Politecnico di Milano (45.50°N, 9.16°E). It comprises 1-minute infrared ASI from Reuniwatt Sky InSight™ thermal infrared imager (640x480 pixels, 180° FOV), along with minute-averaged meteorological measurements (GHI, ambient temperature, and relative humidity) from an on-site weather station. The first step consists in a data preprocessing to uniform the input that is fed to the proposed model, integrating CNN and LSTM as shown in Fig. 1 (right). The CNN component extracts spatial features - representing cloud patterns and distribution - from a sequence of three 128x128 input images that are concatenated with the corresponding normalized meteorological features. This combined feature vector serves as the input to a two-layer LSTM network, which processes the sequence to capture temporal dependencies and long-term relationships between the evolving cloud patterns and weather conditions. This model produces as output the forecast over 5 minutes. To evaluate performance, the model was compared against a Persistence model benchmark, where GHI values are predicted to be equal to the last measurement over the forecast horizon. The results are shown in Table 2.

Table 2: Model comparison over the 1/7/2024 - 26/12/2024 testing period

Model	**MAE (W/m2)**	**RMSE (W/m2)**	**FS (%)**
CNN-LSTM (training: 29/6/2023-31/6/2024)	41.51	86.38	10.70
Persistence	44.30	96.74	0.0

This model achieves lower MAE and RMSE values, with a positive Forecast Skill (FS) of 10.7%, highlighting the potential of hybrid deep learning models for advancing solar nowcasting capabilities, contributing to better integration of renewable energy sources.[1]

Bibliography:

[1] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari, “A physical hybrid artificial neural network for short term forecasting of pv plant power output,” Energies, vol. 8, no. 2, pp. 1138–1153, 2015.

[2] M. Saleptsis, R. Ramaschi, M. Mussetta and S. Leva, “Extreme Gradient Boosting for DayAhead Photovoltaic Output Power Forecast Refinement”, accepted to IEEE PowerTech 2025, Kiel. To be indexed in IEEExplore.

[3] E. Ogliari, M. Sakwa, P. Cusa, "Enhanced Convolutional Neural Network for solar radiation nowcasting: All-Sky camera infrared images embedded with exogeneous parameters," Renewable Energy, vol. 221, 2024.

[1] This study was carried out within the MOST — Sustainable Mobility Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023) – SMART STORAGE PLUG Flagship Project. This manuscript reflects only the authors’ views and opinions, neither the European Union nor the European Commission can be considered responsible for them.