A SYSTEM-INTEGRATING AND ENERGY FORECASTING TOOL FOR RENEWABLE ENERGY SOLUTIONS

Summary

Accurate forecasting of photovoltaic (PV) system performance is increasingly critical as energy markets evolve. Such predictions enable aggregators and energy communities to participate in short-term markets, enhancing revenue and supporting grid stability. This study introduces a performance model capable of predicting PV output within a cluster using only internal generation data. The model is designed to operate independently of external data sources, ensuring robustness and simplicity. The objective of this work is to support energy communities and grid operators in energy planning. Incorporating clear-sky baselines, real-time data, and machine learning, the developed model advances short-term spatial forecasting to support smart and sustainable energy systems. Furthermore, it aims to scale forecasting capabilities from local clusters to a regional or island-wide level, such as Cyprus. This effort aligns with broader goals of fostering smart, sustainable urban development.

Keywords: solar energy markets, grid stability, PV power output, spatiotemporal forecasting

1. Introduction

The ability to accurately forecast photovoltaic (PV) system performance has become increasingly important with the evolution of contemporary energy markets. Reliable short- and mid-term predictions offer market participants the opportunity to unlock new revenue streams beyond traditional, often subsidized, models. These emerging opportunities not only encourage the wider deployment of PV technologies, supporting efforts to reduce carbon emissions, but also contribute to the stability of the energy system and facilitate greater integration of renewables into the grid. For example, system aggregators can access short-term capacity markets, while prosumers and energy communities can participate more effectively in demand response programs, provided that dependable short-term generation forecasts are available.

This study aims to develop a tool to support energy communities and grid operators by enhancing existing PV power output forecasting techniques. The focus is on integrating a broader set of PV generation data to scale forecasting from localized clusters to the island of Cyprus. A distinguishing feature of this approach is its independence from external data sources, relying exclusively on historical and real-time output data from systems within the cluster. The methodology offers several advantages: it is simple to implement (plug-and-play), requires no system metadata (only inverter-generated PV output) and maintains operational capability even when external data sources are unavailable. While it allows optional integration with additional data such as weather forecasts or satellite imagery, it is not dependent on them.

The proposed short-term forecasting models are built upon the authors' previously published work [1][2]. Specifically, the base model uses historical PV output to produce daily clear-sky predictions at the temporal resolution of the input data. This work extends the model by incorporating the effects of weather conditions, particularly cloud cover, which are treated as dynamic attenuation fields moving across space and time. This enhancement enables accurate spatiotemporal PV power output forecasting across the region.

2. Short-Term Power Output Forecasting Tool

A high-resolution solar database, incorporating historical and real-time PV power output data with data-driven spatiotemporal forecasting models for PV power output is the main aim of the study for a comprehensive management of distributed generation within the electricity grid. To achieve this, aggregated data from distributed PV systems across Cyprus will be collected and spatial forecasting techniques based on clear-sky baselines and real-time output will be applied. Forecast accuracy will be further improved through the integration of machine learning algorithms, enabling short-term energy generation forecasting at local and regional scales. It is important to note that the methods have been previously developed and validated separately by the authors on a smaller scale, limited area and fewer PV installations. The current work focuses on scaling the approach to larger geographic regions with increased system density and on enhancing the forecasting techniques to improve overall performance.

The proposed methodology is structured into three key steps and a detailed overview of each step is provided below:

1. **Data Collection**: involves collecting temporally and spatially PV data, retrieved via Application Programming Interfaces (APIs) linked to the inverters of participating partners in the PV-PREDICT project. Data collection is conducted through the respective APIs provided by each inverter manufacturer. To facilitate this process, a real-time data acquisition module is under development, designed to initiate requests at fixed intervals (every 5 or 15 minutes, depending on the manufacturer's specifications) to each active PV station, across all supported inverter brands. The retrieved data is then standardized (normalizing power to kilowatts and timestamps to UTC) before being stored in the central database.

2. **Model Development and Enhancement**: the collected data will be applied to established forecasting methodologies to enable accurate predictions of photovoltaic (PV) power output. These methods encompass the following components:

   (a) **Clear-sky PV power output estimation**, which serves as a baseline to represent the expected performance of a PV system under ideal, unobstructed sky conditions for a given location and time;

   (b) **Weather-induced spatial variability analysis**, which captures the influence of cloud movement on PV output across a spatial neighborhood, such as a cluster of systems within a region. By continuously monitoring multiple PV installations and evaluating deviations from their respective clear-sky baselines, the model can detect and interpret patterns of atmospheric attenuation, effectively inferring cloud movement dynamics across the area. To visualize these effects, an animated map comprising 96 timesteps was generated for each day of the study period, illustrating normalized PV energy production under clear-sky conditions. Figure 1 presents a snapshot from February 22, 2022, at 11:45. In the animations, blue areas represent low PV output, while beige areas correspond to maximum expected energy generation;

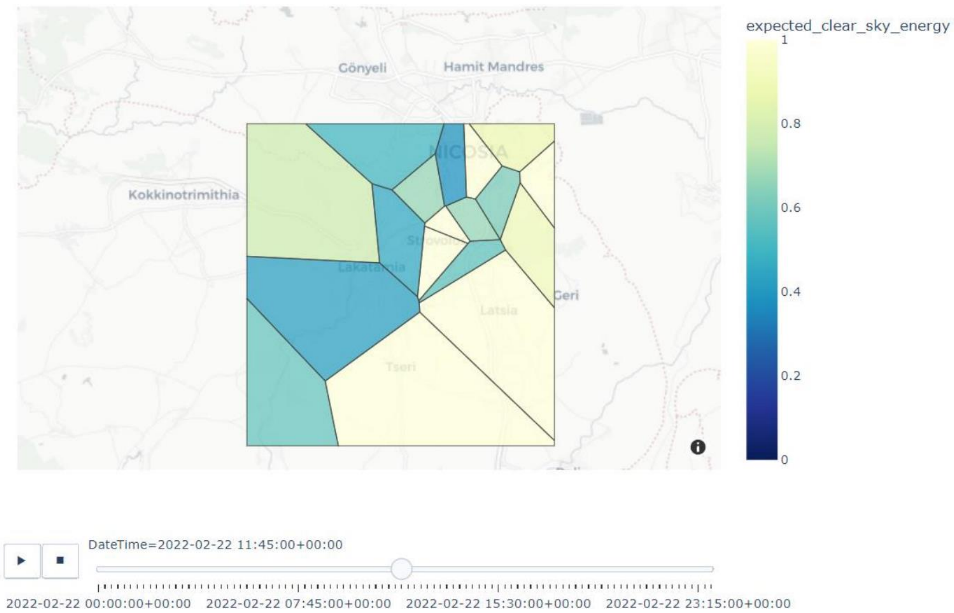

Figure 1: Snapshot of the animated map for February 22, 2022 at 11:45.

   (c) **Spatiotemporal PV power forecasting**, which integrates both spatial and temporal data to estimate power generation across a broader region. This approach is essential for the development of smart grids in urban environments, providing critical insights for solar energy planning, grid management, and the integration of renewable resources.

3. **Tool Creation and Application**: An online platform and database will be established, enabling the use of predicted output data by Transmission System Operators (TSOs) and PV plant managers to enhance grid stability and support further PV integration.

3. Results

A modular, standardized, robust, and freely accessible application architecture has been designed and developed to ensure scalability and portability. This architecture enables the continuous real-time connection of PV systems and the seamless transfer of energy production data from the facilities to a centralized database. The system is built using PostgreSQL/PostGIS for spatial data management, Django for the web framework, and GIS Geoserver for geospatial data visualization and services.

As of today, the database comprises data from 530 PV sites in Cyprus (Figure 2), encompassing both SolarEdge and Huawei installations. Geolocation data (longitude and latitude) for Huawei systems is directly obtained through the Huawei API, whereas for SolarEdge systems, geocoding services from Google and MapBox are employed to translate site addresses into geographic coordinates. The dataset extends back to May 14, 2013, with the bulk of the data concentrated between 2020 and 2025. Currently, the database contains over 30 million individual power measurements, which are continuously updated in real time.

Figure 2: Map showing the distribution of PV systems in Cyprus.

4. Conclusions

This work aims to enable holistic management of small distributed PV production within the electricity grid. The results of the individual methods demonstrate that accurate forecasting is achievable without reliance on external data sources. Currently, the first stage of the proposed methodology has been initiated and is under development. Overall, the results are promising, indicating the potential for a robust, standalone database that can support system aggregators and energy communities in participating in emerging energy markets. This approach can unlock additional revenue streams for PV systems while enhancing the integration capacity of renewable energy into the grid.

References

[1] Peratikou S. & Charalambides A.G. *Short-Term Pv Energy Yield Predictions within City Neighborhoods for Optimum Grid Management.* Energy and Buildings 2024, 323. http://dx.doi.org/10.2139/ssrn.4710380

[2] Peratikou S. & Charalambides A.G. *Estimating Clear-Sky PV Electricity Production Without Exogenous Data.* Solar Energy Advances 2022, 2. https://doi.org/10.1016/j.seja.2022.100015

This work was funded through the Cyprus Research and Innovation Foundation (Project ID: ENTERPRISES/ENERGY/1123/0014).