Towards a seamless approach for photovoltaïc forecasting

Thomas Carriere and Georges Kariniotakis
MINES ParisTech, Centre for processes, renewable energies and energy systems, France (thomas.carriere@mines-paristech.fr)

Trading of photovoltaic (PV) energy generation involves several decision making processes at different times with different objectives. For example, a PV power plant coupled with a Battery Energy Storage System (BESS) has to provide bids in the day-ahead electricity market, but can also provide ancillary services. On the delivery day, it can also participate in intra-day trading sessions, and must decide which quantity to charge or discharge from the BESS in real-time. These successive decision-making processes all require forecasts of the energy production level for different forecast horizons. Besides, such decisions are generally not taken for a single plant at a single location but for a collection of several geographically distributed plants.

However, the models and the inputs used for the different forecast horizons are often different. In situ measurements are more accurate for very-short term forecasts (real-time to one hour ahead forecasts), satellite data is used for short-term forecasts (up to 6 hours ahead), and Numerical Weather Predictions (NWP) are used for long-term forecasts (day-ahead and longer). Models also vary, with auto-regressive approaches being commonly used for very-short term forecasts, while longer forecast horizons use a wide range of machine learning models. PV producers have thus to develop and maintain numerous forecasting models for the different decision-making processes they are involved in, usually fitted for each power plant. This increases further the complexity of the decision-making processes.

In this work we propose a forecasting model that can use all the inputs mentioned before, and weights them according to the forecasting horizon. It can thus operate from very short-term to day-ahead forecast horizons with state-of-the-art performance. It can also directly provide probabilistic forecasts for an aggregation of power plants, thus allowing having a single forecasting model for managing a virtual power plant. The model follows the “lazy learning” paradigm, where generalization from the training set is only computed when a forecast is requested. Thus, the model is resilient to changes in the neighborhood of the plant (surrounding environment, partial outage, soiling, etc.). The model is based on the Analog Ensemble (AnEn) method. However it is structurally expanded to allow the method to use an arbitrary large number of inputs. Each input is then weighted depending on the forecast horizon, which allows dynamically selecting the most relevant inputs depending on the horizon.

The model is evaluated for short-term and day-ahead forecasts, and compared with a Quantile Regression Forest (QRF) and Bayesian Automatic Relevance Determination (ARD) for day-ahead
forecasts, and a linear Auto-Regressive Integrated Moving Average (ARIMA) model for short term forecasts. Results show that the AnEn model is competitive with the QRF and ARD models in day-ahead forecasting, while requiring less computational resources and without a need for regular retraining. It is also better than the ARIMA model for short-term forecasting. An evaluation conditional to the weather variability allow to assess the model performance in the best and worst condition.