



# Regional solar power forecasting

Report IEA-PVPS T16-01:2020

May – 2020

## Executive summary

### Editors

Elke Lorenz (Fraunhofer ISE, Germany), Carmen Köhler (p3r solutions, Germany), Jan Remund (Meteoest AG, Switzerland)

### Authors

*Alessandro Betti (i-EM, Italy), M. Pierro (Uni Roma 2, Italy), C. Cornaro (Uni Roma 2, Italy), D. Moser (EURAC, Italy), M. Moschella (i-EM, Italy), E. Collino (RSE, Italy), D. Ronzio (RSE, Italy), D. van der Meer (Univ. Uppsala, Sweden), J. Widen (Univ. Uppsala, Sweden), L. Visser (Univ. Utrecht, Netherlands), T. AlSkaif (Univ. Utrecht, Netherlands), W. van Sark (Univ. Utrecht, Netherlands)*

High levels of photovoltaic (PV) power penetration pose challenges to the operational performance of the power system. Regional forecasts of PV power allow transmission system operators (TSOs) and distribution system operators (DSOs) to take appropriate measures to maintain balance between supply and demand.

In this work, we compare the accuracy of several up-scaling methods for regional PV power forecast from 1 up to 3 days-ahead using three different case studies characterized by different sizes of the area of interest and located in two different countries:

- Forecast of the aggregated PV power output of 150 PV systems in the province of Utrecht (Netherlands) within a controlled area of  $1.4 \times 10^3 \text{ km}^2$ ;
- Forecast of the regional PV generation of six Italian macro zones (market zones) with controlled areas that range from  $24 \times 10^3$  to  $119 \times 10^3 \text{ km}^2$ ;
- Forecast of the regional PV generation of the whole of Italy with a controlled area of  $300 \times 10^3 \text{ km}^2$ .

The up-scaling approaches are applied to map meteorological forecasts of numerical weather prediction (NWP) models to regional PV power generation. All the up-scaling methods shown here directly predict the regional PV power generation, i.e. they consider the PV power output of the whole PV fleet as if it had been produced by a single “virtual” solar power plant, rather than predicting PV power for representative PV power plants as a basis to predict regional PV power.

In order to improve regional PV power forecasts and to advance the state-of-the-art, it is important that researchers have access to identical datasets, so they may compare their new methods. To that end, this report presents three datasets for Italy and Utrecht, the Netherlands. More specifically, for

Italy, the datasets are made of satellite derived global horizontal irradiance data, numerical weather forecasting of some variables affecting PV production and corresponding PV power data. The Utrecht dataset is comprised of NWP forecasts and aggregated PV power measurements of 150 systems. These datasets have been cleaned in order to be suitable to test different PV power forecasting methods.

The focus of this work is on the comparison of different PV power up-scaling methods, that have been performed on the aforementioned datasets. Since accuracy of a PV power forecast model is greatly affected by the input NWP data, the same NWP forecast has been used within each case study. Among the benchmarks methods there are artificial neural networks, analog ensembles, cooperative ensembles of machine learning algorithms, quantile regression forests, random forests and gradient boosted regression trees. These methods use state-of-the-art machine-learning approaches and allow researchers to compare their new models against each other. It should be noted that probabilistic forecast evaluation is not in the scope of this report, although several models presented herein are capable of producing probabilistic forecasts. Besides the description of the benchmark models, the report also describes several feature engineering techniques and pre-processing procedures aimed at improving the accuracy of the benchmarks presented herein. The accuracy of the benchmark models is assessed extensively, both quantitatively and qualitatively, allowing researchers and engineers to thoroughly compare their new methods against these benchmarks.

For the PV power forecasts at the zonal levels of Italy, the analog ensemble model achieves a root mean squared error (RMSE) skill score between 20% and 36% and a mean bias error (MBE) between -1.9% and 1.5% for all forecast horizons (i.e., 1-3 days ahead). It should be noted that the forecasts and measurements are normalized with respect to the installed capacity in the area throughout this report. Considering the 24h ahead forecasts of the entire PV fleet of Italy, the benchmarks achieve an RMSE skill score between 28% and 47%, while the MBE lies between 0.1% and 1.9%. The so called “smoothing effect” has been also investigated over Italy, demonstrating that the RMSE decreases with the root of the area under control following an exponential or hyperbolic trend and decreasing from almost 5.5% of the installed capacity on average over market zones down to 3.6% over Italy. In addition, blends of the models are also applied to this dataset, which improves the RMSE skill score from 47% (the best model) to 51% using a nonlinear blending technique. Finally, the benchmark models achieve RMSE skill scores between 42% and 44% on the PV power production forecasts over the province of Utrecht, the Netherlands, while the MBE lies between 0% and 0.8%.

The comprehensive analysis presented herein shows to the interested reader the importance of carefully tuning all the different up-scaling steps in the PV power forecast chain in order to achieve the highest accuracy.