

# Forecasting regional wind production based on weather similarity and site clustering for the EEM20 Competition

**Authors: Simon Camal; Kevin Bellinguer; Valentin Mahler; George Kariniotakis**

MINES ParisTech, PSL University,  
Centre PERSEE (Processes, Renewable Energies and Energy Systems),  
Sophia Antipolis, France.

## Summary

Precise probabilistic forecasting tools of wind power generation are essential for solving problems related to the impact of wind generation uncertainty on electrical systems. In this context, the International Conference on the European Energy Market EEM20 set up a competition aiming at forecasting wind power generation for the four Swedish bidding zones. Participants had to generate day-ahead quantile forecasts of the aggregated wind power production from Numerical Weather Predictions (NWP) ensembles provided all over the country [1].

This paper presents the methodology adopted to develop a robust and adaptive model, that has been progressively improved and ultimately won the competition [2]. This approach provides a forecasting tool within the frame of the Horizon2020 Project Smart4RES which addresses a wide range of scales from one wind turbine to regional pools [3].

The proposed approach is based on a physics-oriented pre-processing of regional data, meaning that it is intended to deal with spatial data while making good use of the physics associated to power conversion at the scale of wind turbines. First, weather patterns (i.e. wind amplitude at hub height) are captured at a regional scale by means of an analog-based approach considering spatially distributed NWP [4]. Second, to tackle the high dimensionality issue induced by the NWP data (i.e. nearly 840000 data points are available for each timestamp) a k-means clustering approach is implemented to reduce the spatial distribution of grid points while avoiding the possible redundancy of weather information.

A special focus was given to the relevance of weather variables (e.g. temperature and relative humidity to account for possible icing effects) as well as wind-power-related features (e.g. wind speed vertical profiles and air density to represent wind-to-power conversion process). Predictions are obtained via a Quantile Regression Forests model, a well-proven probabilistic forecasting model in the field of aggregated renewable production [5].

Regarding future work, new physics-oriented features could be proposed (e.g. modelling the ageing effects of wind turbines) as well as other analog predictors. Lastly, the optimal

combination of forecasting models proposed by [6], which ranked second in the competition, could be complemented with the predictions of the present model.

## References

- [1] EEM20 Forecasting Competition [Online]. Available: <https://eem20.eu/forecasting-competition>
- [2] K. Bellinguer, V. Mahler, S. Camal, and G. Kariniotakis, "Probabilistic Forecasting of Regional Wind Production for the EEM20 Competition: a Physics-oriented Machine Learning Approach," International Conference on the European Energy Market, EEM, 2020.
- [3] "Smart4RES - data science for renewable energy prediction." [Online]. Available: <https://www.smart4res.eu>
- [4] K. Bellinguer, R. Girard, G. Bontron, and G. Kariniotakis, "Short-term Forecasting of Photovoltaic Generation based on Conditioned Learning of Geopotential Fields," 55th International Universities Power Engineering Conference (UPEC), 2020, p. 6.
- [5] S. Camal, F. Teng, A. Michiorri, G. Kariniotakis, and L. Badesa, "Scenario generation of aggregated Wind, Photovoltaics and small Hydroproduction for power systems applications", Applied Energy, vol. Volume 242, pp. 1396–1406, May 2019.
- [6] J. Browell, C. Gilbert, R. Tawn, and L. May, "Quantile Combination for the EEM20 Wind Power Forecasting Competition", International Conference on the European Energy Market, EEM, 2020