Data Science for Next Generation Renewable Energy Forecasting - Highlight Results from the Smart4RES Project

George Kariniotakis[1], Simon Camal[1], Fabrizio Sossan[1], Bijan Nouri[2], Jorge Lezaca[2]
Matthias Lange[3], Bastien Alonzo[4], Quentin Libois[4], Pierre Pinson[5], Ricardo Bessa[6], Carla Goncalves[6]

Abstract—Smart4RES is a European Horizon2020 project developing next generation solutions for renewable energy forecasting. This paper presents highlight results obtained during the first year of the project. Data science is used throughout the proposed solutions in order to process the large amount of heterogeneous data available to forecasters, and derive model-free approaches of forecasting and decision-aid tasks. This paper presents a series of solutions addressing relevant for Photovoltaics (PV) and storage applications. High-resolution Numerical Weather Predictions and regional solar irradiance forecasting provide detailed information on local weather conditions and their variability. PV power forecasting benefits from such new data sources, but also the proposed collaborative data exchange. Finally, data-driven methods simplify decision-making for trading in short-term markets and for grid management.

Index Terms—Renewable energy, Forecasting, High resolution, Data Market, Optimization, Grid management

I. INTRODUCTION

In this paper we present the research directions and innovative solutions developed in the European Horizon2020 project Smart4RES (http://www.smart4res.eu) to better model and forecast weather variables and production of renewable energy sources (RES) like wind, solar and run-of-the-river hydro in order to optimise their integration into power systems and electricity markets. Smart4RES gathers experts from several disciplines, from meteorology and renewable generation to market- and grid-integration. It aims to contribute to reach very high RES penetrations in power grids of 2030 and beyond, through thematic objectives including:

1) Improvement of weather and RES forecasting
2) Streamlined extraction of optimal value through new forecasting products, data market places, and novel business models
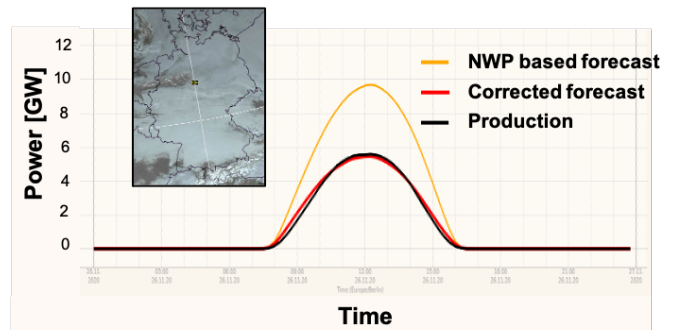3) New data-driven optimization and decision-aid tools for market and grid management applications.

Fig. 1. Impact of unpredicted fog event on the forecast of total PV production in Germany (source: EMSYS)

4) Validation of new models in living labs and assessment of forecasting value vs costly remedies (i.e. storage) to hedge uncertainties.

Forecasting of weather variables and of Renewable Energy Sources (RES) production are both mature disciplines. However, specific weather conditions remain challenging to forecast with existing models, leading to economic impacts on power systems: Figure 1 shows large forecasting errors on the total PV production of Germany, caused by an unpredicted fog event. The imbalance costs associated with these errors reached around 1.6 MEuro for the day. Smart4RES aims at improving weather and RES forecasting (+10-15% in performance), via the development of high-resolution Numerical Weather Predictions (NWP) and the exploitation of local measurements. In particular, an optimized treatment of weather forecast ensembles has the potential to decrease forecasting errors on solar irradiance as found e.g. in [1]. For problems concerned with solar irradiance forecasts at the next minute to the next hours, all-sky-imagers (ASI) provide valuable information on cloud movement above a PV site. Existing approaches consider a limited number of ASIs surrounding the site [Blum], so that valuable information about weather variability at a regional scale is ignored. Finally, Large Eddy Simulation has the potential to resolve local turbulent airflows and reproduce variability over a wind or PV plants [2].

Previous research projects on renewable forecasting established state-of-the-art approaches for forecasting of renewable production considering its application to power systems and electricity markets [3], [4]. However to the author's knowl-

edge, the added value of a new RES forecasting product integrating seamlessly high-resolution NWP inputs has not been established in the existing literature. RES forecasting accuracy also improves with the integration of spatio-temporal information of production from surrounding plants or weather stations [5]. In order to implement such models harvesting local data, valorization schemes of data sharing between distributed RES plants under privacy constraints must be investigated.

Decision-aid models under RES uncertainty integrate RES forecasts in various way, the most common being the 'Forecast-then-Optimize approach': the optimization of trading positions or grid management decision is based on forecasts generated by a forecasting model which has been separately trained beforehand. The resulting chain from data to decision is complex, especially in the presence of multiple RES production sites and the addition of other uncertain variables such as market quantities or load consumption. The simplification of the modelling chain is a major challenge for data-driven decision-aid methods that aim to integrate forecasting into optimization [6] or even directly derive optimization decisions without forecasting [7].

Developments in the project on the challenges mentioned above have been formalized in Use-Cases that cover a large range of time frames, technologies and geographical scales. This paper focuses on the original contributions of Smart4RES for improving the forecasting of weather variables and solar production:

- innovative measuring set-ups (i.e. a network of sky cameras in Germany);
- high-resolution Numerical Weather Predictions
- new PV forecasting products using these advanced weather predictions

Finally, we assess how the new forecasting products may bring value to applications and be improved by collaborative analytics.

II. DATA SCIENCE SOLUTIONS FOR RENEWABLE FORECASTING

A. Advanced Weather Predictions for Improved Renewable Forecasting

Smart4RES develops advanced weather forecasts that may be used in solar and storage applications, providing predictions at high spatio-temporal resolution.

First, *NWP ensembles at high-resolution* (1km-5min) provide detailed information about the uncertainty associated with weather forecasts: Figure 2 shows the standard deviation of solar irradiance ensembles for a winter day in France, with significant variations between regions and within a specific area of a few kilometers surrounding a renewable site. Then, RES forecasting models may combine these ensembles with a second weather input computed also at a large scale, namely *regional solar irradiance forecasts* derived from a network of all-sky-imagers. A network of ASIs and ceilometers called Eye2Sky is being deployed in the Oldenburg region in Germany, which will reach 38 stations at full extent. The high
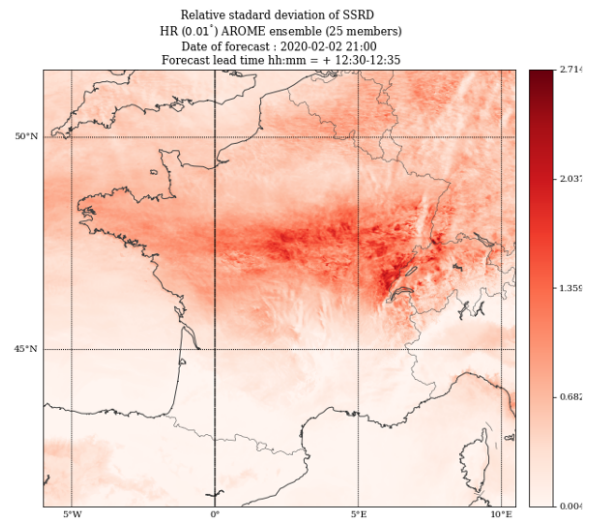

Fig. 2. Standard deviation of solar irradiance ensemble forecasts on a winter day (Source: MeteoFrance)
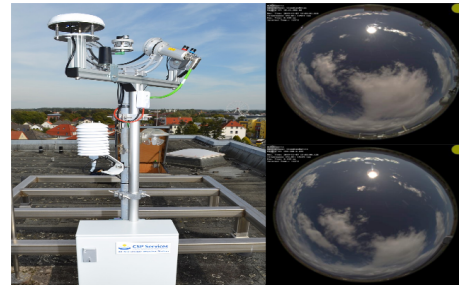

Fig. 3. Ground measurement station and ASI images

spatial density of ASI pairs enables to reduce errors associated with the inference of cloud base height compared to a single pair of ASI [8]. A probabilistic procedure integrates cloud base heights and cloud segmentation relying on machine learning to produce nowcasted irradiance maps, such as the Global Horizontal Irradiance map surrounding a PV site shown in Fig. 4. The predicted irradiance can then be imported in a PV forecasting model. First results indicate increased accuracy compared to a model using conventional irradiance measurement limited to the production site. Finally, another advanced weather forecast proposed by Smart4RES is *Large Eddy Simulation (LES)* applied on RES production site. The LES produces weather forecasts at a 100m-30s resolution accounting for effects of local climate and terrain that cannot be modelled by NWP. The high-frequence variability of weather conditions forecasted by LES is a valuable input for RES prediction models. It can even apply to small islands or parts of regions, providing spatio-temporal information for distributed RES sites.

B. Data-driven decision-aid tools

As mentioned in the Introduction, standard modelling chain of decisions under RES uncertain are complex because they rely on RES forecasting and explicit models of objectives
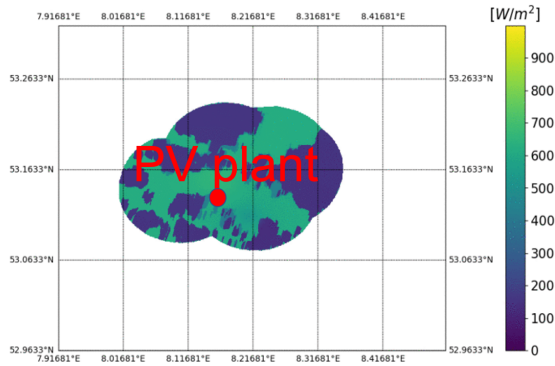
Fig. 4. Irradiance nowcasting obtained by ASI network (Source: DLR)

and constraints, that are consequently solved via stochastic or robust optimization techniques. Besides complexity, two important drawbacks arise: (1) computational time is high when constraints are multiple and variables are coupled, and (2) the resulting modelling chain can be seen as a black-box. Smart4RES proposes to simplify the chain by applying machine learning or operational research approaches which enable to take decisions based on multiple criteria and provide some level of decision interpretability.

A data-driven approach for the predictive management of distribution grids is being developed by INESC TEC. The approach is based on an estimation of the grid sensitivity to uncertain RES production / uncertain load via impedance and admittance bus models. A machine learning model is then trained on the output of these sensitivity models to reproduce the sensitivity in various operational conditions. A more in-depth description of the approach is presented in [9]. This approach is appealing not only for its simplicity of use compared to conventional stochastic optimization of power flows, but also for the interpretability of its results: Fig. 5 shows that obtained sensitivity coefficients of all lines δ_{i-j} are organized in a causal graph enabling to quantify the impact of flexibility activation or of RES/load uncertainty on the congested branch. By doing so, proposed flexibility activations can be easily linked with their local predicted effect on the grid.

Data-driven methods are also able to simplify the mod-elling chain in trading problems. The traditional Forecast-then-Optimize (FO) approach can be replaced by an integrated fore-casting and optimization method which is based on predictive prescriptions (PP). PP are able to derive a weighted sample average approximation of constrained optimization problems represented in Fig. 6. The weights are obtained by fitting a Machine Learning model on the available data (e.g. for trading, market quantities, weather and RES variables) and the decision cost is optimized using a numerical optimization. Being a single model, it is easier to interpret than a series of forecasting model and stochastic optimization. The method has been applied to the trading of the renewable production of

a Virtual Power Plant (VPP) on a day-ahead energy market. Results in Fig. 7 show that the choice of the local weighting algorithm has an influence on the expected trading cost: the cost obtained with the PP method is minimized for the discriminative algorithm Random Forest (RF), whereas k-Nearest Neighbours (kNN) and Kernel Regression (KR) are performing worse than the standard FO method based on VPP production forecasts with RF. The PP method achieves similar result than the FO method even for testing sets of limited sizes, and results converge for a size above 5000 points.

C. Added value of new forecasting products in applications

New forecasting products of weather and RES production shown above can be improved by sharing data with agents pro-ducing local data streams, e.g. of measurements or forecasts. Instead of sharing directly sensitive information, each agent shares with the others encrypted data (Fig. 8). The encrypted data corresponds to distributed regression coefficients, that are conciliated in a further step to consistently model the spatio-temporal information conveyed by the shared data [10]. An iterative process ensures sufficient accuracy of conciliation and predictions. Coefficients adapt to the evolution of weather/RES production, without ever sharing private data. First results obtained on distributed PV in a Portuguese district show an average relative improvement of Normalized RMSE of 5%-10% of privacy-preserving collaborative LASSO-Vector AutoRegressive (VAR) model vs a standard autoregressive model [10]. The value of improvement, which is variable among all agents, can be monetized for all participating agents via a specific data market [11].

In the data-driven grid management approach presented previously, the forecast of the grid sensitivity enables to learn the sensitivity depending on operational conditions without the need of an explicit optimization model of power flows in the grid, hence potentially reducing the costs associated with the purchase and maintenance of power flow simulations.

III. Conclusion

Accurate forecasts of weather and variable generation are key for accurate decision-making. High-resolution weather forecasts are developed in Smart4RES thanks to advanced simulation models at large scale (Ensemble NWP) and small scale (LES), and optimally combined distributed weather measurements for solar nowcasting. Collaborative forecasting investigates the improvement associated to local data sharing between distributed RES plants. This data sharing paves the way to a data market where agents exchange measure-ments, predictions or other types of valuable data. Lastly, Smart4RES proposes data-driven approaches that streamline decision-making by simplifying the model chain of bidding RES production, storage dispatch or predictive management electricity grids. They also provide interpretable hindsight to decision-makers by integrating the decisions of experts (human-in-the-loop) and will be tested in realistic laboratory conditions (software-in-the-loop).
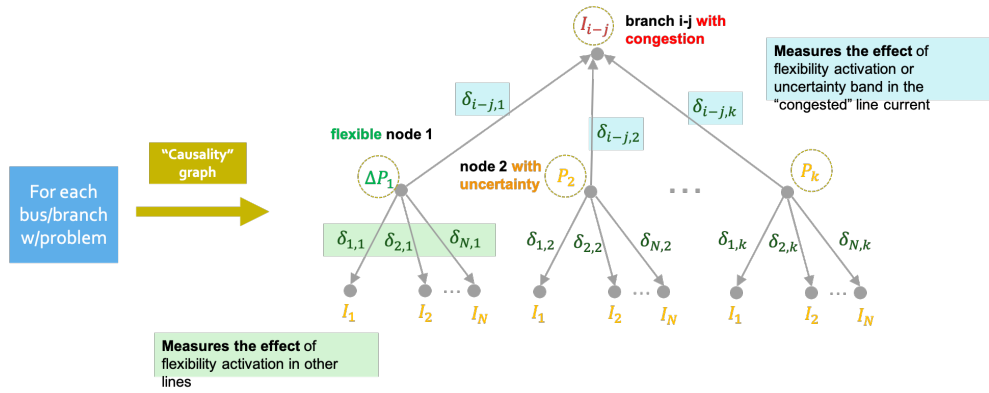
Fig. 5. Causality graph representing the effect of predicted flexibility and uncertainty on grid constraints (Source: INESC TEC)
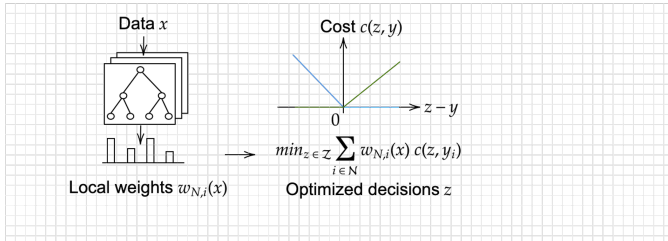

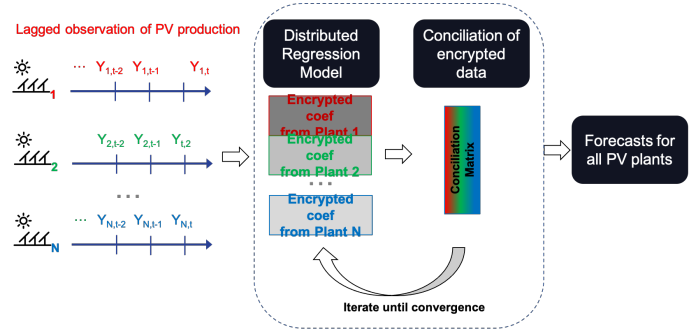Fig. 6. Principle of Predictive Prescriptions (Source: ARMINES - MINES ParisTech)

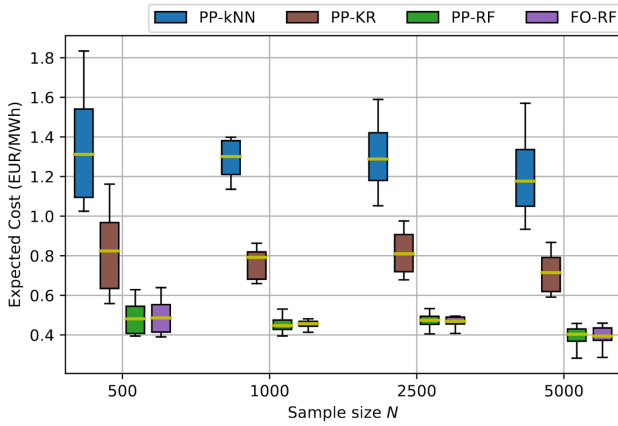Fig. 8. Principle of Privacy-Preserving collaborative Forecasting

Fig. 7. Results of Predictive Prescriptions for the trading of a renewable VPP (Source: ARMINES - MINES ParisTech)

REFERENCES

[1] J. Thorey and J. Thorey, "Ensemble forecasting using sequential aggregation for photovoltaic power applications To cite this version : HAL Id : tel-01697133 Thèse de doctorat de l ' Université Pierre et Marie Curie École doctorale de Sciences Mathématiques de Paris Centre par," 2018.

[2] C. Gilbert, J. W. Messner, P. Pinson, P.-J. Trombe, R. Verzijlbergh, P. van Dorp, and H. Jonker, "Statistical post-processing of turbulence-resolving weather forecasts for offshore wind power forecasting," *Wind Energy*, vol. 23, no. 4, pp. 884–897, 2020.

[3] G. Sideratos and N. D. Hatziargyriou, "Wind power forecasting focused on extreme power system events," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 3, pp. 445–454, 2012.

[4] G. Kariniotakis, I. H. Waldl, I. Marti, G. Giebel, T. S. Nielsen, J. Tambke, J. Usaola, F. Dierich, A. Bocquet, and S. Virlot, "Next generation forecasting tools for the optimal management of wind generation," *2006 9th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS*, 2006.

[5] X. G. Agoua, R. Girard, and G. Kariniotakis, "Short-Term Spatio-Temporal Forecasting of Photovoltaic Power Production," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 538–546, 2018.

[6] J. Filipe, R. J. Bessa, M. Reis, R. Alves, and P. Póvoa, "Data-driven predictive energy optimization in a wastewater pumping station," *Applied Energy*, vol. 252, no. May, p. 113423, 2019.

[7] T. Carriere and G. Kariniotakis, "An Integrated Approach for Value-oriented Energy Forecasting and Data-driven Decision-making. Application to Renewable Energy Trading," *IEEE Transactions on Smart Grid*, vol. PP, no. c, pp. 1–1, 2019.

[8] N. B. Blum, B. Nouri, S. Wilbert, T. Schmidt, O. Lünsdorf, J. Stührenberg, D. Heinemann, A. Kazantzidis, and R. Pitz-Paal, "Cloud height measurement by a network of all-sky imagers," *Atmospheric Measurement Techniques*, vol. 14, no. 7, pp. 5199–5224, 2021.

[9] R. Bessa, L. Teixeira, and H. Mariji, "Integration of wind energy uncertainty in predictive grid management: can we make it simple?," in *Wind Energy Science Conference*, 05 2021.

[10] C. Goncalves, R. J. Bessa, and P. Pinson, "Privacy-preserving distributed learning for renewable energy forecasting," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 3, pp. 1777–1787, 2021.

[11] C. Goncalves, P. Pinson, and R. J. Bessa, "Towards Data Markets in Renewable Energy Forecasting," *IEEE Transactions on Sustainable Energy*, vol. 3029, no. c, pp. 1–1, 2020.