



HAL
open science

Short-term Forecasting of Photovoltaic Generation based on Conditioned Learning of Geopotential Fields

Kevin Bellinguer, Robin Girard, Guillaume Bontron, Georges Kariniotakis

► **To cite this version:**

Kevin Bellinguer, Robin Girard, Guillaume Bontron, Georges Kariniotakis. Short-term Forecasting of Photovoltaic Generation based on Conditioned Learning of Geopotential Fields. 55th International Universities Power Engineering Conference - Virtual Conference UPEC 2020 – "Verifying The Targets", Sep 2020, Torino (Virtual Conference), Italy. hal-02932018

HAL Id: hal-02932018

<https://hal.science/hal-02932018>

Submitted on 7 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Short-term Forecasting of Photovoltaic Generation based on Conditioned Learning of Geopotential Fields

Kevin Bellinger[†], Robin Girard[†], Guillaume Bontron[‡] and Georges Kariniotakis[†]

[†]MINES ParisTech, PSL University, PERSEE - Centre for Processes, Renewable Energies and Energy Systems, CS 10207, rue Claude Daunesse, 06904 Sophia-Antipolis Cedex, France

Email : kevin.bellinger@mines-paristech.fr, robin.girard@mines-paristech.fr, georges.kariniotakis@mines-paristech.fr

[‡]Compagnie Nationale du Rhône, 2 Rue André Bonin, Lyon, France

Email : g.bontron@cnr.tm.fr

Abstract—Due to environmental concerns, renewable energy sources (RES) play an increasingly important role in the energy mix. In France, from 2018 to 2019, an increase of 21.2% and 7.8% of energy production was observed for wind and solar respectively [1]. RES are characterized by high variability and limited predictability, mostly due to their dependence on meteorological factors. This variability presents challenges for RES integration into grids and electricity markets: as the penetration of RES increases, power system balancing becomes more complex, and congestions may occur in the grid. This lack of predictability can also have financial consequences. In some European countries, energy producers have to pay penalties proportional to the forecasting error of the injected power. To address these challenges, it is important to accurately predict the future amount of energy production. In this paper we propose a spot statistical forecasting model for very short-term time horizons (from a few minutes up to 6 hours ahead). This model is based on a combination of heterogeneous inputs with a conditioned learning approach. Spatio-temporal inputs (measurements from geographically distributed PV sites and satellite images) are used to enhance short-term predictability, while a weather analog approach enables adaptability to changes in meteorological conditions by considering the most relevant past observations. The performance evaluations are carried out on a case study featuring nine PV plants located in France, over a one-year period.

Index Terms—Short-term solar power forecasting, analog approach, conditional forecast, spatio-temporal, auto-regressive processes, smart grid.

I. INTRODUCTION

Over the past years, environmental awareness, policies to promote sustainability, and even a reduction in production costs have led to the development of RES in many countries. Nevertheless, RES remain a challenging technology for transmission and distribution system operators. As a weather-dependent source of energy, RES are characterized by high variability. As a result, high penetration rates into the electrical grid can jeopardize the safe and economic operation of the power system. Apart from security concerns, fluctuations in RES can also induce financial penalties for power producers. In the coming years, feed-in tariff policies are due to end in

several European countries which will lead related actors to participate in electricity markets. In addition, operators that sell energy on electricity markets are required to pay financial penalties proportional to any mismatches between the energy sold based on forecasts and the energy actually produced. Thus, to maintain grid stability and avoid economic losses for power producing companies, efficient power forecasting models are needed.

Unlike wind production forecasting which dates back to the 1980s [2], research on PV production forecasting is a more recent field of research, mainly due to the later development of solar power penetration [3]. Nevertheless, the state of the art regarding PV production forecasting has developed rapidly in recent years [4]. The literature provides us with a large number of models tailored for specific lead time ranges. For instance, day-ahead and longer horizon forecasting models are based on down-scaling approaches from Numerical Weather Predictions (NWP), while short-term models typically use past production [5] and remote sensing observations (e.g. sky or satellite images). Recent studies have shown that accurate forecasting benefits from combining endogenous inputs (variables influenced by other factors within the model, in this case, past production) with exogenous ones (variables influenced by external factors to the model) such as satellite images [6] or measurements from geographically distributed PV units [7], [8]. These models use the spatial dependency of weather variables to provide valuable information regarding cloud movements. For a few years, to enhance PV production forecast performances (in terms of accuracy and computational cost), some authors turned to analog-based methods. This set of methods, initially used for down-scaling NWP outputs, supposes that two similar synoptic states lead to similar local situations [9]. This idea can be extended for PV production forecasting assuming that for two similar atmospheric states, related PV production is comparable [10].

The key contribution of this paper is to propose a short-term forecasting model using spatio-temporal inputs conditioned to the synoptic atmospheric state thanks to an analog-based selection. This approach aims at making efficient use of

past production information by providing a coherent learning subset conditioned to meteorological conditions. Our approach extends current methods that mainly use local information from NWP and explores the contribution that this type of information may bring. In contrast to the tendency for black- or grey-box models today, the proposed approach, although data-driven, makes it possible to link the PV prediction output to specific weather situations, thus enhancing the interpretability of the forecast results in terms of both forecasts and prediction errors.

The paper is organized as follows. Section II describes the objectives of the present study and the approaches implemented. Section III presents the inputs used. Then section IV describes the model's implementation while section V discusses the outcomes. Finally, section VI draws the conclusions of this study.

II. MOTIVATION AND APPROACH

A. Objectives

We propose a meteorological data-driven approach based on an analog method to define coherent learning subsets from a weather point of view. To do so, we use a synoptic scale analog predictor, namely the geopotential field which is considered as a wind driver, to condition the learning phase of the forecast model. In other words, an analog based method makes it possible to determine a subset of past PV production, observed under similar weather conditions, to fit the forecasting model. Moreover, the forecast model is tailored for short-term lead times by considering endogenous data and heterogeneous spatio-temporal inputs. The innovative aspect of this paper is that we propose a spatio-temporal weather-dependent model to forecast PV production for short-term horizons. To our knowledge no other contributions address this topic on the same forecast lead time range (i.e. 15' up to 6h ahead).

B. Conditioned Learning Set Based on an Analogy Approach

PV production is intermittent in nature because of variable cloud cover. Accurate short-term predictions require forecast models that are able to operate efficiently on a wide range of weather conditions (from sunny to overcast skies). To take into account constantly variable weather conditions, models must be dynamically adaptable to the current weather situation. [11] proposed an online adaptive linear prediction model using recursive least squares with a forgetting factor. This approach proved to be more accurate than a static approach because it took into account weather variability and the changing conditions of the PV systems.

Rather than considering only recent observations, some authors have chosen to consider PV production observations related to similar weather states, i.e. analog techniques. The main idea behind analogy can be summed up by the saying: *History repeats itself*. Indeed, similar weather states can be observed throughout time. Of course, depending on the number of physical variables at stake or the extent of the spatial

windows over which the analogy is performed, perfect analogy can be illusory. In that case, one must be satisfied with partial similarity. In the literature, this principle has been used through various adaptations. For instance, [12] proposed an analog down-scaling technique to forecast fine-scale variability from a coarse-scale NWP. As for [10], the authors used an analog based approach to provide probabilistic forecast assuming that similar NWP present similar forecast errors.

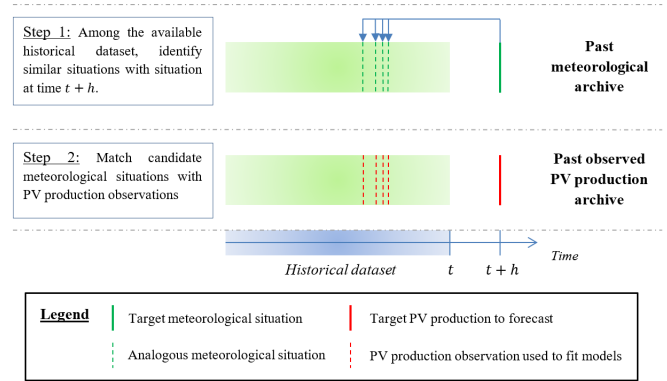


Fig. 1: Schematic representation of the analog process, inspired from [13] and [14].

In the present study, we propose to implement an analog approach to determine a weather-consistent learning subset with the aim of providing a model adapted to future meteorological states. Figure 1 shows in a schematic way how the analogy method is used throughout this study:

- First, we build two past data sets for the same period: a past meteorological archive (containing the analog predictors) and a previously observed PV production archive.
- We then retrieve the forecast of analog predictors at time $t + h$.
- A score of analogy is used to measure similarity between the target meteorological situation, at time $t + h$, and the candidate situations from the past meteorological archive.
- Next, we rank each meteorological situation from the most similar to the least similar. The N best meteorological situations are selected. This approach can be assimilated to a k-Nearest-Neighbor approach.
- The N PV production observations associated with the previous subset are used to train the forecasting model.

1) *Analog variable*: The proposed conditioning approach aims at providing learning samples with coherent PV production patterns. To do so, it is necessary to find situations with similar cloud dynamics. In the meteorology field, geopotential height, which is considered as a wind driver, is a common predictor used to forecast precipitation generation. The geopotential height can be used to derive the pressure gradient that drives the air flow from high to low pressure regions. Thus, geopotential heights are highly correlated to wind flow and cloud development. In this perspective, [15] used geopotential height to forecast global horizontal irradiance (GHI) in

California. It has been previously shown that the formation of local fog and stratus systems is dependent on the synoptic atmospheric flow conditions.

Based on the above, we chose the geopotential field as an analog predictor. This variable can be defined as roughly the height above sea level of a pressure level [16] (e.g. considering the 500 hPa pressure level, if at a specific location the geopotential height is $5,300 \text{ m}^2\text{s}^{-2}$ this means that a 500 hPa atmospheric pressure is achieved at 5,300 m above sea level).

2) *Analog metric*: To measure analogy between target and candidate situations, it is important to define a metric. In the literature on PV forecasting, metrics based on Euclidean distances have been proposed to measure similarity between local variables. In the present study, we consider geopotential height as a spatial variable; thus, it is necessary to use a metric that takes into account the predictor spatial distribution. In the meteorology domain, the score S1 is usually employed [9]:

$$S1 = 100 \frac{\sum_{i=1}^{I-1} \sum_{j=1}^J \left| \Delta_{i,j}^{i,T} - \Delta_{i,j}^{i,C} \right| + \sum_{i=1}^I \sum_{j=1}^{J-1} \left| \Delta_{i,j}^{j,T} - \Delta_{i,j}^{j,C} \right|}{\sum_{i=1}^{I-1} \sum_{j=1}^J \max \left(\left| \Delta_{i,j}^{i,T} \right|; \left| \Delta_{i,j}^{i,C} \right| \right) + \sum_{i=1}^I \sum_{j=1}^{J-1} \max \left(\left| \Delta_{i,j}^{j,T} \right|; \left| \Delta_{i,j}^{j,C} \right| \right)} \quad (1)$$

$$\text{Where: } \begin{cases} \Delta_{i,j}^{i,X} = V_{i+1,j}^X - V_{i,j}^X & X \in \{C, T\} \\ \Delta_{i,j}^{j,X} = V_{i,j+1}^X - V_{i,j}^X \end{cases}$$

- T the target situation (i.e. future state)
- C the candidate situation (i.e. from past records)
- $\Delta_{i,j}^{i,X}$ the east-west geopotential gradient
- $\Delta_{i,j}^{j,X}$ the north-south geopotential gradient
- $V_{i,j}$ the geopotential field at grid node (i,j)

C. Multi-Input Approach

The literature has shown the relevance of exogenous inputs for PV production forecasting. Indeed, recent years have seen an increasing use of satellite images [6], [17] and distributed PV sites [7], [8] to reduce forecasting errors for short lead times. By considering nearby weather conditions, spatio-temporal models tend to outperform temporal-only models.

Due to their wide spatial range, satellite images are used to improve our knowledge of spatio-temporal phenomena occurring close to the site. In contrast to surrounding sites, which are generally sparsely distributed, satellite images provide a much denser source of information with evenly distributed pixels. As a result, it is possible to select a subset of pixels with relevant locations to get the full benefits of the spatio-temporal information (e.g. pixels in the cloud displacement direction). On the other hand, the spatial resolution of satellite images can be too coarse to reveal small cloud structures. Thus, combining distributed PV sites and satellite images seems relevant. It is worth mentioning that satellite images are delivered with a time delay of a few minutes enabling nowcasting approaches.

III. OBSERVATIONAL AND METEOROLOGICAL DATA SETS

The data set considered here is composed of PV production observations from 9 PV plants, satellite images centered on the PV plants, and a geopotential fields reanalysis used as an analog predictor. The PV production observations and satellite images have a temporal resolution of 15 minutes, while reanalysis data have an hourly time step. All of these data range from 2015-01-01 to 2016-12-31.

A. PV Power Production

Production records provided by the Compagnie Nationale du Rhône (CNR), France's leading 100% RES producer, are considered throughout this study. The data set is composed of 9 PV grid-connected systems located in the Rhône valley, mainly along the Rhône River (Figure 2). The installed power capacity ranges from 1.2 to 12 MWp, while the distance between sites ranges from 7.3 to 133 km. Considering this specific data set, spatio-temporal dependencies between sites can be taken to outperform a temporal forecasting approach [18].

Despite the fact that the sites are located in the same region, meteorological conditions can vary from one place to another: the southernmost site is mainly influenced by a Mediterranean climate while the westernmost one is subject to an Alpine climate.

Each site's production observation is normalized by the corresponding installed capacity to enable comparison between PV plants.

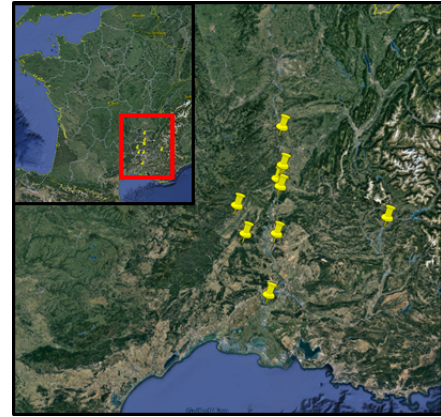


Fig. 2: PV sites located in southeast France.

B. Satellite Images

We considered satellite images obtained from the HelioClim-3 database with the HelioSat-2 method [19]. HelioClim-3 images provide an estimation of the GHI at ground level. The spatial and time resolutions are respectively $0.0625^\circ \times 0.0625^\circ$ and 15 minutes. The images are converted into time series of estimated GHI for each pixel.

C. Analog Predictor

The geopotential height variable comes from the reanalysis data set ERA5-ECMWF [20]. Thus, the expected target situation at time $t + h$ corresponds to observed data whereas in an operational context, NWP models would be used. As a result, the approach implemented corresponds to a *perfect prognosis* case. This choice was made to evaluate the benefits of conditioning the learning to synoptic states, while excluding forecasting errors. However, as pointed out by [15], NWP models are relatively good at predicting synoptic scale variables such as geopotential height fields. Thus, we can assume that conclusions drawn in section VI can be extended, up to a point, to an approach based on predicted rather than reanalysis data.

We chose geopotential fields at pressure levels of 500 and 925 hPa. In the meteorology domain, these fields are known to contain essential information about dynamic and thermodynamic physical processes behind precipitation generation and distribution [21].

IV. FORECAST METHODOLOGY

This section presents the model used to forecast PV production and how the conditioning approach is implemented. All of the developments are made in R language [22].

A. Data Stationarity

PV production is a non-stationary process due to astronomical and meteorological phenomena. A common way to remove the deterministic astronomical component of the signal is to normalize PV production with a clear sky model output, which is a model estimating irradiance with clear sky conditions. Here, the MacClear model [23] is used.

Moreover, in-situ measurements are more subject to error for low solar elevation [24], leading to aberrant normalized values. To deal with this, we imposed a 5° threshold on the solar elevation angle, and only considered observations above this threshold as production observations.

B. Reference Model

The persistence model is often used as a reference in the literature for RES forecasting. This method is simple inasmuch as it does not consider off-site data and does not require modeling processes. The smart persistence model used is defined as follows:

$$\widehat{P}_{t+h}^x = \begin{cases} \bar{P}_t^x & \text{if } \bar{P}_t^x \neq 0 \text{ (i.e. daytime)} \\ \bar{P}_{t+h-24H}^x & \text{if } \bar{P}_t^x = 0 \text{ (i.e. nighttime)} \end{cases} \quad (2)$$

h Forecasting horizon

x Site of interest

\bar{P}_t^x Observed stationarized production at time t

\widehat{P}_{t+h}^x Forecast stationarized production at time $t + h$

To avoid obtaining overoptimistic results, we also implemented a more advanced forecasting model, Random Forest (RF), considering only past production observations of the site of interest.

C. Forecasting Models

The core of the forecasting approach presented here is an auto-regressive (AR) model. ARIMA models are a family that is well suited to short-term predictions and provides easy-to-understand regression coefficients. The initial AR model, fitted on all available data (i.e. independently of the time of day or the weather state), is then extended by conditioning the learning set, which leads to the conditioned auto-regressive (CAR) model. For each forecast, a two-phase conditioning approach is used:

- First, a selection is performed based on a temporal criterion (CAR-T model): the learning set is conditioned to the time of the day. This approach is somewhat similar to [10] who used solar angles (i.e. elevation and azimuth). A comparison between time and solar angle conditioning showed that the first approach leads to slightly better performances.
- Then, a second selection step based on the analog approach presented in II-B is performed (CAR-T.An model): the future geopotential height field at time $t + h$ is used to determine a coherent learning subset among the past records.

Eventually, spatio-temporal (ST) data (neighboring sites and satellite pixels) are added to the model as exogenous inputs, which leads to the CARST-T.An model expressed as:

$$\widehat{P}_{t+h|t,A}^x = \widehat{\beta}_h^0 + \sum_{l=0}^L \sum_{y \in X} \widehat{\beta}_h^{l,y} f_{A,t}(\bar{P}_{t-l}^y) + \sum_{i=1}^N \widehat{\beta}_{h,i}^{Sat} f_{A,t}(\bar{S}_t^i) \quad (3)$$

\bar{S}_t^i Stationarized GHI satellite observation at time t and pixel i

$\widehat{\beta}$ Regression coefficients

L Order of the AR model (here, $L = 2H$)

X Set of neighboring plants

A Analog predictor state at time $t + h$

$f_{A,t}$ Analog based conditioning function

D. Dimensionality Reduction

Taking into account satellite images substantially increases the number of variables. To decrease the related computational burden, only the most informative pixels are retained. To select the relevant pixels, the mutual information (MI) criterion is used [17]. Here, only the 10 most informative pixels are kept.

E. Evaluation Metrics

The Root Mean Square Error (RMSE) is used to assess the forecasting performances:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\widehat{P}_t - P_t)^2} \quad (4)$$

\widehat{P}_t , the daytime production forecast is obtained by multiplying forecast stationarized production by the McClear model and P_t corresponds to non-stationarized observation. To compare model performances with the reference

model, the following comparison skill score is used (with $RMSE_{Perfect\ model}(h) = 0, \forall h$):

$$SS(h) = \frac{RMSE_{Model}(h) - RMSE_{Reference}(h)}{RMSE_{Perfect\ model}(h) - RMSE_{Reference}(h)} \quad (5)$$

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. Sensibility Analysis

In the next step, we carry out a sensibility analysis of the spatial extent of the analog predictor, the number of analogous situations to consider, and the geopotential field (i.e. 500 or 925 hPa). Three spatial windows centered over the Rhône valley region are studied (Fig. 3).

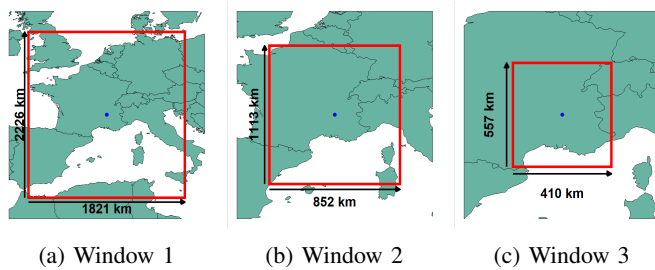


Fig. 3: Grid domains used for analog research

Overall, window 1 (Fig. 3a) leads to better RMSE scores (Fig. 4). For short horizons (i.e. $h \in [15min, 1H00]$), better performances are achieved when considering many analogs (i.e. 100 or 200 analogs), while for longer horizons, it works better to include fewer analogs (50 or 100 analogs). A good compromise seems to be to consider 100 analog. The same analysis is carried out for the 500 hPa geopotential field pressure but lower performances are achieved.

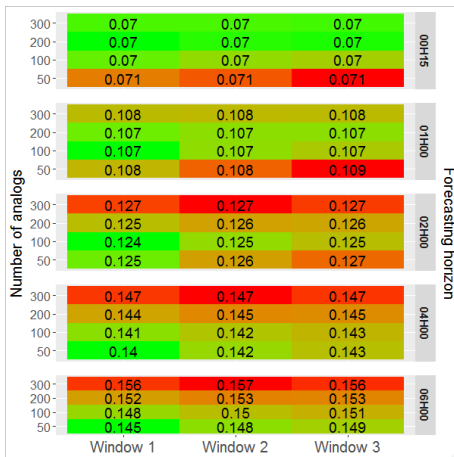


Fig. 4: RMSE scores considering the parameters: spatial window scope, number of analogous situations and forecasting horizons for the 925 hPa geopotential field pressure

B. Influence of Conditioning

With respect to the smart persistence model, the AR model with only endogenous inputs leads to better performances (Fig. 5). Then, conditioning the AR model with the time of the day (i.e. CAR-T model) improves RMSE scores. The AR model conditioned simultaneously by temporal and weather situations (CAR-T.An model) outperforms the previous models especially for longer lead-times. Thus, the proposed conditioning approach allow us to improve forecast performance by up to 28% in comparison with a smart persistence model at a 6H lead-time.

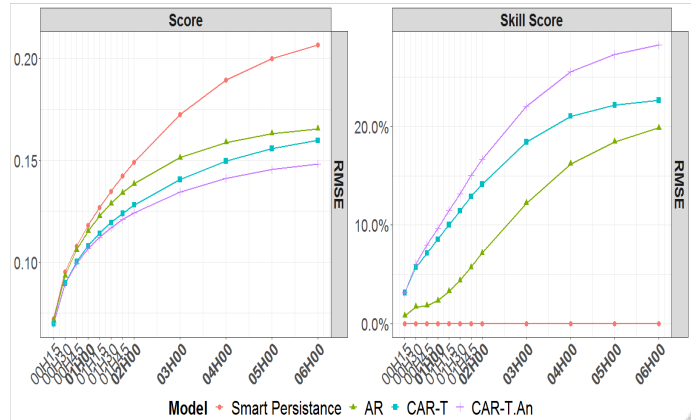


Fig. 5: Averaged performances of conditioning approaches

C. Comparison with an Advanced Forecast Model

Considering only past endogenous observations, the RF model outperforms the AR model (Fig. 6). However, the proposed conditioning approach (CAR-T.An model) improves the performances of the AR model compared to the RF model. Applying the conditioning approach to the RF model (CRF-T.An) improves performances for horizon longer than 1H30 (worse performances for shorter lead times are presumed to result from over-fitting). It is worth mentioning that despite this improved performances, the CRF-T.An model is still less efficient than the CAR-T.An model.

D. Influence of Spatio-temporal Inputs

Finally, we can deduce that including data from the distributed sites and from the satellite images improves the forecasting performances of the CAR model (Fig. 7). The influence of spatio-temporal data is greater for very short-term horizons (i.e. 30min ahead) and decreases progressively for longer horizons, becoming negligible for 6h ahead.

VI. CONCLUSION

This study proposes a deterministic short-term forecast model based on adaptive weather-conditioned learning and fed with spatio-temporal inputs. The conditioning approach is performed thanks to an analog based method using a geopotential field as an analog predictor, which is considered as a wind driver. To improve short-term accuracy, satellite images and distributed PV sites are used as inputs. The

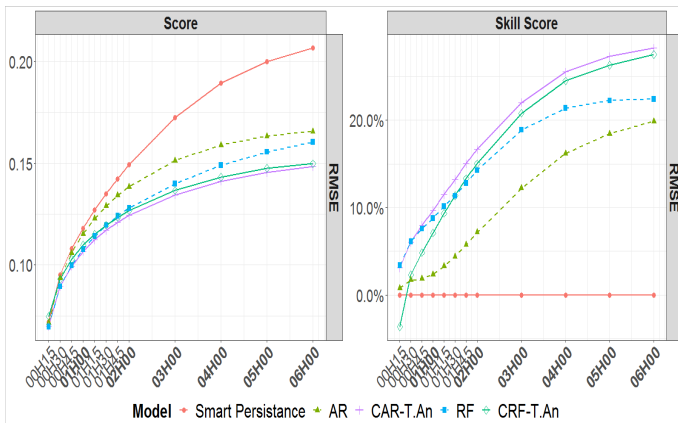


Fig. 6: Averaged performance of the RF model compared with the CAR approach

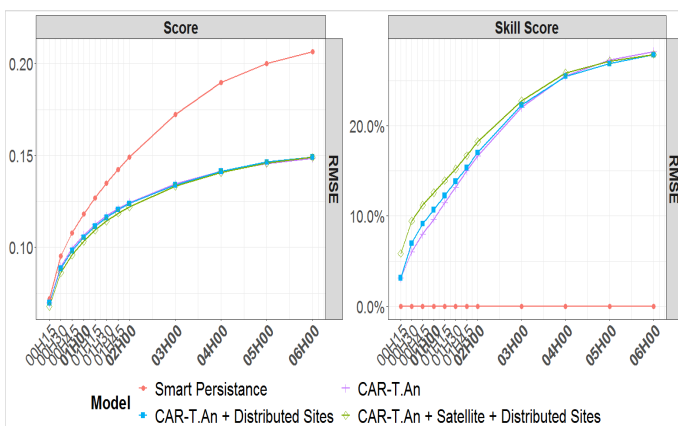


Fig. 7: Averaged performances of the CARST models

results show that the conditioned auto-regressive model (CAR) outperformed the smart persistence reference model by up to 28% for a 6H leading horizon. Thus, we deduce that the conditioning process makes it possible to capture the interaction between PV production and weather conditions. In addition, spatio-temporal data improve the CAR performances by up to 3.5% for short-term horizons. The proposed deterministic approach could be extended to probabilistic forecasting using, for instance, kernel density estimation [17].

VII. ACKNOWLEDGMENT

The authors would like to thank the European Center for Medium-Range Weather Forecasts for providing the reanalysis data as well as the Solar radiation Data (Soda) company for providing the satellite images.

REFERENCES

- [1] RTE, "Electricity Balance - 2019," https://www.rte-france.com/sites/default/files/bilan-electrique-2019_0.pdf.
- [2] G. Kariniotakis, *Renewable Energy Forecasting: From Models to Applications*. Elsevier - Woodhead Publishing, 2017.
- [3] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, Jul. 2016.
- [4] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, Oct. 2016.
- [5] H. T. C. Pedro and C. F. M. Coimbra, "Assessment of forecasting techniques for solar power production with no exogenous inputs," *Solar Energy*, vol. 86, no. 7, pp. 2017–2028, Jul. 2012.
- [6] L. M. Aguiar, B. Pereira, P. Lauret, F. Díaz, and M. David, "Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting," *Renewable Energy*, vol. 97, pp. 599–610, Nov. 2016.
- [7] 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, Apr. 2018.
- [8] R. J. Bessa, A. Trindade, and V. Miranda, "Spatial-Temporal Solar Power Forecasting for Smart Grids," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 1, pp. 232–241, Feb. 2015.
- [9] C. Obled, G. Bontron, and R. Gargon, "Quantitative precipitation forecasts: A statistical adaptation of model outputs through an analogues sorting approach," *Atmospheric Research*, vol. 63, no. 3-4, pp. 303–324, Aug. 2002.
- [10] S. Alessandrini, L. Delle Monache, S. Sperati, and G. Cervone, "An analog ensemble for short-term probabilistic solar power forecast," *Applied Energy*, vol. 157, pp. 95–110, Nov. 2015.
- [11] P. Bacher, H. Madsen, and H. A. Nielsen, "Online short-term solar power forecasting," *Solar Energy*, vol. 83, no. 10, pp. 1772–1783, Oct. 2009.
- [12] P. Mathiesen, D. Rife, and C. Collier, "Forecasting solar irradiance variability using the analog method," in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, Jun. 2016, pp. 1207–1211.
- [13] B. O. Akyurek, A. S. Akyurek, J. Kleissl, and T. S. Rosing, "TESLA - Taylor expanded solar analog forecasting," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Nov. 2014, pp. 127–132.
- [14] L. Delle Monache, F. A. Eckel, D. L. Rife, B. Nagarajan, and K. Searight, "Probabilistic Weather Prediction with an Analog Ensemble," *Monthly Weather Review*, vol. 141, no. 10, pp. 3498–3516, May 2013.
- [15] P. Mathiesen, J. M. Brown, and J. Kleissl, "Geostrophic Wind Dependent Probabilistic Irradiance Forecasts for Coastal California," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 2, pp. 510–518, Apr. 2013.
- [16] N. N. W. Service, "Glossary - NOAA's National Weather Service," <https://w1.weather.gov/glossary/index.php?letter=h>.
- [17] T. Carriere, C. Vernay, S. Pitaval, and G. Kariniotakis, "A Novel Approach for Seamless Probabilistic Photovoltaic Power Forecasting Covering Multiple Time Frames," *IEEE Transactions on Smart Grid*, pp. 1–1, 2019.
- [18] K. Bellinguer, R. Girard, G. Bontron, and G. Kariniotakis, "Short-Term Photovoltaic Generation Forecasting Using Multiple Heterogeneous Sources of Data," in *36th European Photovoltaic Solar Energy Conference and Exhibition*. WIP, Oct. 2019, pp. 1422–1427.
- [19] P. Blanc, B. Gschwind, M. Lefèvre, and L. Wald, "The HelioClim Project: Surface Solar Irradiance Data for Climate Applications," *Remote Sensing*, vol. 3, no. 2, pp. 343–361, Feb. 2011.
- [20] Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate., Copernicus Climate Change Service Climate Data Store (CDS), type: dataset. [Online]. Available: <https://cds.climate.copernicus.eu/cdsapp#!/home>
- [21] R. Romero, G. Sumner, C. Ramis, and A. Genovés, "A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area," *International Journal of Climatology*, vol. 19, no. 7, pp. 765–785, 1999.
- [22] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2020. [Online]. Available: <https://www.R-project.org/>
- [23] M. Lefèvre, A. Oumbe, P. Blanc, B. Espinar, B. Gschwind, Z. Qu, L. Wald, M. S. Homscheidt, C. Hoyer-Klick, A. Arola, A. Benedetti, J. W. Kaiser, and J.-J. Morcrette, "McClear: A new model estimating downwelling solar radiation at ground level in clear-sky conditions," *Atmospheric Measurement Techniques*, vol. 6, pp. 2403–2418, Sep. 2013.
- [24] L. Vallance, B. Charbonnier, N. Paul, S. Dubost, and P. Blanc, "Towards a standardized procedure to assess solar forecast accuracy: A new ramp and time alignment metric," *Solar Energy*, vol. 150, pp. 408–422, Jul. 2017.