A generic methodology to efficiently integrate weather information in short-term Photovoltaic generation forecasting models

Kevin Bellinguer, Robin Girard, Guillaume Bontron, Georges Kariniotakis

A Generic Methodology to Efficiently Integrate Weather Information in Short-term Photovoltaic Generation Forecasting Models

Kevin Bellinguer[a], Robin Girard[a], Guillaume Bontron[b], Georges Kariniotakis[a]

[a] *MINES Paris, PSL University, Centre PERSEE - Centre for Processes, Renewable Energies and Energy Systems, rue Claude Daunesse, Sophia-Antipolis, 06904, France*
[b] *Compagnie Nationale du Rhone, 2 Rue Andre Bonin, Lyon, 69004, France*

Abstract

The power output of Photovoltaic (PV) plants being weather-dependent results to inherent uncertainties about future production. This induces technical challenges for grid operators especially in power systems with high PV penetration, and also financial losses when PV generation is traded on electricity markets. Accurate forecasts for the next hours or days contribute to alleviating these impacts. The literature provides a plethora of forecasting models, among which outstanding approaches combine heterogeneous sources of inputs like measurements, weather forecasts or satellite images. The integration of such inputs into the forecast models may take two forms: either as explanatory features or as state features that condition the model training through a local regression approach. The latter permits to include physics-based information within statistical regression tools and to derive optimised models w.r.t. weather input. Then, these models are extended to integrate spatio-temporal information from satellite observations. We investigate these approaches with the objective to derive the mathematical foundations of a generic methodology to integrate weather information into PV forecasting models. The paper assesses the influence of weather information integration strategies over forecasting performances for two short-term forecasting state-of-the-art models, belonging respectively to linear and non-linear families. Finally general guidelines for forecasters are derived on the procedure to follow when dealing with several sources of information. Evaluations are performed on real-world datasets composed of nine PV plants.

Preprint submitted to Solar Energy *2022-09-28*

Keywords: Short-term solar power forecasting, Grid-connected photovoltaic plants, Analogy, Conditioned forecast, Numerical weather predictions, Spatio-temporal information

Acronyms

AM Analog Methods
ANN Artificial Neural Networks
AR Auto-Regressive
ARIMA Auto Regressive Integrated Moving Average
ARX Auto-Regressive with eXternal inputs
CNN Convolutional Neural Networks
DNN Deep Neural Network
ECMWF Europeran Center for Medium-Range Weather Forecasts
GHI Global Horizontal Irradiance
IFS Integrated Forecast System
kNN k-Nearest Neighbours
LASSO Least Absolute Shrinkage and Selection Operator
mRMR minimal Redundancy Maximal Relevance
nMAE normalised Mean Absolute Error
nRMSE normalised Root Mean Square Error
NWPs Numerical Weather Predictions
PV Photovoltaic
PVPF Photovoltaic Production Forecasting
QRF Quantile Random Forest
RES Renewable Energy Sources
RF Random Forest
SDSI Sattelite Derived Surface Irradiance
SSRD Surface Solar Radiation Downwards
ST Spatio-temporal
SVM Support Vector Machine
T2M 2-m Temperature
TCC Total Cloud Cover
WHCO weather-conditioned

Nomenclature

Indices

t	Time when the forecast is generated
t'	Temporal observations from the learning set
h	Number of time steps of the forecast (i.e. forecast horizon)
i	Index referring to analog predictors

Operator

$\widehat{.}$	Expected quantity
$\bar{.}$	Stationarized quantity
$.^{\intercal}$	Transpose operator
Parameters	
N	Number of analog situations
ϵ_h^N	Bandwidth of the state variable neighbourhood
N^A	Number of analog predictors
$\widetilde{t}$	Half-width of the time window over which the distance is computed
ω_i^A	Weight of analog predictors
L	Order of the AR model
N^{SDSI}	Number of features selected from satellite maps
P_C	Installed capacity
$\beta(Z_{t+h})$	Vector of parameters to be estimated
Variables	
y_{t+h}	Response variable at time $t+h$
X_t	Vector of explanatory features at time t
Z_{t+h}	Vector of state features at time $t+h$
σ_i	Standard deviation
$SSRD_{t+h}$	Solar radiation downwards
$T2M_{t+h}$	Temperature
TCC_{t+h}	Total cloud cover
α_{t+h}^s	Solar azimuth angle
θ_{t+h}^s	Solar elevation angle
P_t	PV production
$SDSI_t$	Satellite-derived surface irradiance
I_t	Observed irradiance
I_t^{CS}	Irradiance under clear-sky conditions
A	Measure of accuracy (e.g. nRMSE or nMAE)
$A_M(h)$	Accuracy score obtained with the model, M, for the horizon h
$A_{Ref}(h)$	Accuracy score of the Persistence model
$A_P(h)$	Accuracy score of a perfect forecast (for the nRMSE and nMAE metrics, a perfect forecast implies $A_P(h) = 0$)
Functions	
D	Distance function
f_{root}	Root regression model employed for the mapping of X_t to y_{t+h}
f_j	j$^{\text{th}}$ regression tree

1. Introduction

1.1. Context

Due to environmental concerns and energy resources depletion, societies are taking action to reduce greenhouse gas emissions. Among these low-carbon strategies, Renewable Energy Sources (RES) appear to be a promising alternative to carbon-based energies. RES penetration is growing continuously thanks to reduced technology costs [1] and supportive mechanisms.

Since Photovoltaic (PV) generation is weather-dependent, it is characterised by high variability and limited predictability. When PV constitutes a significant share in power systems, these features raise challenges for power system operators, which have to ensure a high level of power quality and strike a balance between production and demand. In addition, in a market environment, this weather-induced uncertainty can jeopardise the profitability of PV units.

To deal with the issue of intermittency, several solutions are investigated throughout the literature [2] (e.g. storage systems [3], demand response [4]). RES forecasting appears as a cost-effective option that can anticipate power imbalances and lead to optimal use of flexibility solutions or traditional adjustment means. In the present paper, we investigate Photovoltaic Production Forecasting (PVPF) for short-term horizons: from 15 minutes up to 6 hours ahead.

1.2. PV Power Generation Forecasting

Over the last decade, PVPF has been a very active field of research; in this regard, [5, 6, 7, 8] provide fairly complete literature reviews. Within the literature, two main trends are investigated to improve forecasts accuracy: (1) reducing model uncertainty by employing complex non-linear structures (e.g. Deep Neural Network (DNN)), and (2) combining heterogeneous sources of input information.

For short-term PVPF up to 6 hours ahead, endogenous inputs (i.e. past PV production measurements) are typically the main drivers. For a few years, we observe a paradigm shift from models based on information at the PV plant location to models integrating Spatio-temporal (ST) information such as production measurements from nearby PV plants, observations from nearby weather stations, and satellite-derived data, and turns out to be more accurate than forecasts based solely on temporal information [9, 10, 11, 12, 13]. Satellite-derived data come from space-borne pho-

tographs of Earth that provide observations of the cloud cover, these images are usually post-processed to obtain Global Horizontal Irradiance (GHI) estimations. The increasing use of satellite-derived data in the literature in the last years may be explained by the large spatial coverage of the images that permits to counterbalance lack of spatially distributed measurements due to a low density distribution of PV plants or weather stations. Moreover, this data is suited to operational use thanks to its regular updates (e.g. every 15-minutes). For higher forecasting horizons, such as day-ahead, Numerical Weather Predictions (NWPs) model outputs become the main source of information. They are issued by physics-based models which integrate the complex dynamics of the atmosphere. In the short horizons, NWPs are often neglected to the advantage of ST data. Still several results (e.g. [10]) indicate that they contribute to forecast accuracy improvement by providing the PVPF model with information about future weather trends.

1.3. Integration of Weather Information

PV generation depends on a number of meteorological variables such as irradiance, cloud cover, and ambient temperature. The combinations and interactions of these variables lead to a large range of weather states associated with significant varied dynamics. For this reason, NWPs provide valuable information to PVPF models on the expected atmosphere state and how it will influence production. The predicted weather information can be integrated in the PVPF modelling chain in two different ways: either explicitly or implicitly.

1.3.1. Explicit Integration

The most straightforward method considers NWPs as additional explanatory features within the PVPF model (i.e. data is added linearly to the model). Only one model is fitted for a large range of weather situations thanks to the atmosphere dynamics being explicitly carried by NWPs. This is a computationally inexpensive and easy way to include new information. In comparison with models fitted only with production observations, the use of NWPs as additional regressors improves short-term forecasting performances [14, 10]. Former works by the authors in wind power forecasting showed that integrating NWPs as input is beneficial as they inform on the tendency of weather conditions and this results to an almost double improvement w.r.t. persistence compared to models non considering NWPs.

1.3.2. Implicit Integration

The alternative paradigm is to consider the weather information as state variables. Then, they act as a kind of classification tool that associates PV production data observed under similar atmospheric states. This approach provides a set of expert models dedicated to specific atmospheric states and is adaptive in the sense that the training of the model is conditioned to the weather situation. It can be implemented in two ways: 1) either through a regime-switching model approach, where each model is dedicated to a specific weather type (e.g. sunny, cloudy) [15, 16, 17, 18] or through binning of weather variables [19], and 2) by taking a dynamic approach, where the model's parameters are updated regularly [14, 20, 21].

The second option, which is investigated throughout this paper, consists in training a model for each new situation based on the N most similar past situations. In a space composed by the history of weather features, this approach searches the past situations that are closest to the situation defined by the predicted features. In that sense, this approach presents similarities to a k-Nearest Neighbours (kNN) algorithm. The main drawback of this approach is the need to re-train the model each time a new forecast is generated.

The literature proposes various terminologies to name these approaches (e.g. regime-based models [19], weather status pattern recognition models [16]). In an effort to unify these different approaches, we introduce the terminology *weather-conditioned (WHCO)* to refer to an approach, that operates a weather-based selection or classification in its learning dataset.

The WHCO strategy offers the possibility to condition several types of forecasting models, such as Auto-Regressive with eXternal inputs (ARX) models [19], Artificial Neural Networks (ANN) [22, 17], Convolutional Neural Networks (CNN) [18], and Support Vector Machine (SVM)[15, 16]. From a performance perspective, WHCO models exhibit greater forecasting skills than their counterparts trained on all past production observations.

1.4. Integration of Spatio-temporal Information

In a similar way, ST information can be integrated within a PVPF model as either explanatory or state features.

The first option is the most common approach and considers several sources of ST information: [9, 11] use production measurements of spatially distributed PV units, [10, 21] consider a selection of pixels derived from satellite imagery, while [23] fits solar forecasting models with observations from

nearby irradiance sensors. In [19], the authors combine the idea of WHCO with the use of ST data as explanatory features for 10-second ahead solar forecasts. This study highlights that ST forecasting models benefit from WHCO approaches based on features related to cloud motion. Indeed, wind-conditioned forecasting models are able to select geographically distributed sensors in line with cloud displacement direction, while un-conditioned models select sensors in the direction of the most dominant winds.

To the authors' knowledge, only [24] consider ST information as state variables with the aim of improving the degree of similarity between analog situations.

1.5. Key Contributions

The present paper improves the current state of the art with several key contributions:

- As shown above, the WHCO concept is already present in the literature and is applied using a wide range of forecasting models. As the WHCO strategy can be viewed as a means to include non-linear capabilities, we analyse what is its influence over linear and non-linear models. As an illustration, we focus on two state-of-the-art short-term forecasting models: the ARX and the Random Forest (RF) models.

- The literature highlights that the WHCO strategy outperforms forecasting models trained on all past production observations. Similar conclusions are drawn when NWPs are considered as explanatory features. However, to the best of the authors' knowledge, no comparison has been performed to determine which approach provides the best forecasting performances.

- To answer these two questions, we lay the mathematical foundations of a standardised methodology to weather-condition any regression model. Such approach can be viewed as a means to include physics-based information within statistical models, and thereby to enhance model interpretability. It provides a framework for homogeneous comparisons between models' performances.

- We provide general guidelines for forecasters regarding the best weather information integration strategy to use according to the types of input and forecasting model they have at their disposal.

The paper is organised as follows. First, Section 2 describes the methodology of this study and the approaches implemented, while Section 3 introduces the data sets. Validation results are discussed in Section 4. Finally, Section 5 draws the conclusions and presents research opportunities that could develop the present work.

2. Proposed Methodology

To fulfil the above-mentioned objectives, we need to characterise the interactions that exist between (i) the different ways of integrating weather information, (ii) the nature of explanatory features (i.e. local or ST), and (iii) the model family considered (i.e. linear or non-linear regression models).

To do so, a modular architecture allowing the inhibition or the activation of some specific mechanisms occurring in the forecasting chain is proposed. It is composed of four main building elements (Figure 1): (1) the WHCO block, (2) the forecasting block, (3) the state variables block, and (4) the explanatory features block. This architecture may be seen as a generic data-driven forecasting model enhanced by a physical-based conditioning approach, which enables the model to perform local regression with respect to the atmospheric state.

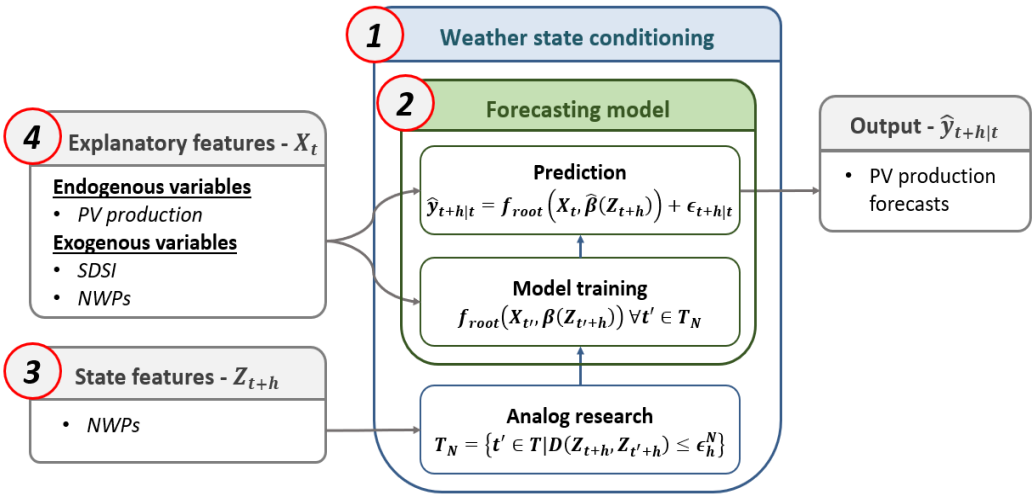

Figure 1: Modular structure used to investigate the interaction between input integration strategies, input types, and model families.

2.1. Weather State Conditioning

2.1.1. The Meteorologist's Perspective

In the meteorology field, the analogy principle stipulates that similar weather states can be observed throughout time. Perfect similarity is hardly attainable due to the atmospheric variability, but similar situations can be found when considering deep datasets, or few weather parameters. This has led to the development of forecasting approaches based on Analog Methods (AM). For instance, such methods can be used as a downscaling approach by supposing that similar large-scale phenomena induce similar local-scale phenomena. In the precipitation forecasting field, AM are used to derive probabilistic relations between large-scale variables (e.g. geopotential fields), named predictors, and local-scale features (e.g. precipitation) denoted as predictands [25].

In the present study, we assume that similar forecasts of the atmospheric states at the PV site position (i.e. predictor) lead to similar PV production measurements (i.e. predictand). Thanks to the analogy principle, we can select a subset of past weather states forecasts which are analog to the expected future atmospheric situation. Then, instead of deriving an estimation of the future production from production observations associated with these analogs (like what is performed in [26, 24]), a dedicated forecasting model deals with the establishment of the statistical law between this production measurements subset and the associated explanatory features. This approach makes it possible to obtain weather-based expert models that dynamically update their coefficients according to the weather situation. Figure 2 illustrates how the AM is used throughout this study. The modelling steps shown in the figure are as follows:

0. First, we build three datasets:
   (a) *The candidate archive* which contains weather forecasts,
   (b) *The response archive* which gathers PV production observations,
   (c) *The explanatory archive* which represents the explanatory features dataset.

1. A score of analogy, D (defined by Equation 2), measures the similarity between the target meteorological situation at time $t + h$ with past forecasts at lead time +h from the candidate situations archive, and

ranks them. The N most similar meteorological situations form the analog situations subset.

2. The N associated PV production observations at lead-time +h are selected as well as the corresponding observations from the explanatory archive.

3. The selected elements from the response and explanatory archives are used to train a forecasting model, while last observations of the explanatory features at time t permit the generation of PV production forecast at time $t + h$.

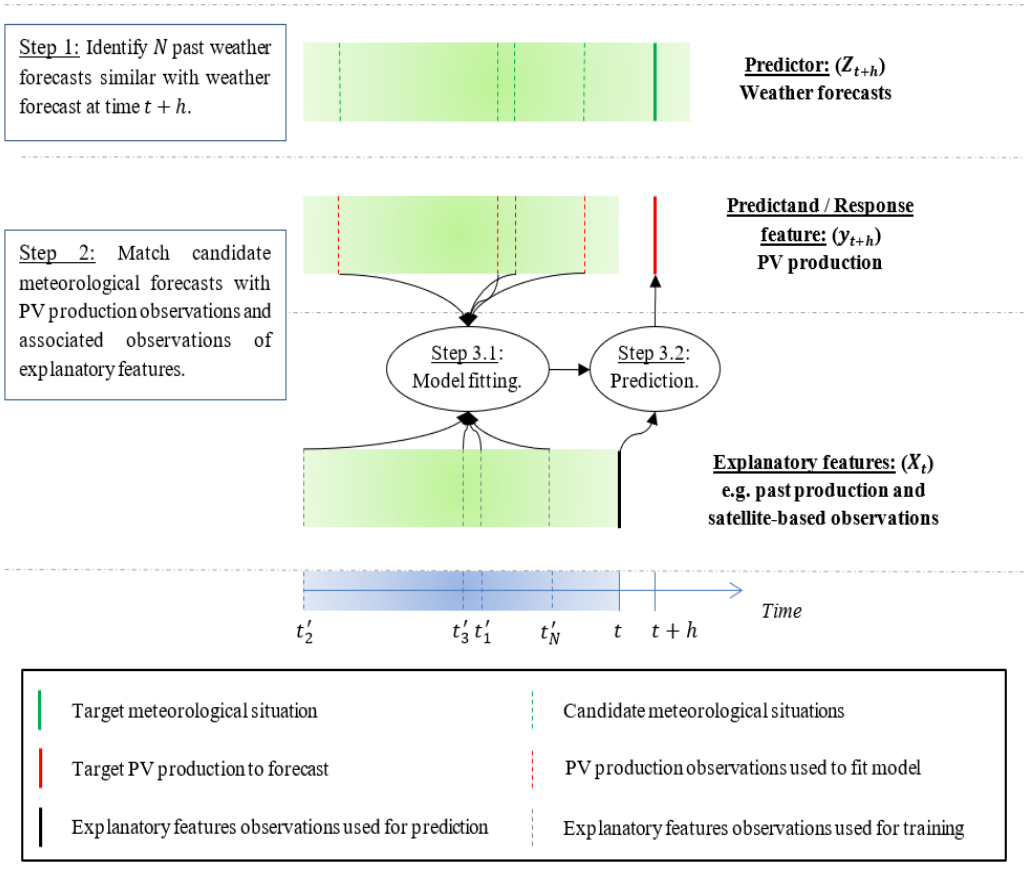

Figure 2: Schematic representation of the training of the analog-based approach, inspired from [27, 28].

2.1.2. Local Regression

From a mathematical point of view, WHCO may be assimilated to a local regression approach [29]. Instead of fitting a regression model (Equation (1)), denoted as *root model*, globally on the whole dataset of available observations $\mathcal{T}$, the fitting is performed locally on a subset $\mathcal{T}_N$. This subset gathers N observations associated with the neighbourhood of the focal point, Z_{t+h}, namely the forecast of weather parameters. Attention is drawn to the fact that the fitting neighbourhood is defined within the state space (i.e. space containing state features, Z_{t+h}), while the model fitting is performed with explanatory features, X_t. This operation is repeated for all the fitting points of the testing set in a rolling manner.

$$\hat{y}_{t+h|t} = f_{root}\left(X_t, \beta(Z_{t+h})\right) + \epsilon_{t+h|t}. \tag{1}$$

Equation (2) presents the distance metric, D, proposed by [26], used here to measure the degree of similarity between the different observations of the state space, and to rank past situations according to their degree of likeness with the focal point. This score outperforms the traditional Euclidean distance thanks to the term under the square root, which takes into account the temporal evolution of the features. In [26], the authors propose a grid search optimisation procedure to determine the optimal set of weights, ω_i^A, leading to the best forecasting performances. In the present configuration, this approach is hardly conceivable due to the high computation cost induced by the RF model fitting. Thus, we presume that the weights are uniform. Eventually, only the N closest elements are kept (Equation (3)).

$$D(Z_{t+h}, Z_{t'+h}) = \sum_{i=1}^{N^A} \frac{\omega_i^A}{\sigma_i} \sqrt{\sum_{j=-\widetilde{t}}^{\widetilde{t}} (z_{i,t+h+j} - z_{i,t'+h+j})^2}. \tag{2}$$

$$\mathcal{T}_N = \{t' \in \mathcal{T} \mid D(Z_{t+h}, Z_{t'+h}) \leq \epsilon_h^N\}. \tag{3}$$

2.2. Forecasting Models

The root model is employed to infer statistical relationships between observations of the response variable and explanatory features belonging to the subset $\mathcal{T}_N$.

Auto Regressive Integrated Moving Average (ARIMA) models [30] constitute a family of models well-suited to short-term PVPF [14, 9, 11]. Here,

the ARX model is considered as the linear root model of our modelling strategy (Equation 4). The high number of available explanatory variables makes the model more complex and may undermine its accuracy. To tackle this issue, the Least Absolute Shrinkage and Selection Operator (LASSO) procedure [31] is implemented to perform feature selection and regularisation (Equation 5).

$$f_{root}\left(X_t, \beta\right) = \beta_0 + \beta' X_t^{\intercal} \tag{4}$$

$$(\hat{\beta}_0, \hat{\beta}') = \underset{\beta_0, \beta'}{\arg\min} \left(\frac{1}{2} \sum_{t=1}^{N} \left(y_t - \beta_0 - \beta' X_t^{\intercal}\right)^2 + \lambda \sum_{j=1}^{p} \left|\beta'_j\right| \right) \tag{5}$$

The second model considered is the RF [32], which is a data-driven model able to perform non-linear mapping between a set of input and output features. It is an ensemble learning method composed of several decision or regression trees grown in parallel, whose the outputs are averaged (Equation 6). Today, RF is one of the mainstream models employed in the field of RES forecasting: as an example, a recent forecasting competition was won by an architecture based on a Quantile Random Forest (QRF) model [33].

$$\hat{Y}_t = \frac{1}{T} \sum_{j=1}^{T} f_j(X_t) \tag{6}$$

2.3. State Variables

To select PV production data observed under similar weather patterns, it is necessary to work with weather parameters that accurately account for the PV generation process. The features considered are the following variables of the NWPs model:Surface Solar Radiation Downwards (SSRD), 2-m Temperature (T2M), and Total Cloud Cover (TCC) at the site position. In addition, the solar azimuth, and elevation angles (α^s and θ^s respectively) are added for two reasons: (1) despite irradiance-based explanatory features being normalised by clear-sky model output (this process is detailed in Section 3.1), [23] highlights that some periodical effects are still present in the normalised outputs; and (2) in a context of WHCO, these inputs enable us to implicitly take into account effects due to dawn (e.g. shading). The vector of state features is built as:

$$Z_{t+h}^{\intercal} = \begin{bmatrix} SSRD_{t+h} \\ T2M_{t+h} \\ TCC_{t+h} \\ \alpha_{t+h}^{s} \\ \theta_{t+h}^{s} \end{bmatrix}_{(5,1)} \tag{7}$$

2.4. Explanatory Variables

The root models in Equation (1) are fed with two kinds of explanatory variables: either endogenous inputs (i.e. PV production) and/or exogenous inputs (i.e. spot NWPs and Sattelite Derived Surface Irradiance (SDSI)). Within the scope of short-term PVPF, endogenous inputs are essential, and consequently they are systematically integrated. In a next step, NWPs and SDSI features are considered individually or jointly to assess their influence on forecasting skills. Equation (8) represents the regressor vector containing all available inputs.

$$X_{t}^{\intercal} = \begin{bmatrix} \overline{P}_{t-h-L:t-h} \\ \overline{SDSI}_{t-h-L:t-h}^{1:N^{SDSI}} \\ \overline{SSRD}_{t} \\ T2M_{t} \\ TCC_{t} \\ \alpha_{t}^{s} \\ \theta_{t}^{s} \end{bmatrix}_{((L+1)*(N^{SDSI}+1)+5,1)} \tag{8}$$

3. Observational and Meteorological Datasets

The models are trained over the year 2015 and evaluated on the period covering 2016. The input explanatory variables and the PV power forecast outputs have a 15-minute granularity, while the state variables consist of hourly predictions. Instead of performing expensive temporal interpolations of the state variables at a 15-minute time step, we assume that the atmospheric state remains constant from time $t - 00h15$ to time $t + 00h30$.

3.1. Data Normalisation

Solar-related features (i.e. PV power, SDSI, SSRD) are non-stationary by nature, which makes them more complex to investigate, while reducing the set of practical tools [5]. In simple terms, stationarity means that the

properties of a time series (e.g. mean, variance) do not change over time. In the case of PV production, non-stationarities result mainly from astronomical phenomena: the sun path induces daily (i.e. day/night cycle) and monthly (seasonal cycle) variability.

Using traditional stationarisation techniques, such as differentiation of the time series, is not enough to remove non-stationarities [11]. In general it is difficult to fully stationarise solar production timeseries. Methods in the literature rather attempt some "normalisation" to attenuate the deterministic components. A common option consists in normalising observed irradiance time series I_t with the output of a clear-sky model I_t^{CS} (Equation (9)), which estimates the part of solar irradiance reaching the ground assuming a cloudless sky. This normalisation permits to exhibit the stochastic component linked to cloud motion. The clear-sky time series are derived from the McClear model [34], while [35] provides the diffuse model to project irradiance on tilted planes. Moreover, observations associated with low-sun situations (i.e. $\theta^S \leq 5°$) are excluded because irradiance levels are too low to be of significance in solar power applications [36].

$$\bar{I}_t = \frac{I_t}{I_t^{CS}}. \tag{9}$$

3.2. PV Production Observations

We consider production records from nine fixed-tilt PV grid-connected systems located in the Rhône valley in France, mainly along the Rhône River (Figure 3) and operated by the Compagnie Nationale du Rhône. The installed power capacity ranges from 1.2 to 12 MWp. Pre-processing is employed to remove obvious outliers from both the training and testing sets.

3.3. Numerical Weather Predictions

The NWPs used in this work are obtained from the highest resolution (HRES) configuration of the Integrated Forecast System (IFS) run by the Europeran Center for Medium-Range Weather Forecasts (ECMWF). This model is run twice a day, at 00:00:00 UTC, and 12:00:00 UTC providing parameters with a 1-hour temporal resolution and a $0.1° \times 0.1°$ spatial resolution. NWPs model outputs are considered at the PV farm's location by performing a bi-linear interpolation of the nearest grid points.

Depending on the lead time, several predictions can be issued for the same time (e.g. predictions for time 13:00:00 are provided by the runs of

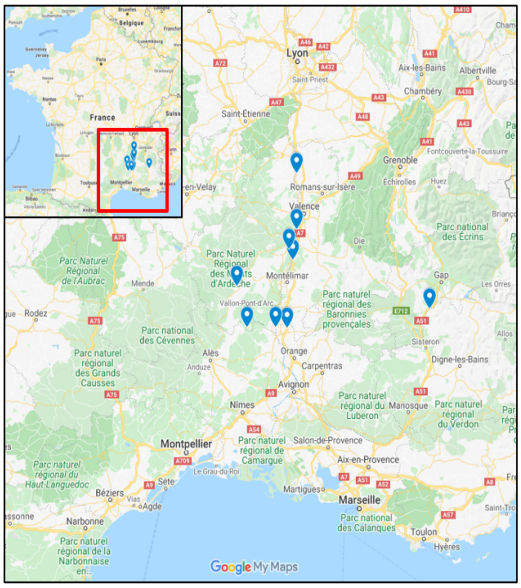

Figure 3: Spatial distribution of PV sites located in southeast France. The blue spots indicate the nine PV plants.

00:00:00 and 12:00:00 on the same day). Thus, two approaches are considered according to the weather information integration strategy. First, one may consider that each run has distinctive features: the number and position of initial observations used to initialise the numerical model vary according to its launching time, which impact the quality of the forecasts. Therefore, when NWPs are considered as state features, it is relevant to compare predictions with similar errors, and thereby, runs delivered at the same time of the day. Alternatively, one may focus on the fact that forecasting precision tends to decrease as the lead time increases. Thus, when NWPs are considered as explanatory variables, predictions from the most recent run are considered.

3.4. Satellite-derived Surface Irradiance

The SDSI data are extracted from the Helioclim-3 database, which stores 15-minute GHI maps with around a 5-km spatial resolution in Europe. This database is generated by the Heliosat-2 method [37], which processes images collected by meteorological geostationary satellites into maps of solar radiation. For the purpose of this study, time series of estimated GHI are derived for each pixel of these spatial maps. A feature selection based on the minimal Redundancy Maximal Relevance (mRMR) framework [38] is implemented

to reduce the computational burden induced by the high dimensionality of SDSI. This selection step, performed on the training set, provides a subspace composed of N^{SDSI} satellite pixels for each forecast horizon.

4. Results

This section investigates the possible interactions between weather information integration strategies with ST inputs and two types of forecasting models. To assess the specific contribution of each data source, we draw on the modular modelling structure introduced in Section 2. All developments are performed using R language [39].

4.1. Benchmark Model and Evaluation

4.1.1. Clearness Index-based Persistence Model

In the literature of RES forecasting, the clearness index-based persistence model, provided by Equation (10), is often used as a reference. This model only uses past measurements and does not involve any modelling process. The main assumption is that the weather situation, and so the related PV generation, remains unchanged for a certain amount of time. Despite being a naive approach, it performs well for very short-term horizons for which the persistence in cloud structures and distribution can be observed and for situations with low weather variability. Thereafter, this model is simply denoted as *persistence*.

$$\widehat{P}^x_{t+h|t} = \begin{cases} \frac{P^x_t}{I^{CS,x}_t} I^{CS,x}_{t+h} & \text{if } P^x_t \neq 0 \text{ (i.e. daytime)} \\ \frac{P^x_{t+h-24h00}}{I^{CS,x}_{t+h-24h00}} I^{CS,x}_{t+h} & \text{if } P^x_t = 0 \text{ (i.e. nighttime)} \end{cases} \tag{10}$$

4.1.2. Forecast Performance Criteria

A large range of performance metrics has been defined by the scientific community, each of which highlights a specific aspect of the forecasting error [40]. In the present study, we use a set of well-established metrics to characterise the quality and accuracy of the models and to enable comparison with other studies.

We consider the normalised Root Mean Square Error (nRMSE) and the normalised Mean Absolute Error (nMAE), defined hereinbelow, where N is the number of paired data. Nighttime data are discarded because they do not offer relevant information. Normalisation is done with the nominal capacity

of the PV plants. Scores are computed individually for the nine PV farms but for a more compact presentation we average them.

$$nRMSE^x(h) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{\widehat{\boldsymbol{P}}^{\boldsymbol{x}}_{\boldsymbol{t+h|t}} - \boldsymbol{P}^{\boldsymbol{x}}_{\boldsymbol{t+h}}}{P^x_c}\right)^2}. \tag{11}$$

$$nMAE^x(h) = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\widehat{\boldsymbol{P}}^{\boldsymbol{x}}_{\boldsymbol{t+h|t}} - \boldsymbol{P}^{\boldsymbol{x}}_{\boldsymbol{t+h}}}{P^x_c}\right|. \tag{12}$$

To quantify the relative improvement of the considered forecasting models over persistence, the following comparison skill score is used:

$$SS_M(h) = \frac{A_M(h) - A_{Ref}(h)}{A_P(h) - A_{Ref}(h)} \times 100\%. \tag{13}$$

A positive (negative) skill score implies that the forecasting model performs better (worse) than the reference model. A skill score equal to zero means that the performances of both models are equal, while a perfect forecast is obtained for a skill score of one.

4.2. Considered Architectures and Terminology

To assist the reader in understanding the configurations assessed, Figure 4 features model denominations and block diagrams representing the model's architecture:

- $Model \in \{AR, RF\}$: the AR and RF models are investigated.
- $X_1, X_2 \in \{\emptyset, NWP, SDSI\}$: forecasting models are fed with PV production observations and/or NWPs and/or SDSI.
- $Z \in \{NWP\}$: forecasts are conditioned with NWPs.

4.3. Weather Information Integration within a Linear Model

4.3.1. Non-linear Dependent Feature

The WHCO approach is a straightforward and efficient way to integrate explanatory features that have a non-linear relationship with the response variable in a linear model. To illustrate this statement, we consider the integration of the azimuth angle, α_s, within the ARX model. Figure 5 shows that performances achieved by considering the azimuth angle as an explanatory feature (i.e. *AR + Azimuth* model) are outperformed by the state feature integration mode (i.e. *CAR(Azimuth)* model).

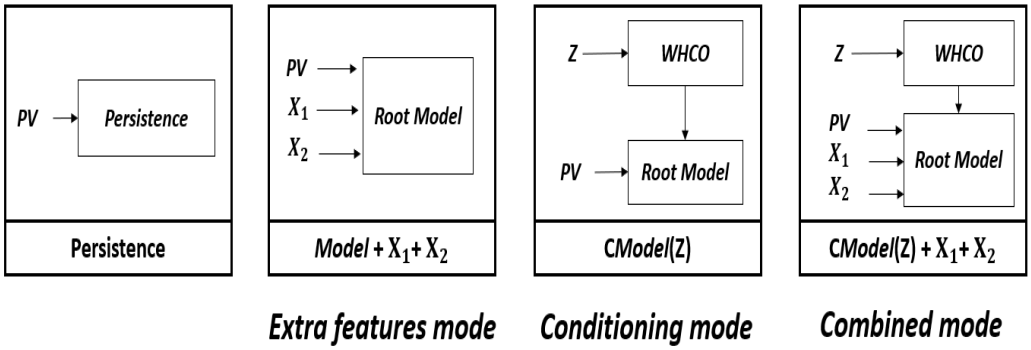

Figure 4: Model designation and corresponding structure.

4.3.2. Two Complementary Approaches

Figure 6 represents forecasting performances achieved by the ARX model fed with SSRD feature. The WHCO approach (i.e. *CAR(SSRD)* model) performs poorly compared to its counterpart (i.e. *AR+SSRD* model). When the two integration modes are employed simultaneously (i.e. *CAR(SSRD)+SSRD* model), resulting performances are the highest for both metrics considered. Therefore, the extra-feature mode should be preferred for features having a linear dependence with the response variable. Yet, far from being two opposed integration modes, the WHCO and the extra-features strategies assess different kinds of information which can complement each another.

4.4. Interaction Between Forecasting Model Families, Sources of Information and Integration Strategy of Weather Data

In this section, we compare the different forecasting architectures to determine the best way of integrating data to obtain optimal forecasting performances. Figure 7 and Figure 8 gather the forecasting skill scores of AR and RF models fed with past PV production observations, and/or NWPs and/or SDSI observations. The WHCO version of these models is also evaluated.

4.4.1. Production Observations

In cases where only PV production observations are available, the nonlinear model turns out to be more efficient than the linear model both in term of nRMSE and nMAE. This statement is observed for all forecasting horizons under study.

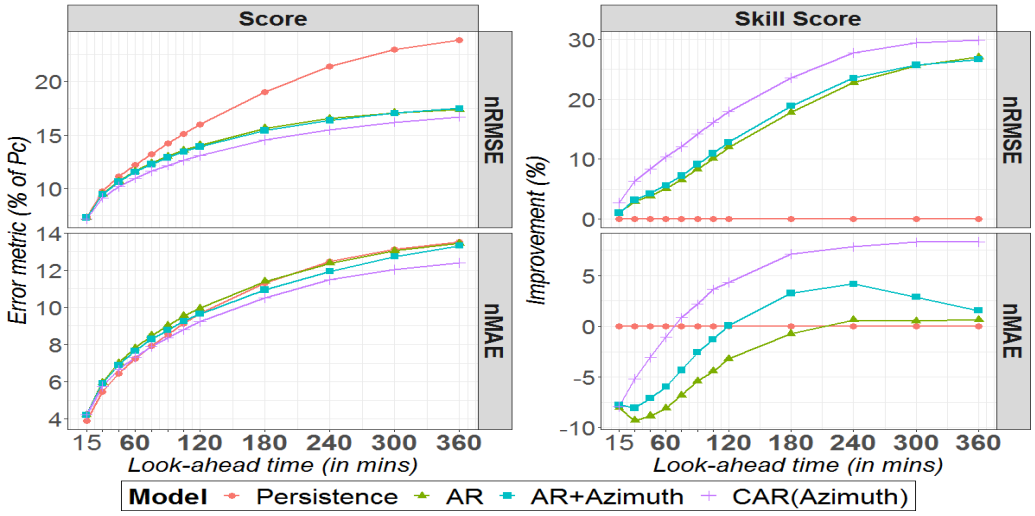

Figure 5: Integration of solar azimuth angle either as an additional explanatory feature or as a state feature in an Auto-Regressive (AR)-based forecasting model.

4.4.2. Production Observations + SDSI

In a ST context, the *RF+SDSI* model outperforms the *AR+SDSI* model for all considered horizons and metrics. A comparison of the skill scores between temporal-based forecasts (i.e. *AR* and *RF*) and ST-based predictions (i.e. *AR+SDSI* and *RF+SDSI*) highlights that the non-linear model is able to extract more information from ST data sources than the linear model (e.g. for a 1-hour lead-time the nRMSE improvement due to ST information is $18.1 - 8.7 = 8.0$ with the RF model, while it is of $10.1 - 5.0 = 3.7$ with the ARX model).

4.4.3. Production Observations + NWPs

Explanatory Features. This approach explicitly considers the information carried by the NWPs data. The *AR+NWP* model manages to extract relevant information in such a way that it can improve nRMSE scores by up to 22.2% in comparison with the *AR* model for a 6-hours lead time. Accuracy improvement due to weather information increases with lead time. Similar conclusions are drawn when comparing the *RF+NWP* model with the *RF* model. Nevertheless, for all considered horizons, the *RF+NWP* model outperforms its linear counterpart.

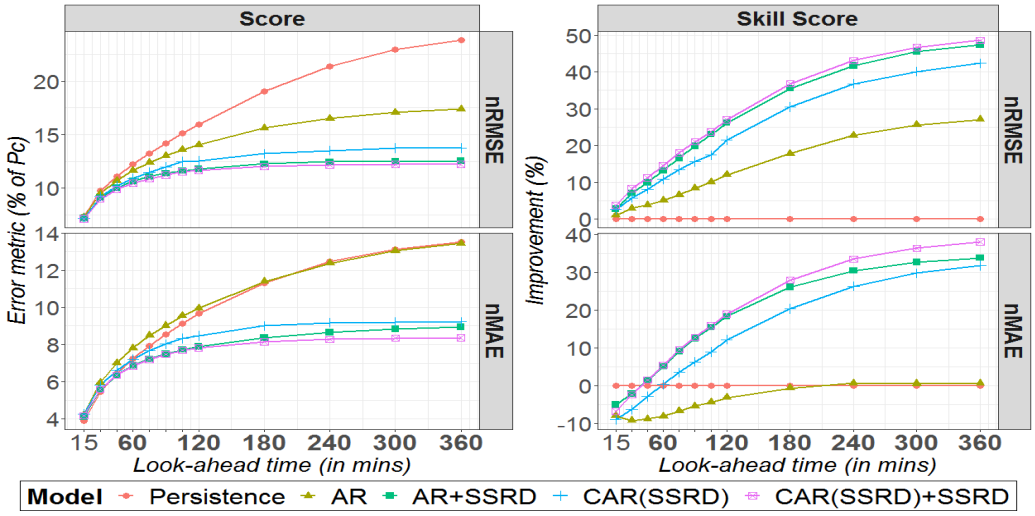

Figure 6: Influence of the feature integration approach of SSRD (i.e. as explanatory feature, state feature or both) on forecasting performances.

State Features. This approach considers the NWPs information as a way to gather PV production measured under similar weather states. As a result, the dynamics are directly carried by production observations. The *CAR(NWP)* model slightly performs better than the *CRF(NWP)* model: on average, a performance increase of +1.64% and +0.68% (in term of nRMSE and nMAE) is observed in favour of the *CAR(NWP)* model.

Figure 9 represents the error distribution of the *Persistence*, *AR*, *CAR*, *RF* and *CRF* models. At first glance, for all models, the error distribution tends to get wider as the look-ahead time gets longer. For short look-ahead times (i.e. $h \leq 30$ minutes), all of the models (except ARX model) have almost perfectly symmetrical, centred distributions around 0. On the contrary, for higher forecasting horizons (i.e. $h \geq 180$ minutes), we observe skewed distributions for the *Persistence*, *AR* and *RF* models. These asymmetrical distributions are corrected with the WHCO process. In addition, the distribution curves of conditioned models tend to be sharper (e.g. the *CAR(NWP)* model never experiences errors greater than 32.5% of the installed capacity while the *AR* model reaches errors representing 40.0% of P_c).

Explanatory and/or State Features. In a next step, we focus on the best way to integrate NWPs in a forecasting model. In the case of the linear model,

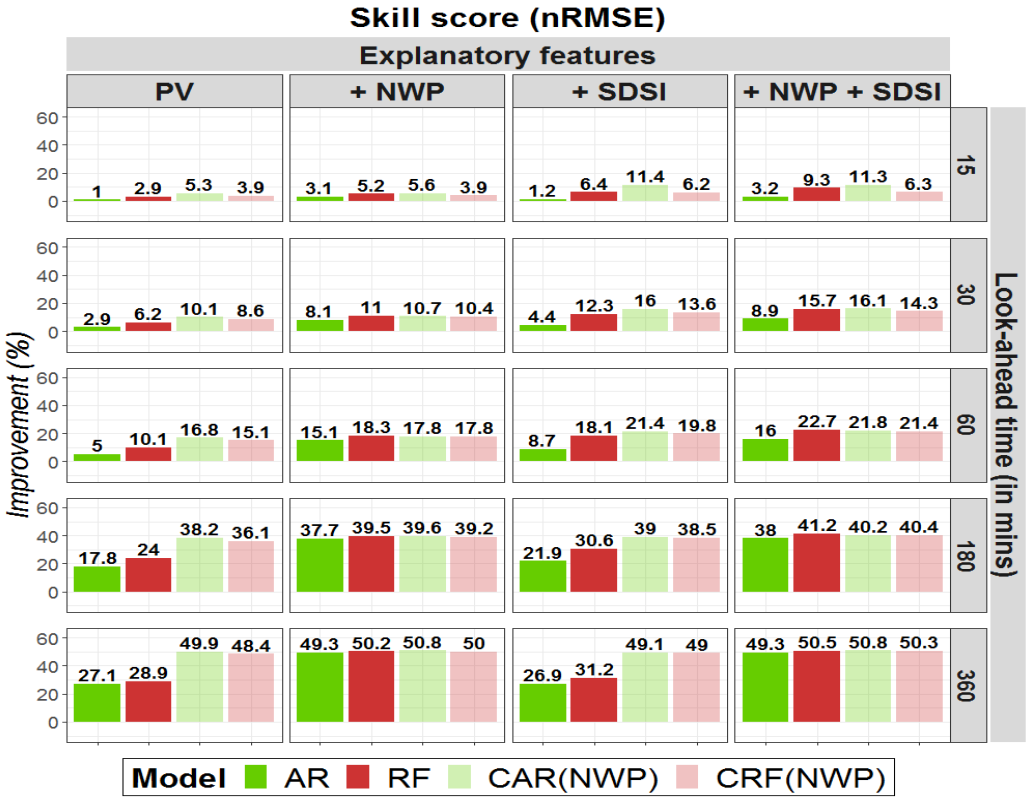

Figure 7: nRMSE skill scores with regard to the persistence model. Dark colours symbolise forecasting models trained on the whole dataset, while light colours stand for WHCO models. Columns represent the explanatory features, while rows indicate the lead time of the forecasts. The number above the bars indicates the exact value of the improvement metric.

the conditioning approach exhibits higher forecasting performances (i.e. the *CAR(NWP)* model is better than *AR + NWP* for both metrics). Based on observations from Section 4.3, this is supposed to result from a better integration of features having a non-linear correlation with the production. On the contrary, when dealing with non-linear forecasting models, it is better to include NWPs as explanatory features. Despite the improved performances due to the conditioning approach, the *CAR(NWP)* model is outperformed by the *RF+NWP* model.

It is possible to further improve the forecasting performance of the *CAR(NWP)*) model by adding NWPs as extra features (which leads to the *CAR(NWP)+NWP*

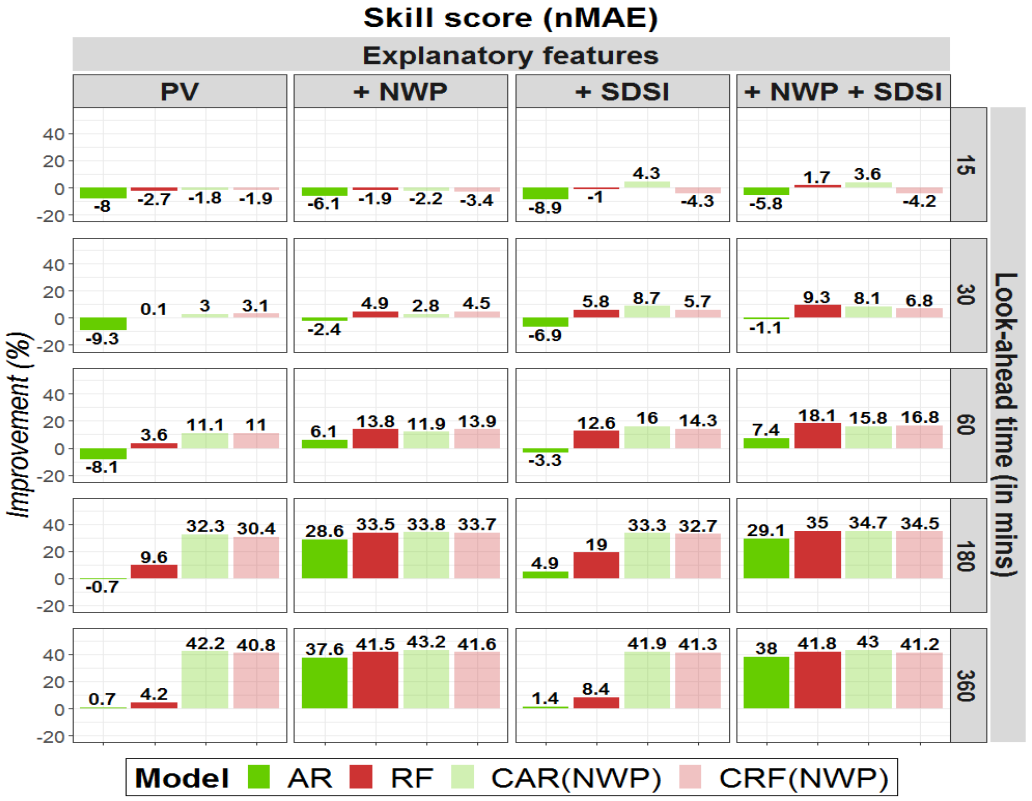

Figure 8: nMAE skill scores with regard to the persistence model. Dark colours symbolise forecasting models trained on the whole dataset, while light colours stand for WHCO models. Columns represent the explanatory features, while rows indicate the lead time of the forecasts. The number above the bars indicates the exact value of the improvement metric.

model). This configuration leads to similar performances as those reached by the *RF+NWP* model (i.e. in average, a 0.06% and −0.46% difference is observed between *CAR(NWP)+NWP* and *RF+NWP* models in term of nRMSE and nMAE). In this respect, feeding the *CRF(NWP)* model with NWPs slightly decrease its forecasting skills, possibly due to overfitting issues.

As a result, when dealing with production observations and NWPs, it is more profitable to consider non-linear models fed with explanatory features or WHCO linear models with weather-based explanatory features. At this point, the choice of the model results mainly from a computational cost and

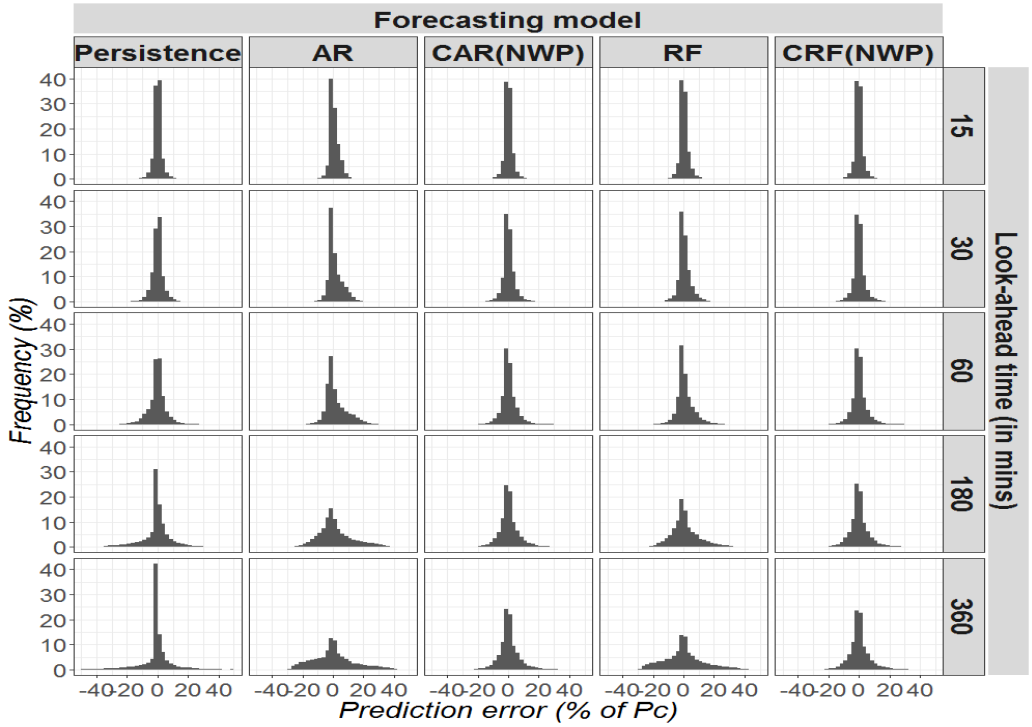

Figure 9: Normalised prediction error distribution (with bins representing 2.5% of the rated power) for the *Persistence*, *AR*, *CAR(NWP)*, *RF* and *CRF(NWP)* models according to the look-ahead times.

interpretability compromise.

4.4.4. Production Observations + NWPs + SDSI

When dealing with production, NWPs and SDSI inputs, it is obvious that including these data as explanatory features in an *AR* model leads to the worst performances both in term of nRMSE and nMAE. Once again, the WHCO approach improves significantly the forecasting performances of the linear model. In addition, we observe that the nRMSE score of the *CAR(NWP)+SDSI+NWP* is slightly better than the one of *CAR(NWP)+SDSI*.

Conditioning non-linear models results in a performance drop compared to the *RF+NWPs+SDSI* model for both metrics.

To conclude, the *CAR(local)+NWPs+SDSI* model appears to be a good option inasmuch as it performs better on very short-term horizons and exhibits similar skill scores to the *RF+NWPs+SDSI* model for higher forecast-

ing horizons.

Table 1 summarises the different findings regarding the optimal model choice considering the inputs taken into account.

Table 1: Summary of the best model configuration depending on the type of inputs.

Inputs	**Best configuration**
PV production	RF
PV production+NWP	CAR(NWP)+NWP / RF+NWP
PV production+SDSI	RF+SDSI
PV production+NWP+SDSI	CAR(NWP)+NWP+SDSI / RF+NWP+SDSI

5. Conclusions

We investigated the influence of NWPs integration within the forecasting chain of PV generation through two strategies: either as explanatory features or as state variables to condition the statistical model to the atmospheric states. To improve forecasting performances for short lead times, ST information is considered in the form of SDSI, from which, geographically dispersed pixels are extracted via a selection procedure newly applied to the PVPF domain. To assess the coupling between inputs integration and type of information, we developed a modular architecture which conditions off-the-shelf regression models to weather parameters. This conditioning approach is a way to include physics-based information within statistical tools and to dynamically update their parameters according to the weather state.

This conditioning procedure appears to be an appealing approach to extend the forecasting performances of linear models such as the AR model (i.e. *CAR(NWP)*). On the contrary, non-linear models such as the RF model perform better when fed with explanatory features (i.e. *RF+NWP*). This finding tends to support that weather conditioning is not adapted with non-linear models, therefore, it should be used carefully in the literature. The WHCO approach reveals interesting performances when it comes to consider ST observations as additional inputs. In that case, the *CAR(NWP)+SDSI+NWPs* model turns out to be more efficient than non-linear approaches such as *RF+SDSI+NWPs* for short lead times. This may be explained by the ability of the model to select SDSI-based features in line with wind propagation [19].

Further works could extend the present investigations by applying the proposed methodology with probabilistic forecasting models or for higher forecast horizons.

Acknowledgement

The authors wish to thank Compagnie Nationale du Rhône (CNR) for financially supporting this research. The authors would like to thank the ECMWF for providing the NWPs data as well as Transvalor for providing the SDSI dataset.

References

[1] IRENA, Renewable power generation costs in 2017, 00008 (2017).

[2] S. R. Sinsel, R. L. Riemke, V. H. Hoffmann, Challenges and solution technologies for the integration of variable renewable energy sources—a review, Renewable Energy 145 (2020) 2271–2285. doi:10.1016/j.renene.2019.06.147.
URL https://linkinghub.elsevier.com/retrieve/pii/S0960148119309875

[3] M. J. M. Al Essa, Power management of grid-integrated energy storage batteries with intermittent renewables, Journal of Energy Storage 31 (2020) 101762, 00011. doi:10.1016/j.est.2020.101762.
URL https://linkinghub.elsevier.com/retrieve/pii/S2352152X20315991

[4] M. S. Viana, G. Manassero, M. E. Udaeta, Analysis of demand response and photovoltaic distributed generation as resources for power utility planning, Applied Energy 217 (2018) 456–466, 00043. doi:10.1016/j.apenergy.2018.02.153.
URL https://linkinghub.elsevier.com/retrieve/pii/S0306261918302873

[5] M. Diagne, M. David, P. Lauret, J. Boland, N. Schmutz, Review of solar irradiance forecasting methods and a proposition for small-scale insular grids, Renewable and Sustainable Energy Reviews 27 (2013) 65–76, 00415. doi:10.1016/j.rser.2013.06.042.
URL http://www.sciencedirect.com/science/article/pii/S1364032113004334

[6] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting, Solar Energy 136 (2016) 78–111, 00326. doi:10.1016/j.solener.2016.06.069.
URL http://www.sciencedirect.com/science/article/pii/S0038092X1630250X

[7] S. Sobri, S. Koohi-Kamali, N. A. Rahim, Solar photovoltaic generation forecasting methods: A review, Energy Conversion and Management 156 (2018) 459–497, 00072. doi:10.1016/j.enconman.2017.11.019.
URL http://www.sciencedirect.com/science/article/pii/S0196890417310622

[8] R. Tawn, J. Browell, A review of very short-term wind and solar power forecasting, Renewable and Sustainable Energy Reviews 153 (2022) 111758, 00000. doi:10.1016/j.rser.2021.111758.
URL https://www.sciencedirect.com/science/article/pii/S1364032121010285

[9] R. J. Bessa, A. Trindade, V. Miranda, Spatial-Temporal Solar Power Forecasting for Smart Grids, IEEE Transactions on Industrial Informatics 11 (1) (2015) 232–241, 00144. doi:10.1109/TII.2014.2365703.

[10] L. M. Aguiar, B. Pereira, P. Lauret, F. Díaz, M. David, Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting, Renewable Energy 97 (2016) 599–610, 00109. doi:10.1016/j.renene.2016.06.018.
URL http://www.sciencedirect.com/science/article/pii/S0960148116305390

[11] X. G. Agoua, R. Girard, G. Kariniotakis, Short-Term Spatio-Temporal Forecasting of Photovoltaic Power Production, IEEE Transactions on Sustainable Energy 9 (2) (2018) 538–546, 00029. doi:10.1109/TSTE.2017.2747765.

[12] D. Yang, Ultra-fast preselection in lasso-type spatio-temporal solar forecasting problems, Solar Energy 176 (2018) 788–796, 00018. doi:10.1016/j.solener.2018.08.041.
URL http://www.sciencedirect.com/science/article/pii/S0038092X18308120

[13] S. Chai, Z. Xu, Y. Jia, W. K. Wong, A Robust Spatiotemporal Forecasting Framework for Photovoltaic Generation, IEEE Transactions on Smart Grid (2020) 1–100014. doi:10.1109/TSG.2020.3006085.

[14] P. Bacher, H. Madsen, H. A. Nielsen, Online short-term solar power forecasting, Solar Energy 83 (10) (2009) 1772–1783, 00514. doi:10.1016/j.solener.2009.05.016.
URL https://linkinghub.elsevier.com/retrieve/pii/S0038092X09001364

[15] J. Shi, W.-J. Lee, Y. Liu, Y. Yang, P. Wang, Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines, IEEE Transactions on Industry Applications 48 (3) (2012) 1064–1069, 00519. doi:10.1109/TIA.2012.2190816.

[16] F. Wang, Z. Zhen, Z. Mi, H. Sun, S. Su, G. Yang, Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting, Energy and Buildings 86 (2015) 427–438, 00126. doi:10.1016/j.enbuild.2014.10.002.
URL http://www.sciencedirect.com/science/article/pii/S0378778814008226

[17] A. Nespoli, E. Ogliari, S. Leva, A. Massi Pavan, A. Mellit, V. Lughi, A. Dolara, Day-Ahead Photovoltaic Forecasting: A Comparison of the Most Effective Techniques, Energies 12 (9) (2019) 1621, 00040. doi:10.3390/en12091621.
URL https://www.mdpi.com/1996-1073/12/9/1621

[18] H. Aprillia, H.-T. Yang, C.-M. Huang, Short-Term Photovoltaic Power Forecasting Using a Convolutional Neural Network–Salp Swarm Algorithm, Energies 13 (8) (2020) 1879, 00018 Number: 8 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en13081879.
URL https://www.mdpi.com/1996-1073/13/8/1879

[19] R. Amaro e Silva, S. E. Haupt, M. C. Brito, A regime-based approach for integrating wind information in spatio-temporal solar forecasting models, Journal of Renewable and Sustainable Energy 11 (5) (2019) 056102, 00001. doi:10.1063/1.5098763.
URL http://aip.scitation.org/doi/10.1063/1.5098763

[20] M. P. Almeida, O. Perpiñán, L. Narvarte, PV power forecast using a nonparametric PV model, Solar Energy 115 (2015) 354–368, 00171. doi:10.1016/j.solener.2015.03.006.
URL http://www.sciencedirect.com/science/article/pii/S0038092X15001218

[21] K. Bellinguer, R. Girard, G. Bontron, G. Kariniotakis, Short-term Forecasting of Photovoltaic Generation based on Conditioned Learning of Geopotential Fields, in: 2020 55th International Universities Power Engineering Conference (UPEC), 2020, pp. 1–6, 00001. doi:10.1109/UPEC49904.2020.9209858.

[22] C. Chen, S. Duan, T. Cai, B. Liu, Online 24-h solar power forecasting based on weather type classification using artificial neural network, Solar Energy 85 (11) (2011) 2856–2870, 00444. doi:10.1016/j.solener.2011.08.027.
URL http://www.sciencedirect.com/science/article/pii/S0038092X11003008

[23] A. Eschenbach, G. Yepes, C. Tenllado, J. I. Gómez-Pérez, L. Piñuel, L. F. Zarzalejo, S. Wilbert, Spatio-Temporal Resolution of Irradiance Samples in Machine Learning Approaches for Irradiance Forecasting, IEEE Access 8 (2020) 51518–51531, 00004. doi:10.1109/ACCESS.2020.2980775.

[24] T. Carriere, C. Vernay, S. Pitaval, G. Kariniotakis, A Novel Approach for Seamless Probabilistic Photovoltaic Power Forecasting Covering Multiple Time Frames, IEEE Transactions on Smart Grid (2019) 1–100022. doi:10.1109/TSG.2019.2951288.

[25] P. Horton, M. Jaboyedoff, C. Obled, Using genetic algorithms to optimize the analogue method for precipitation prediction in the Swiss Alps, Journal of Hydrology 556 (2018) 1220–1231, 00021. doi:10.1016/j.jhydrol.2017.04.017.
URL https://www.sciencedirect.com/science/article/pii/S0022169417302391

[26] S. Alessandrini, L. Delle Monache, S. Sperati, G. Cervone, An Analog Ensemble for Short-term Probabilistic Solar Power Forecast, Applied Energy 157 (2015) 95–110, 00129. doi:10.1016/j.apenergy.2015.08.011.
URL http://www.sciencedirect.com/science/article/pii/S0306261915009368

[27] B. O. Akyurek, A. S. Akyurek, J. Kleissl, T. S. Rosing, TESLA - Taylor expanded solar analog forecasting, in: 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), 2014, pp. 127–132, 00016. doi:10.1109/SmartGridComm.2014.7007634.

[28] L. Delle Monache, F. A. Eckel, D. L. Rife, B. Nagarajan, K. Searight, Probabilistic Weather Prediction with an Analog Ensemble, Monthly Weather Review 141 (10) (2013) 3498–3516, 00000. doi:10.1175/MWR-D-12-00281.1.
URL https://journals.ametsoc.org/doi/10.1175/MWR-D-12-00281.1

[29] W. S. Cleveland, C. Loader, Smoothing by Local Regression: Principles and Methods, in: W. Härdle, M. G. Schimek (Eds.), Statistical Theory and Computational Aspects of Smoothing, Contributions to Statistics, Physica-Verlag HD, Heidelberg, 1996, pp. 10–49, 00566. doi:10.1007/978-3-642-48425-4_2.

[30] G. E. P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day, Inc., San Francisco, CA, USA, 1990, 47056.

[31] R. Tibshirani, Regression Shrinkage and Selection via the Lasso, Journal of the Royal Statistical Society. Series B (Methodological) 58 (1) (1996) 267–288, 27385.
URL https://www.jstor.org/stable/2346178

[32] L. Breiman, Random Forests, Machine Learning 45 (1) (2001) 5–32, 43710. doi:10.1023/A:1010933404324.
URL https://doi.org/10.1023/A:1010933404324

[33] K. Bellinguer, V. Mahler, S. Camal, G. Kariniotakis, Probabilistic Forecasting of Regional Wind Power Generation for the EEM20 Competition: a Physics-oriented Machine Learning Approach, in: 2020 17th International Conference on the European Energy Market (EEM), 2020, pp. 1–6, 00000. doi:10.1109/EEM49802.2020.9221960.

[34] 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, J.-J. Morcrette, McClear a new model estimating downwelling solar radiation at ground level in clear-sky conditions, Atmospheric Measurement Techniques 6 (2013) 2403–2418, 00000. doi:10.5194/amt-6-2403-2013.
URL https://hal-mines-paristech.archives-ouvertes.fr/hal-00862906/document

[35] R. Perez, P. Ineichen, R. Seals, J. Michalsky, R. Stewart, Modeling daylight availability and irradiance components from direct and global irradiance, Solar Energy 44 (5) (1990) 271–289, 01494. doi:10.1016/0038-092X(90)90055-H.
URL http://www.sciencedirect.com/science/article/pii/0038092X9090055H

[36] D. Yang, S. Alessandrini, J. Antonanzas, F. Antonanzas-Torres, V. Badescu, H. G. Beyer, R. Blaga, J. Boland, J. M. Bright, C. F. Coimbra, M. David, Frimane, C. A. Gueymard, T. Hong, M. J. Kay, S. Killinger, J. Kleissl, P. Lauret, E. Lorenz, D. van der Meer, M. Paulescu, R. Perez, O. Perpiñán-Lamigueiro, I. M. Peters, G. Reikard, D. Renné, Y.-M. Saint-Drenan, Y. Shuai, R. Urraca, H. Verbois, F. Vignola, C. Voyant, J. Zhang, Verification of deterministic solar forecasts, Solar Energy 210 (2020) 20–37, 00048. doi:10.1016/j.solener.2020.04.019.
URL https://linkinghub.elsevier.com/retrieve/pii/S0038092X20303947

[37] P. Blanc, B. Gschwind, M. Lefèvre, L. Wald, The HelioClim Project: Surface Solar Irradiance Data for Climate Applications, Remote Sensing 3 (2) (2011) 343–361, 00129. doi:10.3390/rs3020343.
URL https://hal-mines-paristech.archives-ouvertes.fr/hal-00566995/document

[38] K. Bellinguer, R. Girard, G. Bontron, G. Kariniotakis, Short-Term Photovoltaic Generation Forecasting Enhanced by Satellite Derived Irradiance, in: 26th International Conference & Exhibition on Electricity Distribution (CIRED 2021), CIRED, Virtual Event, Switzerland, 2021, tex.ids= bellinguer_short-term_2021-1.
URL https://hal.archives-ouvertes.fr/hal-03407898

[39] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2020).
URL https://www.R-project.org/

[40] L. Vallance, B. Charbonnier, N. Paul, S. Dubost, P. Blanc, Towards a standardized procedure to assess solar forecast accuracy: A new ramp and time alignment metric, Solar Energy 150 (2017) 408–422, 00037. doi:10.1016/j.solener.2017.04.064.
URL http://www.sciencedirect.com/science/article/pii/S0038092X17303687