Intra-day solar probabilistic forecasts including local short-term variability and satellite information

R. Alonso-Suárez[a,*], M. David[b], V. Branco[a,c], P. Lauret[b]

[a] *Laboratorio de Energía Solar, Facultad de Ingeniería, UDELAR. J. H. y Reissig 565, Montevideo, Uruguay*
[b] *University of La Réunion - PIMENT laboratory, 15, avenue René Cassin, 97715 Saint-Denis*
[c] *Laboratorio Solar, ITR Centro-Sur, Universidad Tecnológica. F. Maciel esq. L. Morquio, CP 97000, Durazno, Uruguay.*

Abstract

In this work, three models are built to produce intra-day probabilistic solar forecasts with lead times ranging from 10 minutes to 3 hours with a granularity of 10 minutes. The first model makes only use of past ground measurements. The second model upgrades the first one by adding a variability metric obtained also from the past ground measurements. The third model takes as additional input the satellite albedo. A non parametric approach based on the linear quantile regression technique is used to generate the set of quantiles that summarize the predictive distributions of the global solar irradiance at a horizontal plane (GHI). The probabilistic models are evaluated on several sites that experience very different climatic conditions. It is shown that incorporating variability significantly reduces the width of interval predictions. The addition of satellite information further improves the quality of the probabilistic forecasts.

Keywords: GHI, probabilistic forecast, ground measurement, solar variability, satellite images.

1. Introduction

Achieving a high penetration of solar energy into electricity grids poses a challenge as the inherent variability and lack of predictability of solar power affects the supply/demand balance. As a result, the grid operator needs to take proper actions to maintain this balance either by providing additional reserves or by adjusting the output of controllable generators. One of the strategies to manage the impacts of solar variability is to predict the short-term future solar irradiance and the corresponding solar PV power output at different time horizons. This provides valuable information for grid operators to mitigate the intermittency at lower cost, ensuring the electricity system stability. PV power forecasting is also required to anticipate the PV production availability for trading and to reduce the uncertainty of its selling price. Therefore, improvements in solar forecasting methods are required to increase the value of solar PV, enabling a higher penetration of this technology into electricity grids.

*Corresponding author: R. Alonso-Suárez, r.alonso.suarez@gmail.com

Solar predictions are intended to anticipate at time (t) the future behavior (fluctuations) of ground level solar irradiance as a result of the complex clouds and atmosphere's dynamics. Two types of forecasting methods can be considered namely deterministic and probabilistic. Both methods attempt to predict the future outcome of a variable but information on the probability distribution of the prediction is only present in the probabilistic forecast. Many research efforts have been devoted to the development of models for generating solar deterministic forecasts. However, a forecast is inherently uncertain and a proper assessment of its probability distribution offers the grid operator a more informed decision-making framework. Thus, a deterministic forecast plus predictions intervals for different confidence levels provides more useful information to manage the renewable sources uncertainty in a cost efficient and reliable manner (Botterud, 2017). Contrary to wind power forecasting where probabilistic forecasting appears to be a mature subject (Pinson et al., 2007; Jung and Broadwater, 2014; Iversen et al., 2016), probabilistic solar forecasting is still in its first stages (Antonanzas et al., 2016; Hong et al., 2016).

For one day up to several days ahead, Numerical Weather Predictions (NWP) are routinely used to produce deterministic or probabilistic forecasts. The latter takes the form of ensembles of forecasts obtained by running a NWP model while changing slightly the initial conditions (Leutbecher and Palmer, 2008). Geostationary satellite images can be used for intra-day deterministic forecast by means of sophisticated Cloud Motion Fields (CMF) estimations (Lorenz et al., 2004; Kühnert et al., 2013). As the present article focuses on very short-term intra-day solar forecasts (i.e. for time horizons ranging from some minutes to several hours) that make simple use of satellite information, we will not detail here the research works related to the above mentioned type of models. The present work is under the framework of the time-series analysis forecast methods (Inman et al., 2013; Diagne et al., 2013; Voyant et al., 2017), which are suitable for very short-term forecast. These kind of methods rely on statistical procedures based on linear time series modeling or machine learning techniques, and their inputs should be chosen wisely. Most works in this field, either for deterministic or probabilistic forecast, use solar ground measurements as inputs (Reikard, 2009; Bacher et al., 2009; Mellit and Pavan, 2010; Pedro and Coimbra, 2012; Huang et al., 2013; Lauret et al., 2015; David et al., 2016; Grantham et al., 2016; Chu and Coimbra, 2017; Lauret et al., 2017; David et al., 2018; Pedro et al., 2018). Indeed, these works have shown that the forecast irradiance is highly correlated with the actual and past ground measurements. The availability of ground measurements is not an operational limitation, since automatic weather stations are usually located at large-scale solar energy project's sites and their measurements are reported automatically to dispatch centers. Furthermore, solar irradiance can be estimated by geostationary satellite images at any location in a hourly or intra-hour time basis (Perez et al., 2002; Rigollier et al., 2004; Alonso-Suárez et al., 2012; Qu et al., 2017), and can replace at some extent the operational needs of ground measurements. Some deterministic proposals also include other exogenous variables that can be accessed operationally, such as satellite information (Dambreville et al., 2014; Aguiar et al., 2015). This has not been done for solar probabilistic forecast, which is one contribution of the present work.

Probabilistic forecast techniques aim to predict the future probability distributions of a prediction variable. The predictive distribution can be represented either by a Cumulative Distribution Function (CDF) or a Probability Density Function (PDF). Similarly to deterministic forecasts, different types of probabilistic models can be used depending on the forecast horizon. Golestaneh et al. (2016) empirically observed that solar forecasting errors do not follow any of the most common parametric probability densities. Thus, they proposed a non-parametric approach based on neural networks to produce a very short-term PV power probabilistic forecasting using as inputs the power output and meteorological measurements time-series, including solar irradiance. The proposed forecasting method was evaluated for lead times between a few minutes to one hour ahead and proved to produce a better probabilistic forecast than a selected set of benchmark models. Grantham et al. (2016) also proposed a non-parametric approach, but based on a specific bootstrap method to build predictive global horizontal irradiance (GHI) probability distributions. Using solar irradiance ground measurements and a map of the solar positions in the sky, they produce reliable (used here in the sense of the reliability property, explained in Subsection 4.1) probabilistic forecast for a time horizon of one hour. As the bootstrap method tends to reproduce the distribution of the climatology observed under specific conditions, the predictions are not sharp (see Subsection 4.2 for the description of the sharpness property). Chu and Coimbra (2017) developed an ensemble model based on the k-nearest-neighbors (kNN) algorithm for generating very short-term direct normal irradiance (DNI) probabilistic forecasts from 5 to 20 minutes ahead. This forecast proposal is based on the time-lagged DNI measurements, diffuse irradiance measurements and three features extracted from all-sky camera images. David et al. (2016) used a parametric approach based on a combination of two linear time series model (ARMA and GARCH) to generate 10 minutes to 6 hours ahead GHI probabilistic forecasts using only past ground data. Recently, David et al. (2018) made a comparison of 20 intra-day solar probabilistic models that used only solar irradiance ground measurements and found that models based on linear quantile regression were one of the best performers. The authors also conclude that future improvements could be obtained by including relevant exogenous variables in the models. Also, Lauret et al. (2017) showed that the quality of intra-day probabilistic forecasts can be improved by considering both past ground measurements and day-ahead NWP forecasts (provided by the European Center for Medium-Range Weather Forecast, ECMWF) in the set of explanatory variables. Pedro et al. (2018) used a machine learning approach to generate very short-term probabilistic forecasts of GHI and DNI. The input features of the proposed models were derived from the irradiance time series and sky images. In particular, a variability index was calculated from the GHI time series but the impact of this predictor was not assessed on the quality of the forecasts.

To the best of our knowledge, no works have been published regarding the use of satellite data and/or short-term variability information for intra-day solar probabilistic forecasts. Therefore, this work investigates if a combination of ground telemetry with variability and satellite information could improve the quality of the probabilistic forecasts. As the goal here is to assess the added value of such a combination, we restrict ourselves to a simple probabilistic statistical technique to build the different models, namely the linear quantile regression (Koenker and Bassett, 1978). More precisely, we build three probabilistic models that generate three hours ahead probabilistic forecasts using a 10 minutes granularity. The first model makes use of actual and past solar ground measurements. The second one adds the solar variability as an additional predictor to the past ground measurements while the third one enriched the second by adding spatially-averaged satellite albedo as an exogenous input. In order to evaluate the quality of the probabilistic forecasts, we use the evaluation framework provided by Pinson et al. (2007), recently adapted by Lauret et al. (2019) for the solar irradiance forecasting field.

This article is organized as follows. In Section 2 the sites and data description are presented. Section 3 briefly describes the three probabilistic models based on the quantile regression technique. Section 4 details the different metrics used to evaluate the quality of the solar probabilistic forecasts and Section 5 presents the results based on the evaluation framework. Finally, in Section 6 we provide our concluding remarks.

2. Data

To study the effect of adding short-term variability to probabilistic forecast, only solar irradiance ground measurements are required. For this, we used high quality data recorded at several sites that exhibit completely different sky conditions and solar variability regimes. The utilization of satellite information was tested in two of these sites, where we have geostationary satellite information that coincide with the high quality measurements. In this section we describe the solar irradiance data sets, the way we compute the short-term variability input from the past solar time-series and the satellite information.

2.1. Ground measurements

High quality solar irradiance measurements are essential for any resource assessment or forecasting study. In this work we used data from seven stations in different part of the world which equipment meets the BSRN quality criteria. Two of these sites are from the NOAA's SURFRAD solar radiation network (Desert Rock and Fort Peck), three are from insular sites (Oahu, Fouillole and Tampom) and two are from subtropical climates in South America (Salto and São Martinho da Serra). Also, four of these stations are in the northern hemisphere and three are in the southern hemisphere. From the complete set of measurements that these stations record, namely, the global irradiance at a horizontal plane (GHI), the diffuse irradiance at a horizontal plane (DHI) and the direct normal irradiance (DNI), only the GHI component is considered in this work. This variable is measured in all the stations using spectrally flat class A Pyranometers (according to the new ISO 9060:2018 standard) and receive the daily maintenance required for a high quality measurement.

It is common practice in the field to work with the clear sky index, k_c, rather than the GHI data. The GHI time series, G_h, have a diurnal and seasonal geometrical behavior that result from the Sun's apparent movement and introduces a deterministic component into the times series that can be removed using accurate estimations of the clear sky irradiance. Indeed, an estimation of the clear sky GHI, G_h^{csk}, can be used to calculate k_c as defined in Eq. (1), removing the geometrical behavior and the cloudless atmosphere's slow variations, isolating the fluctuations only due to cloudiness;

$$k_c(t) = \frac{G_h(t)}{G_h^{\mathrm{csk}}(t)}. \tag{1}$$

With this methodology, the forecasting models can be dedicated to predict the stochastic component of the GHI due to the presence or not of cloud cover, leaving the geometric and the deterministic part to be represented by the clear sky model. Here, the McClear model's (Lefèvre et al., 2013) clear sky estimations are used. McClear's clear sky estimates are publicly available with worldwide coverage in the CAMS platform (Copernicus Atmosphere Monitoring Service, http://www.soda-pro.com), from where we downloaded them. This model is based on sophisticated radiative transfer calculations, making use of aerosol optical properties, water vapor and ozone data obtained also from the CAMS products. This allows the model to reproduce the daily clear sky atmosphere variability and to provide accurate GHI clear sky estimates. The input variables' uncertainty can have an important impact on a clear sky model accuracy (Polo et al., 2014; Zhong and Kleissl, 2015; Gueymard, 2019), so prevision on their availability and quality need to be taken into account to choose it wisely. The validation can be done by using local measurements of these variables, if available, such as those provided by the AERONET network (Gueymard and Yang, 2020). McClear clear sky estimates have been evaluated in several parts of the world by using high-quality solar irradiance ground measurements, showing very good accuracy for solar assessment (Cros et al., 2013; Ineichen, 2016; Antonanzas-Torres et al., 2019; Laguarda et al., 2020). The operational version of this model takes the form of a look-up-table based on different input variables, which allows to quickly compute the estimations provided by the platform. It shall be noted that other clear sky models may be used, for instance, the modified Kasten model (Kasten, 1980; Ineichen and Perez, 2002) or the ESRA model (Rigollier et al., 2000), among others. Here, for the sake of simplicity, we choose to retrieve the clear sky estimates from the CAMS website. A recent article by Yang (2020) analyzed several clear sky models for its value to calculate k_c for solar forecast, using measurements from sites with different climate characteristics. It was found that from the forecasting performance point of view, there is not much difference in using one or other clear sky model, and hence the McClear model's estimates are the best option due to its easy access.

State-of-the-art forecasting methods described in Section 1 generally make use of the clear sky index, k_c, rather than the clearness index, k_t, being the latter the normalization of the GHI by its corresponding extraterrestrial irradiance at horizontal plane. Although k_t is a popular variable for decomposition and transposition irradiance models, mostly due to its simple calculation, it exhibits significant correlations with the solar zenith angle and the air mass, resulting in artifacts during early mornings and late afternoons. In fact, k_t does not provide the well-behaved deseasonalized signal that is required for forecasting methods. The clear sky index overcomes this issue by providing an almost stationary time series that is representative of the cloudiness only. Hence, we prefer here the use of k_c instead of k_t, which is commonly admitted in the solar forecasting community (Inman et al., 2013; Diagne et al., 2013; Sengupta et al., 2015; Yang et al., 2018) and usually preferred by solar forecasting researchers (Yang, 2020).

Ground measurements were filtered at 1 minute intervals according to the BSRN criteria to exclude atypical and physically impossible samples from GHI time series. The 10 minutes GHI averages required for this work were obtained from these filtered 1 minute data by averaging in the corresponding time slot. Also, we discard some few outliers (less than 1%) by setting maximum thresholds for the 10-minutes clear sky index and the short-term local variability, calculated as it is explained in Subsection 2.2. Finally, we discard data with solar altitude lower than 10^o, as pyranometers measurement uncertainties are higher for low solar elevations due to the cosine error of the equipment.

Table 1 summarizes the main characteristics of the different sites considered in this work, including its location, provider and the site's climate type. The code (first row) will be used as a short reference to the station. The short-term variability of the time series is also provided in Table 1, and it is derived from each station filtered data as the standard deviation of the 10 minutes clear sky changes: $\sigma_c = \text{std}\{\Delta k_c\}$. A site with variability $\sigma_c > 0.2$ is considered as experiencing very unstable GHI conditions (Hoff and Perez, 2012). As shown in Table 1, the solar irradiance variability of the Oahu, Fouillole and Tampon stations are around or above this threshold, which is typical of insular sites where partly cloudy conditions prevail. On the other hand, Desert Rock is a site where clear sky days prevail, thus it has a small GHI variability. The rest of the sites have intermediate conditions in terms of GHI variability (see Table 1), with a balance mix of clear sky, partly cloudy and cloudy days. Finally, for each station, two years of data, as shown in the last two rows of Table 1, are used as follow: the oldest year is used to train the quantile regression models and the most recent year is used to test the results.

Table 1: Main characteristics of the sites under study.

site name	**Salto**	**Desert Rock**	**Fort Peck**	**Oahu**	**Fouillole**	**Tampon**	**São Marti. da Serra**
code	LE	DR	FP	OA	FO	TM	MS
provider	LES	NOAA	NOAA	NREL	LARGE	PIMENT	SODA
latitude	31.3 S	36.6 N	48.3 N	21.3 N	16.6 N	21.3 S	29.4 S
longitude	57.9 W	119.0 W	105.1 W	158.1 W	61.5 W	55.5 E	53.8 W
elevation	56 m	1007 m	634 m	11 m	6 m	550 m	489 m
climate type	subtropical	arid	continental	tropical	tropical	tropical	subtropical
variability	0.140	0.113	0.145	0.207	0.200	0.194	0.151
training year	2016	2012	2012	2010	2010	2012	2014
testing year	2017	2013	2013	2011	2011	2013	2015

2.2. Local short-term variability

One of the variables inspected in this work is the 10 minutes variability considered for the near-time past measurements to time (t). This variable aims to introduce information that quantifies the instability of the cloud regime close to time (t), the moment of the forecast. Thus, a time lagged variability index is used, defined as,

$$\sigma_c^{l_v}(t) = \text{std}\{\Delta k_c(t), \ldots, \Delta k_c(t - l_v)\}, \tag{2}$$

where $\Delta k_c(t)$ is the time series of the 10 minutes changes of k_c and l_v is the amount of past samples considered in the calculation of the standard deviation. We use $l_v = 5$, the same quantity of lags that will be used for building the quantile regression models, so this local short-term variability accounts for the past one hour from time (t). At the beginning of the day this calculation may not be indicative of the current sky instability, as the moving window would consider samples from the present day mixed with samples of the previous day (night-time values are removed by the solar altitude filter), which may have different sky conditions. To solve this, the computation of $\sigma_c^{l_v}$ only takes into account the non-filtered daylight samples of the current day, so the time window that is used for its calculation at the beginning of the day is adapted to meet this requirement. The variability for the first sample of the day (after filtering, it will be a daylight sample) is set to zero as a default value.

2.3. Satellite information

Visible channel satellite images measure the solar radiance reflected at each pixel of the Earth surface. By knowing the extraterrestrial irradiance at the top of the atmosphere, it is possible to derive the Earth surface reflectance, known as albedo, ρ_p. This information provided by geostationary satellites is the main input for satellite based models for intra-hour solar irradiance assessment (Perez et al., 2002; Ceballos et al., 2004; Rigollier et al., 2004; Alonso-Suárez et al., 2014; Qu et al., 2017). In this work, rather than using the solar satellite estimates, we use directly the satellite albedo as an input for the probabilistic models. This seeks to quantify the impact of satellite cloudiness in the forecast performance without using a solar irradiance model that may be adding uncertainty to the problem. In absence of snow or any other kind of high albedo terrain, a low albedo represents the ground surface and a high albedo represents the presence of clouds, as they reflect more radiation to the outer space than the ground. Such is the case of the two South American sites where satellite information is considered, whose ground albedo is around 5-15 % during all the year. Albedo samples ranging from 20 % to 100 % represent the presence of cloudiness, in its various types.

For the two considered sites, LE and MS, satellite information is used by averaging the satellite albedo pixels in a 10 min$\times$10 min latitude-longitude cell around the site. This cell size is known to be the optimal to reduce the uncertainty of hourly irradiation assessments in the region (Laguarda et al., 2020; Alonso-Suárez, 2017) and it is intended to represent the average hourly cloudiness using an ergodic assumption. As the satellite cloudiness is spatially averaged, information of the near future cloudiness over the target site is directly included. This is a very simple way to introduce satellite derived data as exogenous variable into solar forecasts. In this way, both satellite information and short-term variability lags represent approximately one hour. Finally, it has to be noted that satellite images for South America during the considered period (till year 2017) have a rate of two per hour, so the required 10 minutes granularity was obtained by linear interpolation using the available albedo samples. The images were downloaded in raw format from the NOAA CLASS website (https://www.avl.class.noaa.gov) and the calibration procedures for obtaining the corrected albedo were applied as recommended by NOAA (Wu and Sun, 2005).

3. Probabilistic forecast

Solar probabilistic forecast is assessed by estimating at time (t) the future solar irradiance probability distribution at time $(t + \Delta t)$, where Δt is the forecast horizon. Hence, a probabilistic forecast can be defined as the prediction of its future cumulative distribution function (CDF). Two main possibilities exist to estimate this CDF. The first one consists in assuming a parametric law for the CDF, generally a Gaussian distribution, and tune its parameters based on a training data set, i.e. the mean and the standard deviation for the Gaussian assumption. The second one is a non-parametric approach where no assumption about the shape of the future distribution is made. In this case, the CDF can be obtained by estimating each quantile separately. This set of discrete quantiles can be estimated with simple techniques like the Linear Quantile Regression (LQR) or with more sophisticated machine learning techniques like Gradient Boosting Decision Trees (GBDT) or Support Vector Machines (SVM), among others.

In this work we follow the second option for estimating the future CDF, so, a set of quantiles have to be estimated for each forecast horizon. Formally, a quantile q_τ at probability level τ is defined as,

$$q_\tau = F^{-1}(\tau) = \inf\ \{y : F(y) >= \tau\}\,, \tag{3}$$

where F is the CDF of a random variable Y, defined as $F(y) = \Pr(Y \leq y)$. In other words, a quantile q_τ indicates that there is a τ probability that the observation falls below the quantile q_τ. In this work, these quantiles are estimated for the clear sky index (k_c) and each time horizon, for the nominal proportions $\tau = [0.1, 0.2, \ldots, 0.9]$, resulting in different sets of k_c quantiles $q_\tau = [q_1, q_2, \ldots, q_9]$ that define each CDF. More precisely, the output (at time t) of the different probabilistic models that are generated in this work is the ensemble of nine clear sky index quantiles $q_\tau(t+\Delta t)$ for each forecasting time horizon $\Delta t = 10, 20, \cdots, 180$ minutes. For the sake of completeness and proper calculation of the performance metrics, the quantiles q_0 and q_{10} that correspond to the nominal proportions $\tau = 0.0$ and $\tau = 1.0$ where set to zero and to the maximum k_c value in the time series, respectively. Then, using Eq. (1), the k_c quantiles are converted to the required GHI quantiles. Prediction intervals with different nominal coverage rates can be inferred from the set of quantiles. Prediction intervals give a range of values in which the true value of GHI is expected to lie with a certain probability, namely, the nominal coverage rate. In this study, we chose prediction intervals with a nominal coverage rate of 80% as it leads to a good compromise as stated by (Pinson et al., 2007). Hence, the $(1-\alpha) \times 100\%$ central prediction interval is generated by taking the $\frac{\alpha}{2}$ quantile as the lower bound and the $1-\frac{\alpha}{2}$ quantile as the upper bound. In this case, the 80% confident interval ($\alpha = 0.2$) is obtained by considering the q_1 and q_9 GHI quantiles.

In this work, we explore the effect of including the past short-term variability and present time spatially-averaged satellite albedo as explanatory variables for solar irradiance probabilistic forecast. Thus, we use a rather simple technique for probabilistic forecast, namely the Linear Quantile Regression (Koenker and Bassett, 1978) as explained in the following Subsection 3.1, which allows us to focus on the performance comparison when such variables are included to a simple baseline model that only use past ground measurements. We only use information that is known at time (t) for this estimation, for instance, no other source of forecast were included as an input, such as Numerical Weather Predict (NWP) or satellite Cloud Motion Field (CMF) predictions. We assess a 3-hours ahead probabilistic forecast in a non-parametric way without any a priori assumption of the CDF shape and using a 10-minutes granularity in the input data and forecast output. All the models, described in Subsection 3.2, produce the set of quantiles q_τ using the LQR method. Subsection 3.3 presents the baseline persistence ensemble that is commonly used to provide a reference for the probabilistic forecasts performance, in the same manner as regular persistence is commonly used to provide a reference for the deterministic forecast.

3.1. The Linear Quantile Regression (LQR) method

This method estimates the quantiles of the cumulative distribution function of some response variable Y (called predictand) by assuming a linear relationship between the quantiles of Y (q_τ) and a set of explanatory variables X (called predictors):

$$q_\tau = \beta_\tau \; X + \epsilon, \tag{4}$$

where β_τ is a vector of parameters to optimize for each probability level τ and ϵ represents a random error term. The explanatory variables are the columns of the matrix X, and correspond to the known information at time (t) which will be related to the specific problem. For instance, in this work, past ground measurements, local past short-term variability and current time satellite cloudiness information can take part into the X matrix. The models, defined by their matrix X, are presented in Subsection 3.2. The response variable Y, whose quantiles have to be estimated, is the future clear sky index at time horizon Δt.

Following (Koenker and Bassett, 1978), and based on a training set of N samples, the quantiles $q_\tau = F^{-1}(\tau)$ can be estimated as the solution of the following optimization problem,

$$\hat{q}_\tau = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{N} \Psi_\tau\ (Y_i - q_\tau), \tag{5}$$

where $\Psi_\tau(u)$ is the quantile loss function defined as,

$$\Psi_\tau(u) = \begin{cases} u \times \tau & \text{if } u \geq 0, \\ u \times (\tau - 1) & \text{if } u < 0, \end{cases} \tag{6}$$

with τ representing the quantile probability level. Hence, in quantile regression, the quantiles are estimated by applying asymmetric weights to the mean absolute error. Using Eq. (4), the optimization problem can be translated to a set of regression parameters (β_τ) as,

$$\hat{\beta}_\tau = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{N} \Psi_\tau\ (Y_i - \beta X_i). \tag{7}$$

Thus, the quantity $\hat{q}_\tau = \hat{\beta}_\tau X$ is the estimation of the τ^{th} quantile obtained by the LQR method. The pairs of observed predictands and the set of predictors (Y_i and X_i) for the adjustment of $\hat{\beta}_\tau$ are taken from the training set. It must be noted that the quantile regression method estimates each quantile separately (i.e. the minimization of the quantile loss function is made for each τ separately). As a consequence, one can obtain quantile regression curves that may intersect, i.e. $\hat{q}_{\tau 1} > \hat{q}_{\tau 2}$ when $\tau_1 < \tau_2$. To avoid this issue, we used the rearrangement method described by (Chernozhukov et al., 2010).

3.2. Implemented regression models

We propose to analyze three different regression models based on the LQR method. Each model produces a probabilistic forecast by its own and is based on its own set of input variables. The first model is used in this work as a performance reference level and it follows Lauret et al. (2017) proposal to produce an intra-day probabilistic forecast using only actual and past ground measurements. It was found that, in addition to the actual measurement, five past input measurements were relevant to produce deterministic forecasts (Lauret et al., 2015). Thus, this first model (L5) is expressed in terms of its X matrix for the LQR regression method as,

$$\text{L5} : X = [\ k_c(t),\ k_c(t-1),\ \ldots,\ k_c(t-l_x)\ ]\,, \tag{8}$$

where $l_x = 5$ is the amount of lags used for the model.

The second model (L5-V) takes the short term variability defined in Eq. (2) as an additional variable:

$$\text{L5-V} : X = \left[\ k_c(t), \ k_c(t-1), \ \ldots, \ k_c(t-l_x), \ \sigma_c^{l_v}(t) \ \right]. \tag{9}$$

The third model (L5-S) takes the current time satellite albedo $\rho_p(t)$ averaged over a 10 min × 10 min latitude-longitude cell around the specific forecast site, as explained in Subsection 2.3. Its input matrix is expressed as:

$$\text{L5-S} : X = [\ k_c(t), \ k_c(t-1), \ \ldots, \ k_c(t-l_x), \ \rho_p(t) \ ]. \tag{10}$$

Finally, the last proposed model (L5-VS) takes all the input variables:

$$\text{L5-VS} : X = \left[\ k_c(t), \ k_c(t-1), \ \ldots, \ k_c(t-l_x), \ \sigma_c^{l_v}(t), \ \rho_p(t) \ \right]. \tag{11}$$

It has to be noted that depending on the implementation of the LQR algorithm, the independent term for the regression shall be or not explicitly included into the X matrices. All the models explored in this work includes the independent term, i.e. a column of ones, as they exhibit better results in terms of performance.

3.3. Persistence ensemble

The persistence is commonly used to define a performance reference for the forecast, as it is the simplest way to establish a prediction. For probabilistic forecast, a simple baseline persistence procedure can be adapted from the classical deterministic persistence by using the past ground measurements, namely, the persistence ensemble (Alessandrini et al., 2015; Chu and Coimbra, 2017; David et al., 2016). The persistence ensemble (PeEn) considered here is the nine GHI lagged measurements that precede the forecasting issuing time (including time (t)). The nine measurements are ordered to define the quantile values for the irradiance forecast, and thus, defining a persistence predicted CDF. Authors usually differ in how many past measurements or which methodology to use to define the PeEn. There are different proposals to define the PeEn, for instance the recent CH-PeEn (Yang, 2019) which uses historical data (not only past ground measurements) to define the predicted CDF for each time of the day. The choice of one or other methodology can affect the evaluation, in particular, the Skill Score (CRPSS) defined in Subsection 4.3. In this work we want to assess the impact of including the new predictors to the baseline L5 model, hence, we use the simplest nine past measurements to define nine predicted quantiles as it will not affect the comparison.

4. Performance metrics

Probabilistic forecast cannot be evaluated by using the standard deterministic performance metrics, such as the Mean Bias Error (MBE), Root Mean Squared Error (RMSE) and Forecasting Skill (FS), among other proposals. The computation of these quantities requires to compare two deterministic samples: a deterministic forecast vs a deterministic ground truth. As stated by (Wilks, 2009, Subsection 8.2.5) the conversion of a probabilistic forecast into a non-probabilistic (deterministic) forecast by any means (i.e. by taking the median or mean value) is an information degradation process that is, in all cases, in detriment of the forecast users. Hence, as the forecast objective in this article is probabilistic, we will focus on its detailed evaluation, without considering any form of converting this probabilistic forecast into a deterministic one. To assess the performance of probabilistic forecast, i.e. how well the CDFs are predicted, dedicated metrics are used. In all cases, the assessment is done by comparing the predictions (in this case, the predicted CDFs) with the corresponding future observations. In the field of meteorology or atmospheric sciences Jolliffe and Stephenson (2012) and Wilks (2009) list several attributes related to the quality of probabilistic forecasts. Mainly, three important properties are required for a skillful probabilistic forecast system, known as reliability, sharpness and resolution. Here, we follow the evaluation framework originally proposed by Pinson et al. (2007) for wind forecast, and recently adapted by Lauret et al. (2019) for solar forecast. This framework consists in assessing the reliability attribute before evaluating the others. In other words, a system which provides a probabilistic forecast must be, primarily, reliable, since a lack of reliability would introduce a systematic bias in subsequent decision-making. Pinson et al. (2007) also proposed to evaluate the overall skill of the methods by calculating a scoring rule that permits to objectively rank the different competing methods, but this tool must be used after knowing that the method is reliable. The sharpness, in this framework, refers to the concentration of the predictive distribution. Thus, it does not depend on the observations and it is an indicator of the forecast itself. It must be noticed that assessment of sharpness alone is not sufficient and must be made in conjunction with the reliability analysis. To sum up, as stated by Gneiting et al. (2007), the goal is to produce sharp forecast while ensuring that these forecasts are reliable. Regarding the third property (resolution), it consists (in this framework) in evaluating the ability of the forecast system to generate predictive distributions with prediction intervals that vary in size, depending on the forecast conditions (Pinson et al., 2007). For instance, resolution can be evaluated by showing how the width of the predictive distributions vary with increasing forecast horizon. Also, the level of uncertainty may depend on other physical external conditions and in the case of solar irradiance may vary according to the Sun's position in the sky (see for the instance the work of Grantham et al. (2016)).

The requirement of reliability will be assessed with the help of reliability diagrams (see Subsection 4.1). As noted above, sharpness is related to the concentration of the predictive distributions and will be measured by the average width of the prediction intervals or the so-called PINAW metric (Khosravi et al., 2013), presented in Subsection 4.2. Finally, the CRPS (Continuous Rank Probability Score) (Hersbach, 2000), explained in Subsection 4.3, gives an evaluation of the global skill of the proposed probabilistic models and provides a tool to rank the models once the reliability feature is met.

4.1. Reliability Property

To assess the reliability property, we use the methodology defined in Pinson et al. (2010) that is specially designed for density forecasts of continuous variables. The reliability diagram is a graphical verification display used to verify the reliability component of a probabilistic forecast system. This type of diagram illustrates the observed probabilities against the nominal ones (i.e., the probability levels of the different quantiles). By doing so, deviations from perfect reliability (the diagonal) are immediately revealed in a visual manner. However, similarly to Pinson et al. (2010), and for ease results' communication, we adopt here the type of reliability diagrams that show the difference between the observed probabilities and the nominal ones, for each probability level. Subsection 5.1 is devoted to assess the reliability of the models based on the explained diagrams.

4.2. Sharpness

Sharp probabilistic forecasts must present prediction intervals that are shorter on average than other reliable methods, like the climatological forecast for instance. A metric to quantify the sharpness is the normalized average interval width, know as PINAW (Prediction Interval Normalized Averaged Width). A shorter PINAW indicates that the system can produce smaller prediction intervals for a given coverage level (when the system is reliable), thus providing more information to decision-makers. Its simplified computation is,

$$\mathrm{PINAW}(\Delta t, \alpha) = \frac{\sum_{i=1}^{N} \left(\hat{G}_{h,i}|_{1-\frac{\alpha}{2}}(t+\Delta t) - \hat{G}_{h,i}|_{\frac{\alpha}{2}}(t+\Delta t) \right)}{\sum_{i=1}^{N} G_{h,i}(t+\Delta t)}, \tag{12}$$

where Δt is the forecast horizon and α is the confidence level that corresponds to the nominal coverage rate $(1-\alpha) \times 100\%$. $\hat{G}_h|_{1-\frac{\alpha}{2}}(t+\Delta t)$ and $\hat{G}_h|_{\frac{\alpha}{2}}(t+\Delta t)$ are the time series of the superior and inferior predicted GHI quantiles for the confidence level α. As stated before, PINAW is assessed in this work for a confidence level $\alpha = 0.2$ which correspond to a 80% nominal coverage. $G_h(t+\Delta t)$ is the actual GHI measurement at the prediction time. Notice that the N value required for calculating both averages is canceled in Eq. (12).

4.3. CRPS

The CRPS quantifies the deviation between the cumulative distributions functions (CDF) of the predicted and observed data (Hersbach, 2000). Thus, it quantifies the deviation of the probabilistic forecast from the perfect forecast. The formulation of the CRPS is:

$$\mathrm{CRPS} = \frac{1}{N} \sum_{i=1}^{N} \int_{-\infty}^{+\infty} \left(\hat{F}_i(x) - F_{i,x_o}(x) \right)^2 dx, \tag{13}$$

where $\hat{F}(x)$ is the predictive CDF of the variable of interest x (the GHI in this case) and $F_{x_o}(x)$ is a CDF step function that jumps from zero to one at the point where x equals the future observation x_0

($F_{x_0}(x) = 1|_{\{x \geq x_0\}}$). The integrated squared difference between the two CDFs is averaged over the N pairs of forecast/observation data. A different value of CRPS is computed for each time horizon Δt.

The CRPS is negatively oriented (smaller values represent a better performance) and it rewards the concentration of probability around the step function located at the observed value. Thus, the CRPS penalizes the lack of sharpness of the predictive distributions, as well as biased forecasts. Note that the CRPS has the same dimension as the predicted variable (W/m^2 in this case). The CRPS constitutes an attractive summary to quantify the predictive performance as it can address reliability and sharpness simultaneously. Indeed, if the CRPS metric is evaluated with a deterministic step function forecast, it turns to be the classical MAE (Mean Absolute Error) deterministic metric (Hersbach, 2000).

Similarly to the Forecast Skill (FS) metric commonly used to assess the quality of deterministic forecasts, we also include the CRPSS (Continuous Rank Probability Skill Score). This Skill Score (in %) is given by,

$$\mathrm{CRPSS} = 100 \times \left(1 - \frac{\mathrm{CRPS_m}}{\mathrm{CRPS_o}}\right), \tag{14}$$

where $\mathrm{CRPS_o}$ stands for the persistence ensemble and $\mathrm{CRPS_m}$ stands for the model under evaluation (here, the LQR models). A negative value of CRPSS indicates that the probabilistic method fails to outperform the persistence ensemble, while a positive value of CRPSS means that the forecasting method improves it. Further, the higher the CRPSS, the better the improvement.

5. Results and discussion

In this section, and for ease communication, we shall present the results for four selected sites from Table 1, namely, Salto (LE), Desert Rock (DR), Oahu (OA) and São Martinho da Serra (MS) . The four sites are representative of various climatic conditions. These stations comprise one low solar irradiance variability site (DR), one high variability site (OA) and two intermediate variability sites (LE and MS). Also, LE and MS are the only sites where we have satellite information. The full results for all the sites can be found in the Appendix A.

Following the evaluation framework, we will first evaluate the reliability property by analyzing the reliability diagrams. Then, the sharpness property will be assessed with the PINAW metric. Finally, we will evaluate the overall skill of the probabilistic models with the CRPS and CRPSS, being the latter more suitable for results' comparison in different sites than the former. The CRPS will be normalized and expressed as a percentage by dividing its absolute value (in W/m^2) with the GHI measurements average for the considered test period.

5.1. Reliability assessment

The reliability diagrams provide a visual inspection between estimated and nominal proportions. One form of visualization is to plot the deviations, i.e. the difference between both proportions, as a function of the nominal proportion. In such case, the perfect reliability corresponds to a zero line plot. Fig. 1 shows the deviations in the quantile forecasts for the four selected sites. The quantitative values of the plots are given in the Appendix A in Table A.2 and Table A.3, including the other sites.

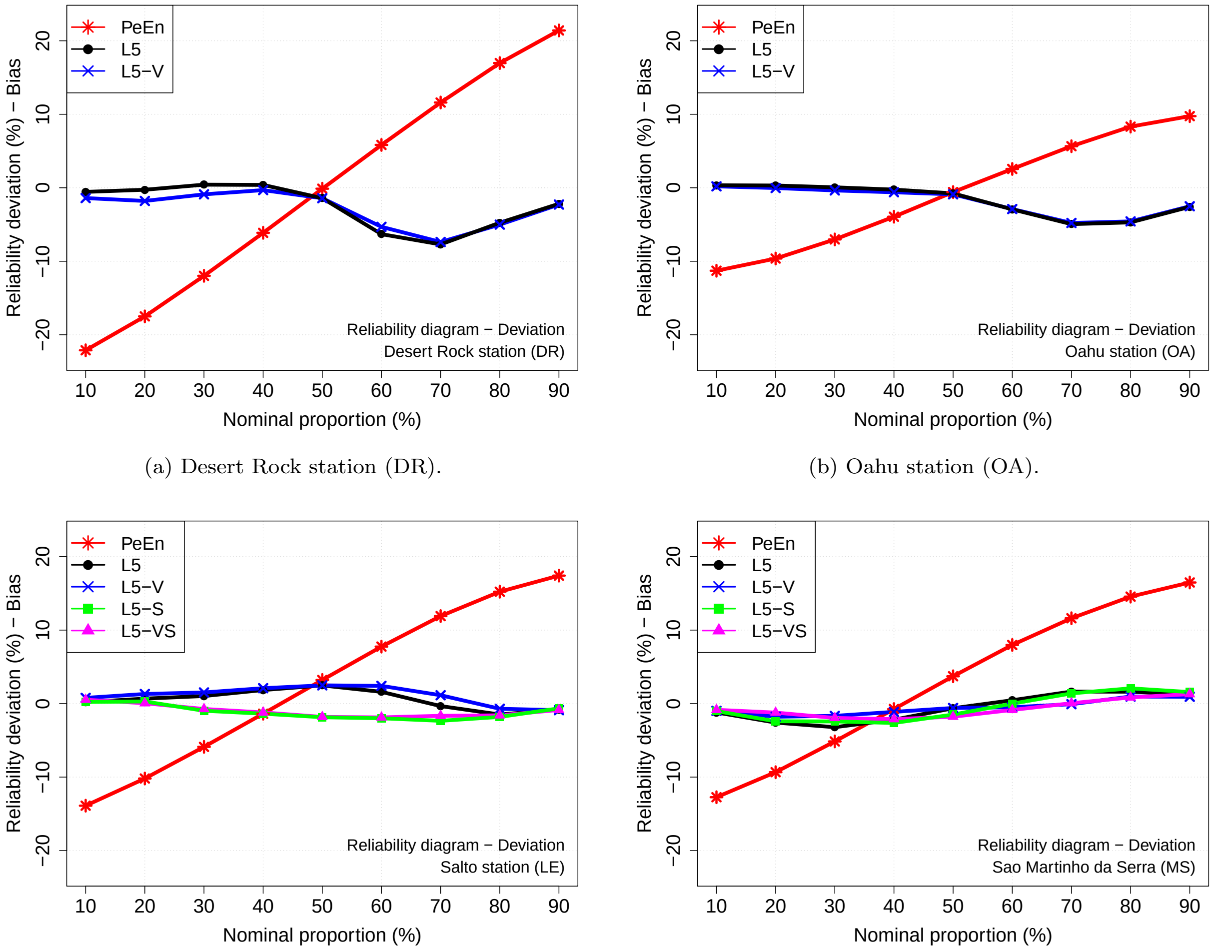
From the reliability diagrams, one can clearly argue that the Persistence (PeEn) model is not reliable. Indeed, the pattern exhibited by the PeEn model (that overforecast quantiles for nominal proportion above $\simeq 50\%$ and underforecast quantiles for nominal proportion below $\simeq 50\%$) is typical of models that generate too narrow predictive distributions. As the persistence is not reliable, we should normally discard this model at this point of the evaluation framework.

Regarding the other models (L5, L5-V, L5-S and L5-VS), one may conclude that for the sites of Salto (LE) and São Martinho da Serra (MS) the reliability property is verified as the deviation from the ideal case is lower than 2%. Conversely, for the sites of Desert Rock (DR) and Oahu (OA), some deviations ($< 8\%$ for DR and $< 5\%$ for OA) are noted for proportions between 60% and 80%. This means that both L5 and L5-V models overestimate the corresponding quantiles. This behavior may be explained by small differences between the training year and the testing year. However, in this work, the goal is to assess the added-value brought by additional predictors such as short-term variability or cloudiness satellite information in the quality of the probabilistic forecasts. As shown by Fig. 1, it is clear that short-term variability and satellite information do not modify significantly the reliability property.

5.2. Sharpness assessment

Fig. 2 plots the PINAW diagrams for the different sites and models, as a function of the time step. The quantitative values are given in Table A.4 and Table A.5 (see Appendix A), including the other sites. The PINAW metric that is shown here correspond to the 80% coverage rate between the 10% and 90% quantiles. As expected, PINAW increases with increasing forecast horizon. One may also notice that prediction intervals are wider for the high variability OA site than for the low variability DR site (see the different y axis scale in Fig. 2). Also, the slope along the first time steps (i.e. the first 6 time steps, corresponding to 1 hour) is higher for the OA site than for DR site. LE and MS, that represent intermediate solar variability conditions, are in a middle category in terms of PINAW: its slope behavior is more similar to that of the DR site but the span is more similar to that of OA site. The PINAW values for the last time steps in LE and MS are more similar to that of OA (in comparison to DR), but it is different for the first time steps, being significantly lower for these sites than for OA. Fig. 2 also shows a known behavior of the PeEn model: it leads to low PINAW values but at the expense of unreliability (Pedro et al., 2018). As stated in the previous subsection, this model is already discarded from the evaluation framework for not being reliable, and it is shown in Fig. 2 only for the sake of completeness.

The sharpness evaluation indicates that short term variability reduces the interval forecasts (blue line) albeit the picture is less clear for the Oahu site. In fact, in Subsection 5.4 it is shown that short-term



(a) Desert Rock station (DR).

(b) Oahu station (OA).

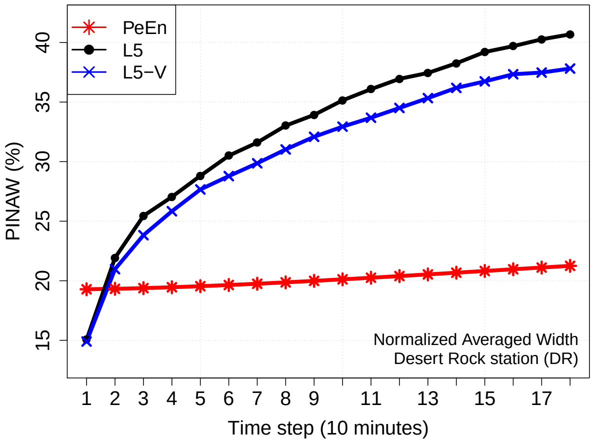
(c) Salto station (LE).

(d) São Martinho da Serra station (MS).

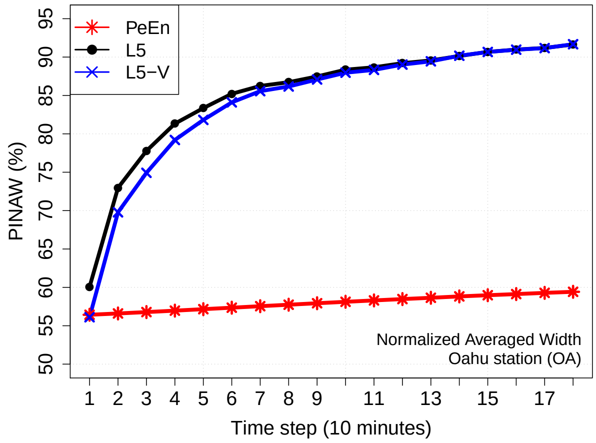
Figure 1: Reliability deviations of the inspected models in the four selected sites.

variability is a valuable predictor for probabilistic forecast in medium and low variability sites, as it gives an indication of the actual sky instability. In particular, it indicates if the previous samples correspond or not to clear sky conditions. This information can be taken by the probabilistic forecast methods to provide narrower prediction intervals if clear sky conditions are present and more wider if not.

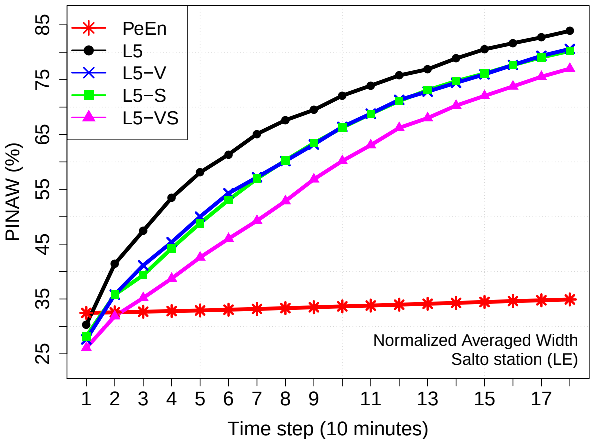
The spatially averaged satellite albedo tested for the two intermediate variability sites, LE and MS, also shows an improvement in the PINAW metric (see Figs. 2c and 2d). The models that only include satellite information obtain a PINAW trend similar to that of models that only include short-term variability. For the MS site, including only satellite information is slightly better (in terms of PINAW) than including only short-term variability. In both sites, the best model, in terms of PINAW, is the one which include both predictors.



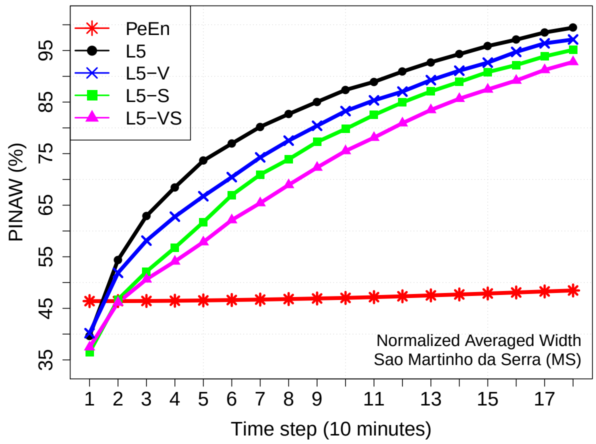
(a) Desert Rock station (DR).



(b) Oahu station (OA).



(c) Salto station (LE).



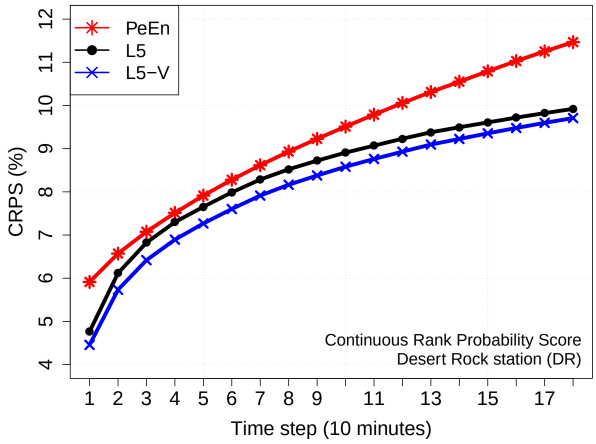
(d) São Martinho da Serra station (MS).

Figure 2: PINAW metric of the inspected models in the different solar variability sites.

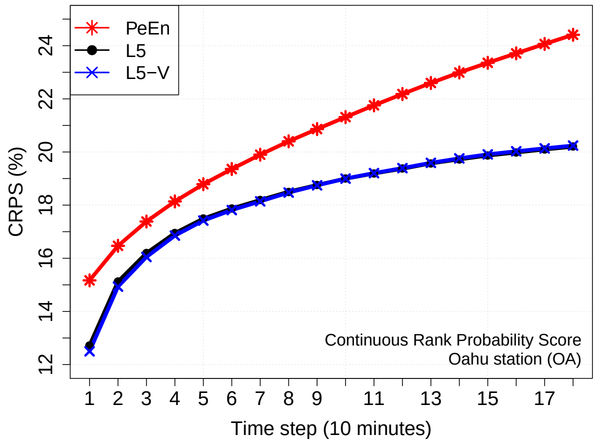
5.3. Overall skill with the CRPS

Fig. 3 and Fig. 4 show the CRPS and CRPSS for all models and the four selected sites, respectively, as an overall performance evaluation. As expected, the performance of the probabilistic models downgrades with increasing lead time (see Fig. 3); the lower the CRPS the better the performance is. Expressed in relative terms to the PeEn model, the CRPSS does not exhibits a monotonic trend (see Fig. 4). In all cases, CRPSS is high for the first two lead times and for large lead times, and have a minimum in the middle, showing a 'U' type shape. This means that the gain with respect to the PeEn model is the lowest at some lead time in the middle of the extremes. In all cases, this minimal gain is positive and above 2.7% (L5 model in LE site), which means that all models outperform the PeEn model. The quantitative values for the curves contained in Fig. 3 and Fig. 4 are given in the Appendix A from Table A.6 to Table A.9, including the other sites.

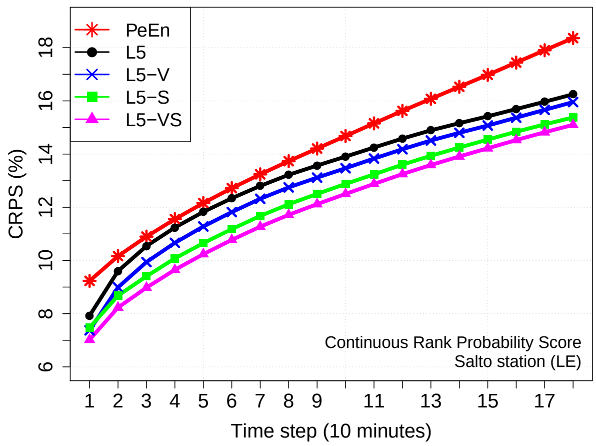
To better visualize the gain of including the exogenous variables in the LQR model, Fig. 5 provides the CRPSS difference between the L5-V, L5-S and L5-VS models with respect to the L5 model. This allows to isolate the improvement of adding the different variables to the baseline lagged model. These curves show a trend where the maximum gains are observed for the shorter lead times, peaking at the 2-3 time horizons. It shall be noticed that the lead times where the impact of adding the extra variables is the highest, is also roughly the same lead times where the L5 baseline model shows the less improvement with respect to the PeEn model. Hence, the effect of adding the extra variables not only improves the overall performance of the probabilistic forecast for these lead times, but also it moves the CRPS minimum towards higher lead times (see Fig. 4). This effect is clearly observable for the DR, LE and MS sites, and to a minor extent for the OA site.



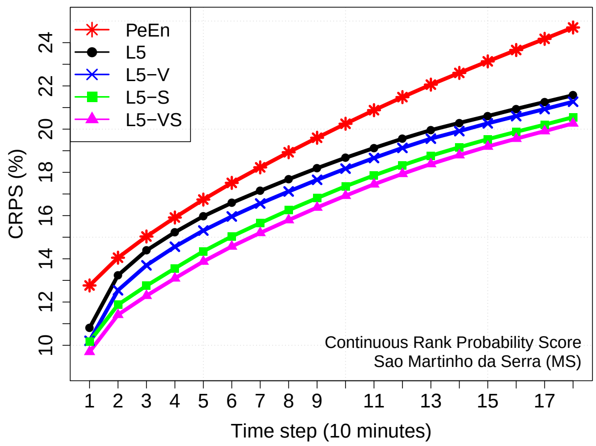
(a) Desert Rock station (DR).



(b) Oahu station (OA).

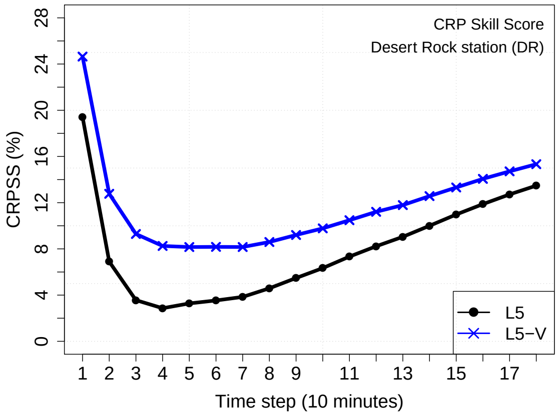


(c) Salto station (LE).

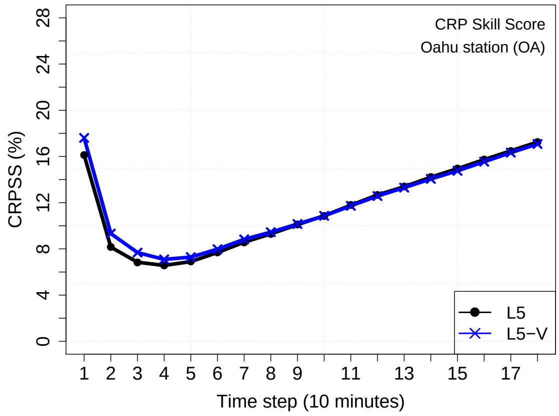


(d) São Martinho da Serra station (MS).

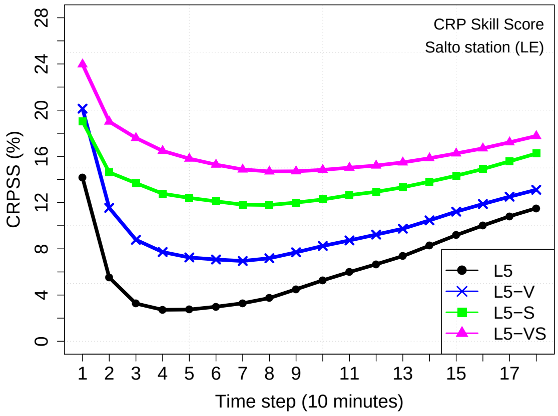
Figure 3: CRPS of the inspected models in the different solar variability sites.



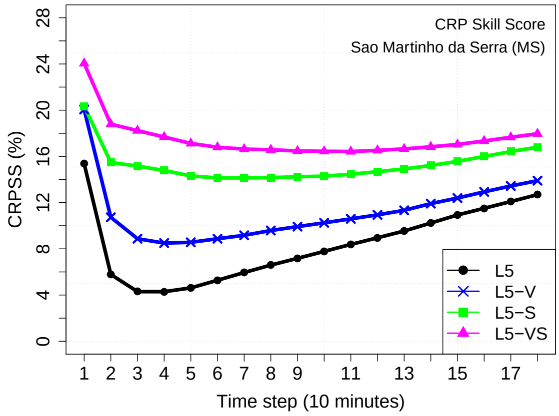
(a) Desert Rock station (DR).



(b) Oahu station (OA).

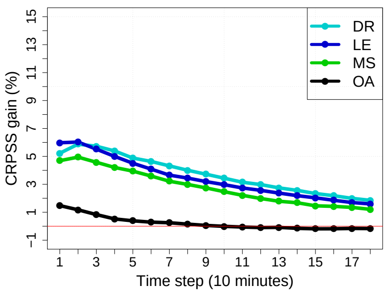


(c) Salto station (LE).

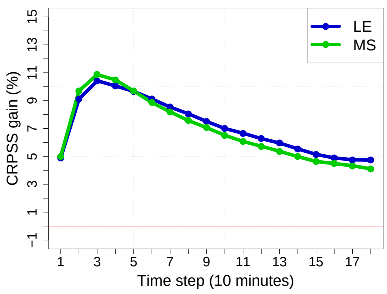


(d) São Martinho da Serra station (MS).

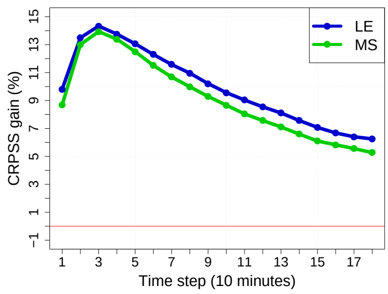
Figure 4: CRPSS of the inspected models in the different solar variability sites.



(a) L5-V (only variability).



(b) L5-S (only satellite).



(c) L5-VS (variability and satellite).

Figure 5: CRPSS gain when including exogenous variables to the baseline L5 LQR probabilistic model. The three figures have the same y axis scale for easy comparison.

The inclusion of short-term variability as a predictor improves the overall performance in the low variability site of DR and in the two intermediate variability sites of LE and MS. Inspecting Fig. 5a, the CRPSS improvement with respect to the baseline L5 model is similar for these three sites, peaking at $\simeq$ 5-6% for the shorter lead times (from 1 to 3 lead times) and decreasing asymptotically to a 1-2%. However, adding short-term variability has almost no effect for the high variability site of OA: there is a little positive effect for the first four lead times and no effect for longer time horizons (see Fig. 4b and Fig. 5a). The CRPSS improvement for the shorter lead times is around 1%, but is negligible for longer ones and even negative.

The inclusion of the averaged satellite albedo tested here in the LE and MS sites outperforms the inclusion of the short-term variability, as it can be seen in Fig. 3, and more clearly in Fig. 4 and by comparison of Figs. 5a and 5b. The trend is similar in both stations. In Fig. 5b, the CRPSS gain when adding only the satellite information raises to a peak value of $\simeq$ 10-11% at $\Delta t = 30$ minutes. The improvement is about 5% both for the first lead time and the largest ones, reaching the latter in a gradual manner. Based on this, it is clear that the spatially averaged satellite albedo is a useful input variable for probabilistic solar forecast in sites that exhibit an intermediate variability.

The best results are obtained by including both, short-term variability and satellite albedo, as shown in Fig. 3, Fig. 4 and Fig. 5c. The CRPSS gain reaches $\simeq$ 14%, again, at $\Delta t = 30$ minutes, but is increased for the first lead time (9-10% gain) in comparison with the last lead time (6-7% gain), as can be seen in Fig. 5c. In this case, a slightly better improvement is observed for LE (see Fig. 5c). This is consistent with the improvement observed for the L5-V model (where only short-term variability is added), for which the CRPSS improvement is also slightly better for the LE site than for the MS site (see Fig. 5a). To sum up, satellite albedo has a similar impact in both sites, but short-term variability is slightly more useful in the LE site, and the plots for the L5-V, L5-S and L5-VS are consistent with each other. Further, as it can be identified a peak gain of 5-6% for the L5-V model, of $\simeq$ 10% for the L5-S model and of $\simeq$ 14% for the L5-VS model, the performance of the 'complete' L5-VS model can be understood as roughly adding the performance gain of each variable. This is also approximately verified for the first and largest lead times.

5.4. Analysis by sky condition

In this section, an analysis is presented on whether it is better to include short-term variability or satellite albedo based on the actual sky conditions. Visual inspection of the 80% prediction intervals in the solar time series shows that the inclusion of short-term variability tends to tight the intervals under clear sky condition while expanding them under cloudy conditions. This behavior does not hold under all solar variability sites. On the other hand, the inclusion of satellite albedo improves the prediction intervals under cloudy condition, while not having a clear effect under clear sky condition. Therefore, this section is intended to quantify the impact of each variable in the probabilistic forecast performance for clear and cloudy skies.

For this analysis, it is required to classify the samples into clear sky and non clear sky. This was done by using the clear sky index and the short-term variability index, identifying the region in the two-dimensional plot in where the clear sky samples lay. The region was defined tight enough to ensure that only clear sky samples were selected for the clear sky data set. Visual inspection of the time-series was carried out to check that no cloudy-contaminated samples were included in the clear sky data set. Thus, the non clear sky data set (samples which are not in the clear sky data set) corresponds to partly cloudy and overcast sky conditions.

In the following analysis, the PINAW metric and the CRPSS are shown, as they are the best indicators to illustrate the discussion. The CRPS is omitted for brevity: it contains the same information as the CRPSS, but the differences are more difficult to visualize. We emphasize that the models are already checked for reliability.

5.4.1. Analysis for L5-V

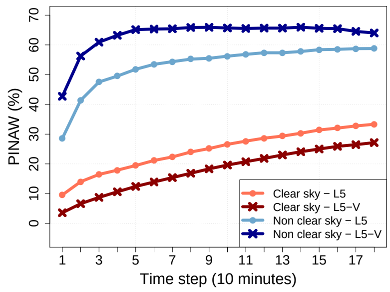
Fig. 6 compares the performance of the L5 and L5-V models discriminated by sky condition. For each station (each row of Fig. 6 plots), the PINAW metric is shown in the left panel for both sky conditions and the CRPSS is shown in the two other panels (center and right) for each sky condition. The inclusion of short-term variability reduces the average width of clear sky prediction intervals as it can be seen in the 'red' curves of the left panels. The improvement is important for the low variability site (DR) and the two intermediate variability sites (LE and MS). In the high variability site (OA), where clear sky samples are the minority in the time series, there is also a reduction in the clear sky intervals average width, but it is concentrated in the first time horizons (up to $\simeq 6$ lead times). On the other hand, the effect of including short-term variability in the LQR model is not as straightforward for cloudy sky condition (see the 'blue' curves in the left panels). In the high variability site it has negligible effect while in the low variability site it increases the average interval width. This is not necessary bad, as larger intervals may account for a better characterization of the probability density under cloudy skies. The intermediate variability sites show a mixed behavior between these two regimes: interval average width increases in the shorter lead times (up to $\simeq 30$ minutes ahead) but after this point the effect is rather negligible. The previous considerations are confirmed by the CRPSS. The inclusion of short-term variability improves the overall performance in both sky conditions and almost all sites, with the only exception of cloudy conditions for the high variability site of OA (see center and right panels of Fig. 6). Hence, short-term variability is a useful input variable for linear probabilistic forecast in low and intermediate variability sites, reducing the prediction intervals' average width under clear sky condition and increasing them under cloudy condition, but accounting for a performance gain in both sky conditions. However, its usefulness for high variability sites is restricted only to clear sky samples, which are rare in these data sets, and thus accounting for a small or negligible overall improvement throughout the whole time series. It is also interesting to note that the inclusion of short-term variability in OA, LE and MS turns negative values of CRPSS into positive ones for clear sky conditions (see center panels), therefore the variable has a significant impact in these cases (even for the high variability site of OA).

5.4.2. *Analysis for L5-S*

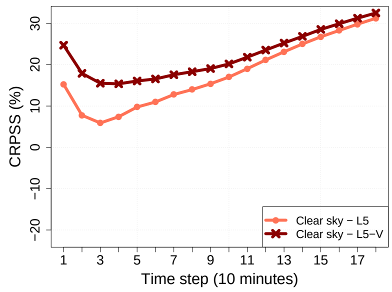
The performance comparison between the L5 baseline model and the L5-S model is provided in Fig. 7. The inclusion of spatially averaged satellite albedo is only tested here for intermediate variability sites (LE and MS, one in each row of Fig. 7). In these sites, it significantly improves the probabilistic forecast performance under cloudy sky condition, both reducing the average interval width (see Figs. 7a and 7d) and obtaining a better probability distribution prediction (see Figs. 7c and 7f). It also succeeds to slightly decrease the average interval width for clear sky condition (see the 'red' curves in the left panels of Fig. 7). However, the effect in the probability distribution prediction under clear sky condition depends on the lead time. For this condition, including satellite information downgrades the CDF prediction for the shorter lead times (up to 6-7 time horizons ahead) while improving the prediction for the longer ones (center panels). The downgrade observed for the shorter time horizons may be explained by surrounding sparse clouds that are spatially averaged but does not 'move' in the near future into the middle of the cell, hence, affecting the satellite average albedo but not the solar irradiance at the specific point. It can be summarized that, for intermediate variability sites, adding spatially averaged satellite albedo reports an improvement for probabilistic solar forecast under cloudy conditions but its impact is mixed under clear sky condition, being positive at the longer lead times and negative at the shorter lead times.

5.4.3. *Analysis for L5-VS*

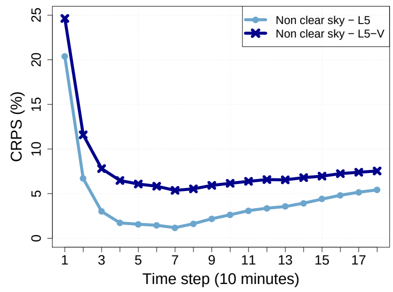
In the 'complete' L5-VS model, the inclusion of both variables compensates the drawbacks of each other. The performance comparison with the L5 baseline model is presented in Fig. 8. As it can be seen in the center and right panels, the L5-VS model outperforms the CDF prediction of the L5 model for both sky conditions and all the lead times. At the same time, it succeeds to reduce most of the average interval widths, with the only exception of the first and second time horizons under cloudy condition. The mixed behavior observed under clear sky when only adding the satellite input is upgraded by the inclusion of the short-term variability. The small effect observed under cloudy condition when only adding short-term variability is upgraded by the inclusion of the satellite information. The slight PINAW increase at the shorter time horizons for cloudy condition due to short-term variability is partially compensated when adding satellite information. If Figs. 6 to 8 are compared for the intermediate variability sites (LE and MS), again, the performance of the L5-VS model can be roughly explained as the added value of each variable separately.



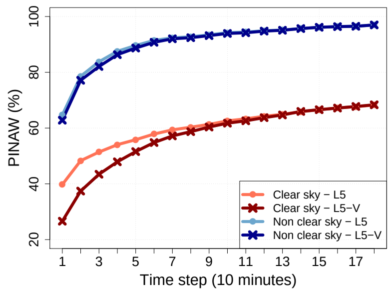
(a) DR site: PINAW metric.



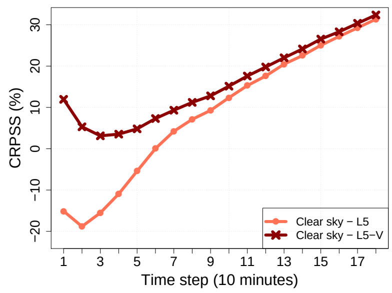
(b) DR site: clear sky CRPSS.



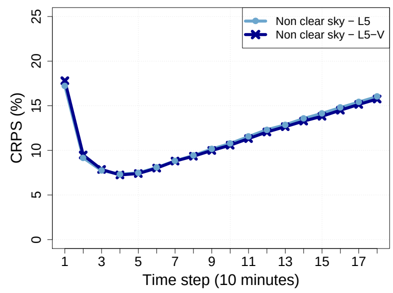
(c) DR site: non clear sky CRPSS.



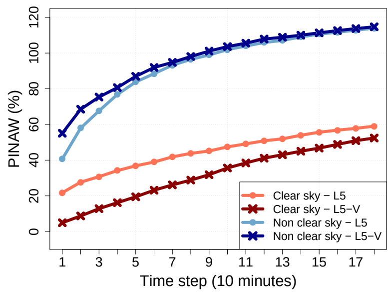
(d) OA site: PINAW metric.



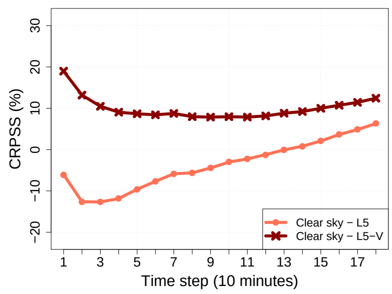
(e) OA site: clear sky CRPSS.



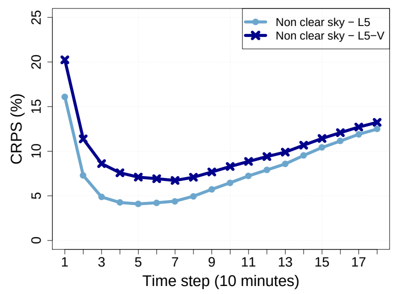
(f) OA site: non clear sky CRPSS.



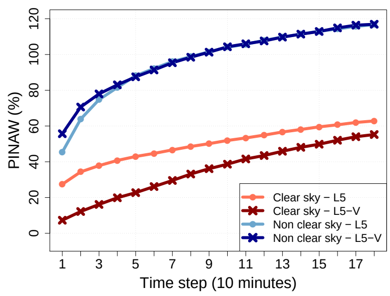
(g) LE site: PINAW metric.



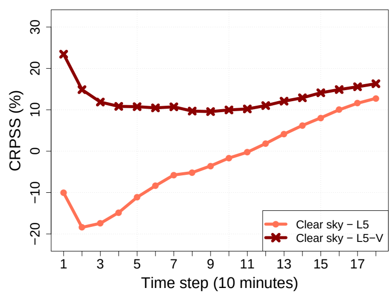
(h) LE site: clear sky CRPSS.



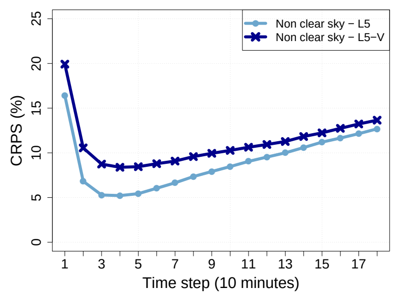
(i) LE site: non clear sky CRPSS.



(j) MS site: PINAW metric.

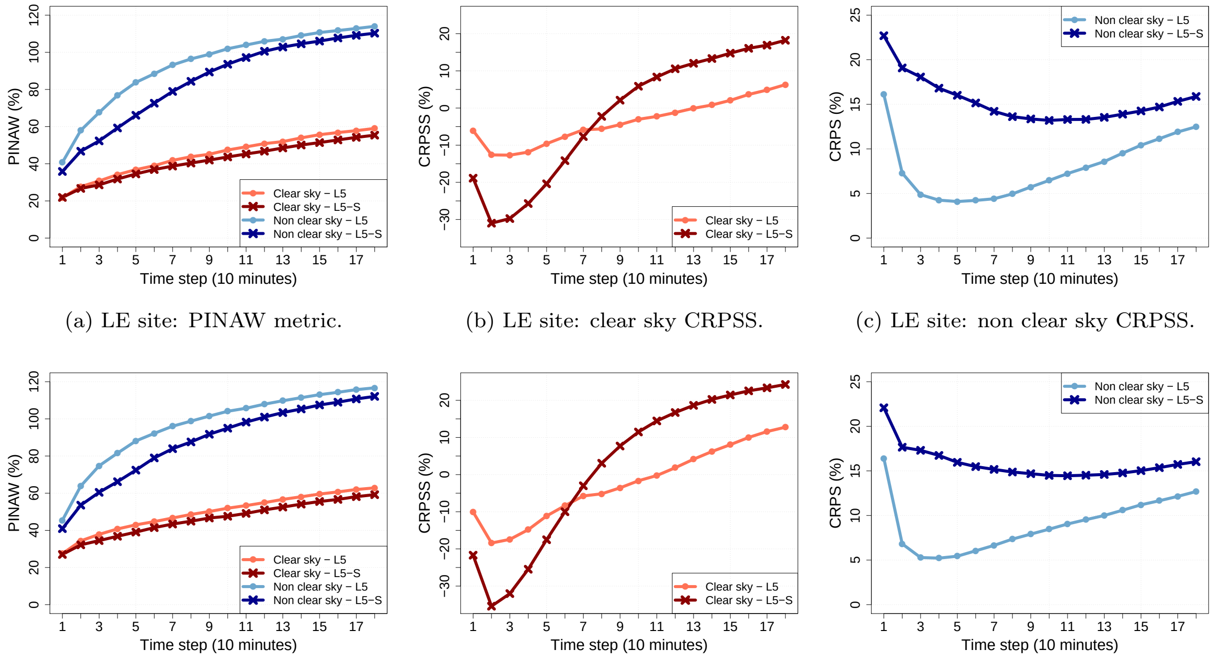


(k) MS site: clear sky CRPSS.



(l) MS site: non clear sky CRPSS.

Figure 6: Performance assessment for the L5-V model discriminated by sky condition. The figures show the impact of including short-term variability to the baseline L5 model for the four selected sites.



(a) LE site: PINAW metric. (b) LE site: clear sky CRPSS. (c) LE site: non clear sky CRPSS.

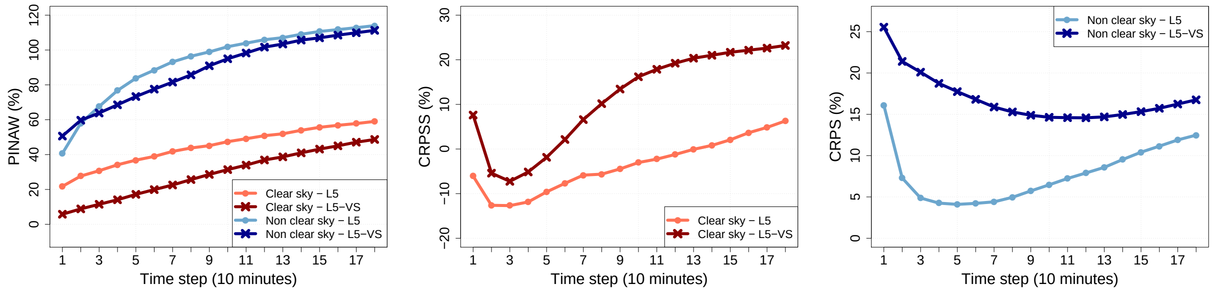
(d) MS site: PINAW metric. (e) MS site: clear sky CRPSS. (f) MS site: non clear sky CRPSS.

Figure 7: Performance assessment for the L5-S model discriminated by sky condition. The figures show the impact of including spatially averaged satellite albedo to the baseline L5 model for the two intermediate solar variability sites under study.

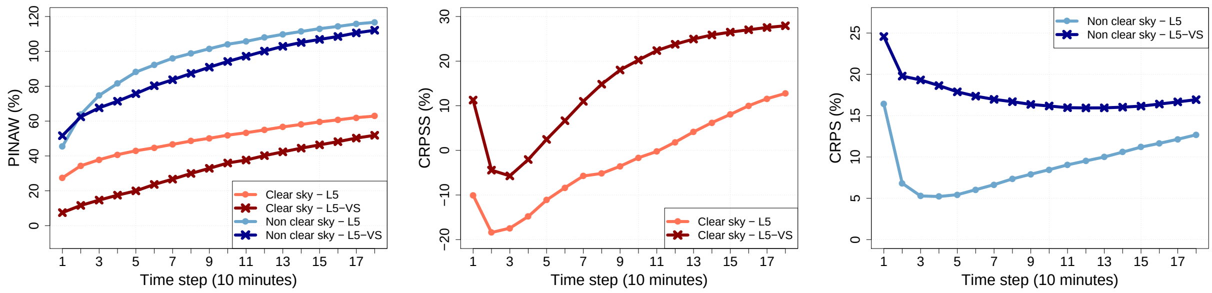
6. Conclusions

In this work, we quantified the impact of adding two exogenous variables to intra-day linear solar probabilistic forecast: (a) local short-term solar irradiance variability and (b) spatially averaged satellite albedo. Evidence that these information improve the forecast in both cases is provided. The Linear Quantile Regression method was used with lagged past measurements and the extra variables to forecast the global horizontal irradiance up to 3 hours ahead with a 10 minutes granularity. Short-term variability is tested in 7 sites, accounting for different climates that exhibit high, intermediate and low solar variability conditions. Satellite albedo is tested in two intermediate variability sites where we had satellite images availability. A detailed performance assessment was provided and the added value of each variable was quantified, both separately and jointly. Furthermore, an analysis depending on the sky condition was presented, allowing us to visualize the impact of each variable in the probabilistic forecast at clear and non clear sky conditions.

The models that include these new extra variables (L5-V, L5-S and L5-VS) were compared to a baseline L5 model which only uses five past measurements. All the models, from L5 to L5-VS, met the reliability property and outperformed the Persistence Ensemble (PeEn) in predicting the future CDFs for all lead times (lower CRPS and positive CRPSS). It was shown that the inclusion of each variable improves the overall (throughout all the time-series) probabilistic forecast performance of the L5 model, but the impact is not



(a) LE site: PINAW metric. (b) LE site: clear sky CRPSS. (c) LE site: non clear sky CRPSS.



(d) MS site: PINAW metric. (e) MS site: clear sky CRPSS. (f) MS site: non clear sky CRPSS.

Figure 8: Performance assessment for the 'complete' L5-VS model discriminated by sky condition. The figures show the impact of including both inspected variables to the baseline L5 model for the two intermediate variability sites under study.

the same depending on the lead times, sky condition and solar variability of the site.

The inclusion of short-term variability improves the overall probabilistic forecast performance for all lead times in the low and intermediate variability sites. In high variability sites, the impact is small or negligible, and only restricted to the first five lead times (approximately) in where a slight improvement is found. This improvement is mostly related with narrower prediction intervals under clear sky condition, for which PINAW is reduced and CRPSS is increased. The general conclusion is that short-term variability improves the linear probabilistic forecast under clear sky condition, consistent with a negligible overall effect in high solar variability sites where clear sky samples are rare. Also, it improves the performance under cloudy condition, but in a minor extent than under clear sky and not for high variability sites.

The inclusion of the satellite albedo to the L5 baseline model (L5-S) showed a very important improvement in the CDFs prediction accuracy, outperforming the overall value of short-term variability for intermediate variability sites. There is also an improvement in the overall average interval width, but at a less extent. When discriminating the impact into clear and non clear skies, it is observed that the satellite predictor has an important effect on increasing the CDFs prediction accuracy under cloudy sky while at the same time reducing its average width. For clear sky the average width is also reduced, but slightly, and the effect on the CDFs accuracy depends on the time horizon. The drawbacks observed when adding each variable separately, namely, higher cloudy sky PINAW when adding short-term variability and lower clear sky CPRSS at shorter lead times when adding satellite information, are compensated. In fact, it is possible to roughly understand the PINAW and CRPSS behavior of the L5-VS model by adding the value of each variable with respect to the L5 model.

The results show that both variable are useful inputs for solar probabilistic predictions, helping to improve different features of the probabilistic forecast. However, this does not mean that they are useful predictors for any kind of forecast. For instance, in Marchesoni-Acland et al. (2019) the same variables are inspected as exogenous inputs for an optimal auto-regressive moving-average (ARMAX) solar deterministic forecast in the same forecast horizons and granularity of the present work. It was found that, at least for intermediate solar variability sites, short-term variability is not a useful variable. On the contrary, space-averaged satellite albedo does improve the deterministic forecast. This means that a variable that is useless for deterministic forecast, should not necessary be discarded for probabilistic forecast. Such is the case of the past short-term variability, which improves the probabilistic forecast mainly by reducing the average interval width under clear sky condition, but does not report an improvement in a deterministic forecast under a similar linear framework. The usefulness of input variables may also depend on the nature of each technique.

Solar variability is a predictor that can be easily derived from clear sky index data series and, as shown, can efficiently improve the quality of probabilistic forecast. Recent bibliography (Yang, 2020) suggests that, among accurate clear sky models, the choice of the clear sky model is rather arbitrary for forecasting techniques and hence the forecaster can choose the clear sky model of his convenience to reproduce the methodology. Satellite albedo requires geostationary satellite images availability, which are not trivially accessed. However, specialized industry companies which provide solar forecasting services usually have access to these satellite images for Cloud Motion Vector forecast. As we showed here, the overall impact of adding satellite information outperforms the one of short-term variability (at least in intermediate variability sites), so spatially averaged satellite albedo is a valuable and simple variable to include in solar probabilistic frameworks. The final answers on whether satellite averaged albedo improves the probabilistic forecast for high or low variability sites still remains. The final test of this hypothesis requires to access satellite images for other regions, which is part of our current work. The impact in the probabilistic forecast of the cell size or displacement from the specific site are also part of our current work. The methodology presented here may be extended to longer time horizons, such as 4-hours ahead or more, by adding NWP data.

Acknowledgments

The authors gratefully acknowledge the laboratory PIMENT from the University of La Réunion, the laboratory LARGE from the University of Guadeloupe, the National Renewable Energy Laboratory (NREL), and the NOAA's SURFRAD meteorological network for providing part of the ground measurements used in this work. The authors acknowledge the financial support given by Uruguay's National Research and Innovation Agency (ANII) under the FSE-ANII-2016-131799 grant.

A. Performance detailed evaluation

This Appendix provides the quantitative values of the performance assessment for all inspected sites. Tables A.2 and A.3 present the bias (reliability deviations) and Tables A.4 and A.5 the PINAW metric. Tables from Table A.6 to Table A.9 present the CRPS and CRPSS. This Appendix also presents the overall metrics' plots for the FP, FO and TM sites (see Fig. A.9), which were not included in the general analysis for the sake of simplicity. FP is an intermediate variability site and hence its results are similar to those of LE and MS. FO and TM, being insular sites, exhibit high variability and their results are similar to OA.

station	**LE**		**DR**		**FP**		**OA**		**FO**		**TM**		**MS**	
prob.	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**
10%	+0.2	+0.8	−0.6	−1.4	+0.6	+0.4	+0.3	+0.2	+0.3	+0.2	+0.1	−0.2	−1.2	−1.0
20%	+0.7	+1.3	−0.3	−1.8	+0.5	0.0	+0.3	0.0	+0.3	v0.0	−0.1	−0.4	−2.6	−1.8
30%	+1.0	+1.5	+0.4	−0.9	−0.6	−0.7	+0.1	−0.4	+0.1	−0.4	+0.1	−0.3	−3.2	−1.7
40%	+1.8	+2.1	+0.4	−0.3	−1.0	−0.7	−0.2	−0.6	−0.2	−0.6	+0.1	−0.1	−2.2	−1.1
50%	+2.5	+2.5	−1.4	−1.4	−0.9	−0.9	−0.8	−0.9	−0.8	−0.9	+0.3	+0.4	−0.6	−0.6
60%	+1.6	+2.4	−6.3	−5.3	−1.1	−1.4	−2.9	−2.9	−2.9	−2.9	+0.4	+0.4	+0.5	−0.5
70%	−0.3	+1.1	−7.7	−7.4	−0.6	−1.1	−4.9	−4.8	−4.9	−4.8	−0.1	−0.1	−0.1	−0.1
80%	−1.5	−0.7	−4.8	−5.0	−0.2	−0.7	−4.7	−4.6	−4.7	−4.6	−0.6	−0.5	+1.6	+1.0
90%	−0.8	−0.9	−2.2	−2.3	−0.2	−0.1	−2.6	−2.5	−2.6	−2.5	−1.2	−1.1	+1.2	+0.9

Table A.2: Bias of L5 and L5V models

station	**LE**		**MS**	
prob.	**L5-S**	**L5-VS**	**L5-S**	**L5-VS**
10%	+0.2	+0.5	−1.0	−0.8
20%	+0.3	+0.1	−2.4	−1.2
30%	−1.0	−0.7	−2.4	−2.0
40%	−1.4	−1.2	−2.6	−2.1
50%	−1.9	−1.8	−1.5	−1.8
60%	−2.0	−1.9	0.0	−0.9
70%	−2.3	−1.7	+1.4	+0.1
80%	−1.8	−1.6	+2.1	+0.8
90%	−0.6	−0.9	+1.6	+1.3

Table A.3: Bias of L5S and L5VS models.

station	LE		DR		FP		OA		FO		TM		MS	
$\mathbf{\Delta t_i}$	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**	**L5**	**L5-V**
10 min	30.3	27.6	15.1	14.9	34.8	31.9	60.1	56.1	58.7	57.1	62.5	59.2	39.6	40.2
20 min	41.4	35.8	21.9	21.0	48.3	43.2	73.0	69.8	74.1	72.5	83.5	79.3	54.4	51.9
30 min	47.4	41.1	25.4	23.8	55.2	49.7	77.8	74.9	81.2	80.0	92.8	90.0	62.9	58.1
40 min	53.4	45.4	27.0	25.8	60.5	54.7	81.3	79.2	85.0	84.2	97.3	94.8	68.5	62.8
50 min	58.1	50.0	28.8	27.7	63.3	58.7	83.4	81.8	87.4	87.3	100.8	98.5	73.6	66.7
60 min	61.4	54.3	30.5	28.8	66.6	62.3	85.2	84.1	89.3	89.4	102.5	100.7	77.0	70.4
70 min	65.1	57.2	31.6	29.9	69.0	65.5	86.2	85.6	91.1	90.7	104.0	103.0	80.2	74.3
80 min	67.6	60.2	33.0	31.0	70.7	68.0	86.7	86.2	92.7	92.6	105.5	104.5	82.7	77.5
90 min	69.5	63.2	33.9	32.1	72.8	70.3	87.5	87.1	94.1	94.0	106.8	105.9	85.0	80.4
100 min	72.1	66.4	35.1	32.9	74.7	72.2	88.4	88.0	95.0	94.9	108.0	107.1	87.4	83.3
110 min	73.9	68.8	36.1	33.7	76.4	73.6	88.7	88.3	96.2	95.7	109.1	108.8	88.9	85.3
120 min	75.8	71.3	37.0	34.5	77.3	74.9	89.2	89.0	97.0	96.5	110.3	109.6	90.9	87.0
130 min	76.9	72.8	37.4	35.3	78.1	76.0	89.6	89.4	97.9	97.7	110.8	110.5	92.7	89.3
140 min	78.9	74.4	38.3	36.2	79.2	77.5	90.1	90.2	98.4	98.1	111.5	111.1	94.3	91.1
150 min	80.5	76.0	39.2	36.7	80.1	78.5	90.7	90.7	98.8	98.6	112.3	111.9	95.9	92.7
160 min	81.7	77.7	39.7	37.3	81.1	79.3	91.0	91.0	99.5	99.6	112.7	112.6	97.1	94.7
170 min	82.7	79.3	40.3	37.5	82.1	80.8	91.2	91.2	99.8	99.8	112.8	112.5	98.5	96.4
180 min	83.9	80.6	40.7	37.8	82.9	81.9	91.6	91.7	99.8	99.9	113.2	113.1	99.4	97.1

Table A.4: PINAW of L5 and L5V models.

station	LE		MS		station	LE		MS	
$\mathbf{\Delta t_i}$	**L5-S**	**L5-VS**	**L5-S**	**L5-VS**	$\mathbf{\Delta t_i}$	**L5-S**	**L5-VS**	**L5-S**	**L5-VS**
10 min	28.2	26.1	36.5	37.5	100 min	66.3	60.2	79.8	75.5
20 min	35.8	31.9	46.7	46.1	110 min	68.8	63.1	82.5	78.1
30 min	39.3	35.2	52.1	50.6	120 min	71.1	66.2	84.9	80.9
40 min	44.2	38.7	56.8	54.1	130 min	73.1	68.0	87.1	83.5
50 min	48.8	42.6	61.7	57.8	140 min	74.7	70.3	88.9	85.7
60 min	53.0	46.0	67.0	62.1	150 min	76.1	72.0	90.8	87.5
70 min	56.9	49.3	70.9	65.4	160 min	77.7	73.8	92.2	89.2
80 min	60.3	52.9	73.9	68.9	170 min	79.1	75.5	93.9	91.2
90 min	63.4	56.8	77.3	72.3	180 min	80.2	77.0	95.1	92.8

Table A.5: PINAW of L5S and L5VS models.

station	LE		DR		FP		OA		FO		TM		MS	
mean	464 W/m^2		540 W/m^2		388 W/m^2		487 W/m^2		513 W/m^2		450 W/m^2		421 W/m^2	
Δt_i	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V
10 min	7.9	7.4	4.8	4.5	8.8	8.4	12.7	12.5	13.8	13.5	13.7	13.3	10.8	10.2
20 min	9.6	9.0	6.1	5.7	11.1	10.6	15.1	14.9	16.6	16.3	16.9	16.4	13.2	12.5
30 min	10.5	9.9	6.8	6.4	12.2	11.8	16.2	16.0	17.8	17.6	18.5	18.1	14.4	13.7
40 min	11.2	10.7	7.3	6.9	13.1	12.6	16.9	16.9	18.6	18.4	19.7	19.3	15.2	14.6
50 min	11.8	11.3	7.7	7.3	13.7	13.3	17.5	17.4	19.2	19.0	20.7	20.3	16.0	15.3
60 min	12.3	11.8	8.0	7.6	14.3	13.9	17.9	17.8	19.7	19.5	21.5	21.1	16.6	16.0
70 min	12.8	12.3	8.3	7.9	14.8	14.4	18.2	18.1	20.1	20.0	22.2	21.9	17.1	16.6
80 min	13.2	12.7	8.5	8.2	15.2	14.9	18.5	18.5	20.5	20.4	22.9	22.6	17.7	17.1
90 min	13.6	13.1	8.7	8.4	15.6	15.3	18.8	18.7	20.8	20.7	23.5	23.2	18.2	17.7
100 min	13.9	13.5	8.9	8.6	15.9	15.6	19.0	19.0	21.1	21.0	24.1	23.9	18.7	18.2
110 min	14.2	13.8	9.1	8.8	16.3	16.0	19.2	19.2	21.3	21.2	24.7	24.4	19.1	18.7
120 min	14.6	14.2	9.2	8.9	16.6	16.3	19.4	19.4	21.6	21.5	25.1	24.9	19.6	19.1
130 min	14.9	14.5	9.4	9.1	16.9	16.6	19.6	19.6	21.7	21.7	25.6	25.4	20.0	19.6
140 min	15.2	14.8	9.5	9.2	17.1	16.9	19.7	19.8	21.9	21.9	25.9	25.7	20.3	19.9
150 min	15.4	15.1	9.6	9.4	17.4	17.2	19.9	19.9	22.1	22.0	26.2	26.1	20.6	20.3
160 min	15.7	15.4	9.7	9.5	17.6	17.4	20.0	20.0	22.2	22.2	26.5	26.4	20.9	20.6
170 min	16.0	15.7	9.8	9.6	17.8	17.6	20.1	20.1	22.4	22.3	26.8	26.6	21.3	20.9
180 min	16.2	16.0	9.9	9.7	17.9	17.8	20.2	20.2	22.5	22.4	27.0	26.9	21.6	21.3

Table A.6: CRPS of L5 and L5V models

station	LE		MS		station	LE		MS	
mean	464 W/m^2		421 W/m^2		mean	464 W/m^2		421 W/m^2	
Δt_i	L5-S	L5-VS	L5-S	L5-VS	Δt_i	L5-S	L5-VS	L5-S	L5-VS
10 min	7.5	7.0	10.2	9.7	100 min	12.9	12.5	17.4	16.9
20 min	8.7	8.2	11.9	11.4	110 min	13.2	12.9	17.9	17.4
30 min	9.4	9.0	12.8	12.3	120 min	13.6	13.2	18.3	17.9
40 min	10.1	9.7	13.6	13.1	130 min	13.9	13.6	18.8	18.4
50 min	10.7	10.2	14.3	13.9	140 min	14.2	13.9	19.2	18.8
60 min	11.2	10.8	15.0	14.6	150 min	14.5	14.2	19.5	19.2
70 min	11.7	11.3	15.7	15.2	160 min	14.8	14.5	19.9	19.6
80 min	12.1	11.7	16.3	15.8	170 min	15.1	14.8	20.2	19.9
90 min	12.5	12.1	16.8	16.4	180 min	15.4	15.1	20.6	20.3

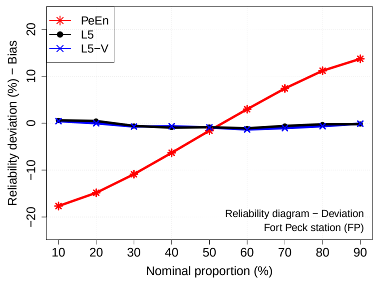
Table A.7: CRPS of L5S and L5VS models

station	LE		DR		FP		OA		FO		TM		MS	
Δt_i	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V	L5	L5-V
10 min	14.2	20.1	19.4	24.6	21.2	24.9	16.1	17.6	16.1	18.1	19.9	22.5	15.4	20.1
20 min	5.5	11.5	6.9	12.8	10.6	14.6	8.2	9.3	7.2	9.0	11.0	13.4	5.8	10.7
30 min	3.3	8.8	3.6	9.3	8.2	11.8	6.8	7.7	5.8	7.3	9.5	11.6	4.3	8.9
40 min	2.7	7.7	2.9	8.3	7.7	11.0	6.6	7.1	5.8	7.0	9.5	11.2	4.3	8.5
50 min	2.7	7.3	3.3	8.2	7.8	10.8	6.9	7.3	6.4	7.4	10.2	11.8	4.6	8.6
60 min	3.0	7.1	3.5	8.2	8.1	10.8	7.7	8.0	7.1	7.9	11.3	12.7	5.3	8.9
70 min	3.3	6.9	3.8	8.2	8.8	11.2	8.6	8.8	7.7	8.4	12.4	13.7	5.9	9.2
80 min	3.7	7.2	4.6	8.6	9.4	11.6	9.3	9.4	8.4	8.9	14.6	14.7	6.6	9.6
90 min	4.5	7.7	5.5	9.2	10.2	12.2	10.1	10.2	9.1	9.5	15.6	15.7	7.2	9.9
100 min	5.3	8.2	6.3	9.8	11.1	12.8	10.9	10.9	9.8	10.2	15.6	16.5	7.8	10.3
110 min	6.0	8.7	7.3	10.5	11.9	13.3	11.8	11.7	10.5	10.9	16.7	17.5	8.4	10.6
120 min	6.7	9.2	8.2	11.2	12.5	13.8	12.7	12.6	11.1	11.4	17.7	18.4	9.0	10.9
130 min	7.4	9.7	9.0	11.8	13.1	14.3	13.4	13.3	11.9	12.1	18.7	19.4	9.5	11.3
140 min	8.3	10.5	10.0	12.6	13.6	14.6	14.2	14.1	12.5	12.7	19.7	20.3	10.2	11.9
150 min	9.2	11.2	11.0	13.3	14.2	15.1	14.9	14.8	13.2	13.4	20.7	21.2	10.9	12.4
160 min	10.0	11.9	11.9	14.1	14.9	15.6	15.7	15.5	13.8	14.0	21.5	21.9	11.5	12.9
170 min	10.8	12.5	12.7	14.7	15.5	16.2	16.5	16.3	14.4	14.5	22.4	22.7	12.1	13.4
180 min	11.5	13.1	13.5	15.3	16.1	16.7	17.2	17.1	15.0	15.1	23.2	23.5	12.7	13.9

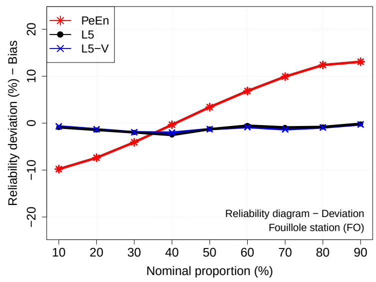
Table A.8: CRPSS of L5 and L5V models

station	LE		MS		station	LE		MS	
Δt_i	L5-S	L5-VS	L5-S	L5-VS	Δt_i	L5-S	L5-VS	L5-S	L5-VS
10 min	19.0	24.0	20.4	24.1	100 min	12.3	14.8	14.3	16.4
20 min	14.6	19.0	15.5	18.8	110 min	12.6	15.0	14.5	16.4
30 min	13.7	17.6	15.2	18.2	120 min	12.9	15.2	14.7	16.5
40 min	12.8	16.5	14.8	17.7	130 min	13.3	15.5	14.9	16.6
50 min	12.4	15.8	14.3	17.1	140 min	13.8	15.8	15.2	16.8
60 min	12.1	15.3	14.1	16.8	150 min	14.3	16.3	15.6	17.0
70 min	11.8	14.9	14.1	16.6	160 min	14.9	16.7	16.0	17.3
80 min	11.8	14.7	14.2	16.6	170 min	15.6	17.2	16.4	17.7
90 min	12.0	14.7	14.2	16.5	180 min	16.2	17.8	16.8	18.0

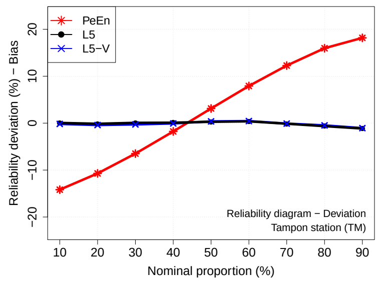
Table A.9: CRPSS of L5S and L5VS models



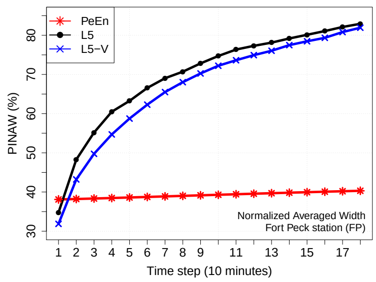
(a) FP site: reliability diagram.



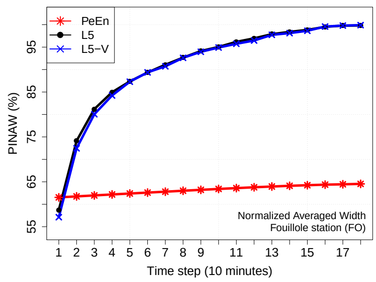
(b) FO site: reliability diagram.



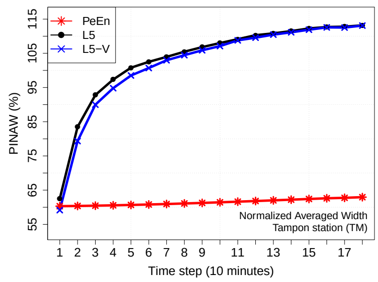
(c) TM site: reliability diagram.



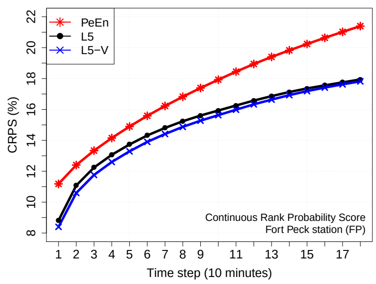
(d) FP site: PINAW metric.



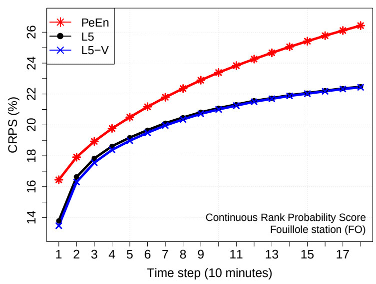
(e) FO site: PINAW metric.



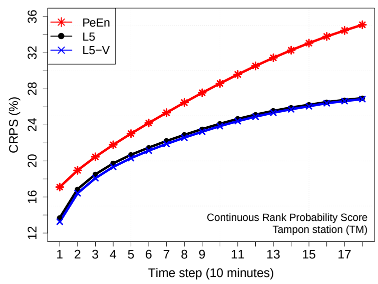
(f) TM site: PINAW metric.



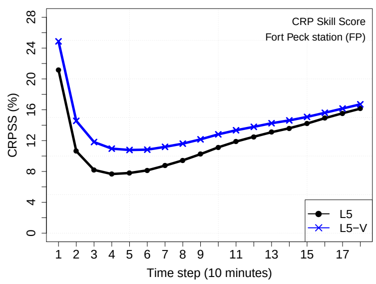
(g) FP site: CRPS.



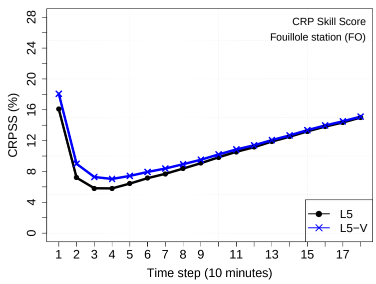
(h) FO site: CRPS.



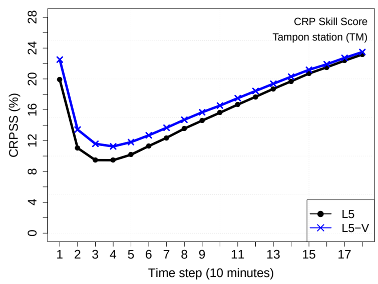
(i) TM site: CRPS.



(j) FP site: CRPSS.



(k) FO site: CRPSS.



(l) TM site: CRPSS.

Figure A.9: Performance metrics for the other sites (FP, FO and TM).

References

Aguiar, L. M., Pereira, B., David, M., Díaz, F., and Lauret, P. (2015). Use of satellite data to improve solar radiation forecasting with bayesian artificial neural networks. *Solar Energy*, 122:1309–1324.

Alessandrini, S., Monache, L. D., Sperati, S., and Cervone, G. (2015). An analog ensemble for short-term probabilistic solar power forecast. *Applied Energy*, 157:95–110.

Alonso-Suárez, R. (2017). *Estimación del recurso solar en Uruguay mediante imágenes satelitales.* PhD thesis, Facultad de Ingeniería, Universidad de la República. Available at: https://hdl.handle.net/20.500.12008/20200.

Alonso-Suárez, R., Abal, G., Musé, P., and Siri, R. (2014). Satellite-derived solar irradiation map for Uruguay. In *Elsevier Energy Procedia*, volume 57, pages 1237–1246.

Alonso-Suárez, R., Abal, G., Siri, R., and Musé, P. (2012). Brightness-dependent Tarpley model for global solar radiation estimation using GOES satellite images: application to Uruguay. *Solar Energy*, 86(11):3205–3215.

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., de Pison, F. M., and Antonanzas-Torres, F. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136:78–111.

Antonanzas-Torres, F., Urraca, R., Polo, J., Perpiñán-Lamigueiro, O., and Escobar, R. (2019). Clear sky solar irradiance models: A review of seventy models. *Renewable and Sustainable Energy Reviews*, 107:374–387.

Bacher, P., Madsen, H., and Nielsen, H. A. (2009). Online short-term solar power forecasting. *Solar Energy*, 83(10):1772–1783.

Botterud, A. (2017). Chapter 10 - forecasting renewable energy for grid operations. In Jones, L. E., editor, *Renewable Energy Integration (Second Edition)*, pages 133–143. Academic Press, Boston, second edition edition.

Ceballos, J. C., Bottino, M., and de Souza, J. (2004). A simplified physical model for assessing solar radiation over Brazil using GOES 8 visible imagery. *Journal of Geophysical Research: Atmospheres*, 109(D2).

Chernozhukov, V., Fernández-Val, I., and Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica*, 78(3):1093–1125.

Chu, Y. and Coimbra, C. F. (2017). Short-term probabilistic forecasts for direct normal irradiance. *Renewable Energy*, 101:526–536.

Cros, S., Liandrat, O., Sébastien, N., Schmutz, N., and Voyant, C. (2013). Clear sky models assessment for an operational PV production forecasting solution. In *Proceedings of the 28th European PV Solar Energy Conference and Exhibition*, Villepinte, France.

Dambreville, R., Blanc, P., Chanussot, J., and Boldo, D. (2014). Very short term forecasting of the global horizontal irradiance using a spatio-temporal autoregressive model. *Renewable Energy*, 72:291–300.

David, M., Luis, M. A., and Lauret, P. (2018). Comparison of intraday probabilistic forecasting of solar irradiance using only endogenous data. *International Journal of Forecasting*, 34(3):529–547.

David, M., Ramahatana, F., Trombe, P., and Lauret, P. (2016). Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models. *Solar Energy*, 133:55–72.

Diagne, M., David, M., Lauret, P., Boland, J., and Schmutz, N. (2013). Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews*, 27:65–76.

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268.

Golestaneh, F., Pinson, P., and Gooi, H. B. (2016). Very short-term nonparametric probabilistic forecasting of renewable energy generation–with application to solar energy. *IEEE Transactions on Power Systems*, 31(5):3850–3863.

Grantham, A., Gel, Y. R., and Boland, J. (2016). Nonparametric short-term probabilistic forecasting for solar radiation. *Solar Energy*, 133:465–475.

Gueymard, C. A. (2019). Chapter 5 - Clear-sky radiation models and aerosol effects. In Polo, J., Martín-Pomares, L., and Sanfilippo, A., editors, *Solar Resources Mapping: Fundamentals and Applications*, pages 137–182. Springer, Cham.

Gueymard, C. A. and Yang, D. (2020). Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations. *Atmospheric Environment*, 225:117216.

Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5):559–570.

Hoff, T. E. and Perez, R. (2012). Modeling PV fleet output variability. *Solar Energy*, 86(8):2177–2189. Progress in Solar Energy 3.

Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., and Hyndman, R. J. (2016). Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of Forecasting*, 32(3):896–913.

Huang, J., Korolkiewicz, M., Agrawal, M., and Boland, J. (2013). Forecasting solar radiation on an hourly time scale using a Coupled AutoRegressive and Dynamical System (CARDS) model. *Solar Energy*, 87:136–149.

Ineichen, P. (2016). Validation of models that estimate the clear sky global and beam solar irradiance. *Solar Energy*, 132:332–344.

Ineichen, P. and Perez, R. (2002). A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy*, 73(3):151–157.

Inman, R. H., Pedro, H. T., and Coimbra, C. F. (2013). Solar forecasting methods for renewable energy integration. *Progress in Energy and Combustion Science*, 39(6):535–576.

Iversen, E. B., Morales, J. M., Møller, J. K., and Madsen, H. (2016). Short-term probabilistic forecasting of wind speed using stochastic differential equations. *International Journal of Forecasting*, 32(3):981–990.

Jolliffe, I. and Stephenson, D. (2012). *Forecast Verification. A practitioner's guide in atmospheric science*. John Wiley and Sons, Ltd, London, 2nd edition.

Jung, J. and Broadwater, R. P. (2014). Current status and future advances for wind speed and power forecasting. *Renewable and Sustainable Energy Reviews*, 31:762–777.

Kasten, F. (1980). A simple parameterization of the pyrheliometric formula for determining the Linke turbidity factor. *Meteor. Rundschau*, 33:124–127.

Khosravi, A., Nahavandi, S., and Creighton, D. (2013). Prediction intervals for short-term wind farm power generation forecasts. *IEEE Transactions on Sustainable Energy*, 4(3):602–610.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Kühnert, J., Lorenz, E., and Heinemann, D. (2013). Chapter 11 - Satellite-based irradiance and power forecasting for the german energy market. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 267–297. Academic Press, Boston.

Laguarda, A., Giacosa, G., Alonso-Suárez, R., and Abal, G. (2020). Performance of the site-adapted CAMS database and locally adjusted cloud index models for estimating global solar horizontal irradiation over the Pampa Húmeda. *Solar Energy*, 199:295–307.

Lauret, P., David, M., and Pedro, H. T. C. (2017). Probabilistic solar forecasting using quantile regression models. *Energies*, 10(10).

Lauret, P., David, M., and Pinson, P. (2019). Verification of solar irradiance probabilistic forecasts. *Solar Energy*, 194:254–271.

Lauret, P., Voyant, C., Soubdhan, T., David, M., and Poggi, P. (2015). A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. *Solar Energy*, 112:446–457.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Qu, Z., Wald, L., Homscheidt, M. S., and Arola, A. (2013). McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques, European Geosciences Union*, 6:2403–2418.

Leutbecher, M. and Palmer, T. N. (2008). Ensemble forecasting. *J. Comput. Phys.*, 227(7):3515–3539.

Lorenz, E., Hammer, A., and Heinemann, D. (2004). Short term forecasting of solar radiation based on satellite data. In

EUROSUN2004 (ISES Europe Solar Congress), pages 841–848, Freiburg, Germany.

Marchesoni-Acland, F., Lauret, P., Gómez, A., and Alonso-Suárez, R. (2019). Analysis of ARMA solar forecasting models using ground measurements and satellite images. In *IEEE 46th Photovoltaic Specialists Conference (PVSC)*, pages 2445–2451.

Mellit, A. and Pavan, A. M. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at trieste, italy. *Solar Energy*, 84(5):807–821.

Pedro, H. T. and Coimbra, C. F. (2012). Assessment of forecasting techniques for solar power production with no exogenous inputs. *Solar Energy*, 86(7):2017–2028.

Pedro, H. T., Coimbra, C. F., David, M., and Lauret, P. (2018). Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts. *Renewable Energy*, 123:191–203.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F. (2002). A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, 73(5):307–317.

Pinson, P., McSharry, P., and Madsen, H. (2010). Reliability diagrams for non-parametric density forecasts of continuous variables: Accounting for serial correlation. *Quarterly Journal of the Royal Meteorological Society*, 136(646):77–90.

Pinson, P., Nielsen, H. A., Møller, J. K., Madsen, H., and Kariniotakis, G. N. (2007). Non-parametric probabilistic forecasts of wind power: required properties and evaluation. *Wind Energy*, 10(6):497–516.

Polo, J., Antonanzas-Torres, F., Vindel, J., and Ramirez, L. (2014). Sensitivity of satellite-based methods for deriving solar radiation to different choice of aerosol input and models. *Renewable Energy*, 68:785–792.

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., and Wald, L. (2017). Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. *Meteorologische Zeitschrift*, 26(1):33–57.

Reikard, G. (2009). Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Solar Energy*, 83(3):342–349.

Rigollier, C., Bauer, O., and Wald, L. (2000). On the clear sky model of the ESRA - European Solar Radiation Atlas - with respect to the Heliosat method. *Solar Energy*, 68(1):33–48.

Rigollier, C., Lefevre, M., and Wald, L. (2004). The method Heliosat-2 for deriving shortwave solar radiation from satellite images. *Solar Energy*, 77(2):159–169.

Sengupta, M., Habte, A., Kurtz, S., Dobos, A., Wilbert, S., Lorenz, E., Stoffel, T., Renné, D., Gueymard, C., Myers, D., Wilcox, S., Blanc, P., and Perez, R. (2015). Best practices handbook for the collection and use of solar resource data for solar energy applications. Tech. Report NREL/TP-5D00-63112, NREL-IEA.

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., and Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569–582.

Wilks, D. S. (2009). *Statistical methods in the atmospheric sciences*. International geophysics series. Elsevier, Amsterdam, 2nd edition.

Wu, X. and Sun, F. (2005). Post-launch calibration of GOES imager visible channel using MODIS. In *Proc. SPIE*, volume 5882.

Yang, D. (2019). A universal benchmarking method for probabilistic solar irradiance forecasting. *Solar Energy*, 184:410–416.

Yang, D. (2020). Choice of clear-sky model in solar forecasting. *Journal of Renewable and Sustainable Energy*, 12(2):026101.

Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T., and Coimbra, C. F. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy*, 168:60–101. Advances in Solar Resource Assessment and Forecasting.

Zhong, X. and Kleissl, J. (2015). Clear sky irradiances using rest2 and modis. *Solar Energy*, 116:144–164.