

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

3 R. Alonso-Suárez^{a,*}, M. David^b, V. Branco^{a,c}, P. Lauret^b

4 ^aLaboratorio de Energía Solar, Facultad de Ingeniería, UDELAR. J. H. y Reissig 565, Montevideo, Uruguay

5 ^bUniversity of La Réunion - PIMENT laboratory, 15, avenue René Cassin, 97715 Saint-Denis

6 ^cLaboratorio Solar, ITR Centro-Sur, Universidad Tecnológica. F. Maciel esq. L. Morquio, CP 97000, Durazno, Uruguay.

7 **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.

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

9 **1. Introduction**

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

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

20 Solar predictions are intended to anticipate at time (t) the future behavior (fluctuations) of ground level
21 solar irradiance as a result of the complex clouds and atmosphere's dynamics. Two types of forecasting
22 methods can be considered namely deterministic and probabilistic. Both methods attempt to predict the
23 future outcome of a variable but information on the probability distribution of the prediction is only present
24 in the probabilistic forecast. Many research efforts have been devoted to the development of models for
25 generating solar deterministic forecasts. However, a forecast is inherently uncertain and a proper assess-
26 ment of its probability distribution offers the grid operator a more informed decision-making framework.
27 Thus, a deterministic forecast plus predictions intervals for different confidence levels provides more useful
28 information to manage the renewable sources uncertainty in a cost efficient and reliable manner (Botterud,
29 2017). Contrary to wind power forecasting where probabilistic forecasting appears to be a mature subject
30 (Pinson et al., 2007; Jung and Broadwater, 2014; Iversen et al., 2016), probabilistic solar forecasting is still
31 in its first stages (Antonanzas et al., 2016; Hong et al., 2016).

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

55 work.

56 Probabilistic forecast techniques aim to predict the future probability distributions of a prediction vari-
57 able. The predictive distribution can be represented either by a Cumulative Distribution Function (CDF) or
58 a Probability Density Function (PDF). Similarly to deterministic forecasts, different types of probabilistic
59 models can be used depending on the forecast horizon. [Golestaneh et al. \(2016\)](#) empirically observed that
60 solar forecasting errors do not follow any of the most common parametric probability densities. Thus, they
61 proposed a non-parametric approach based on neural networks to produce a very short-term PV power prob-
62 abilistic forecasting using as inputs the power output and meteorological measurements time-series, including
63 solar irradiance. The proposed forecasting method was evaluated for lead times between a few minutes to
64 one hour ahead and proved to produce a better probabilistic forecast than a selected set of benchmark
65 models. [Grantham et al. \(2016\)](#) also proposed a non-parametric approach, but based on a specific bootstrap
66 method to build predictive global horizontal irradiance (GHI) probability distributions. Using solar irradi-
67 ance ground measurements and a map of the solar positions in the sky, they produce reliable (used here in
68 the sense of the reliability property, explained in [Subsection 4.1](#)) probabilistic forecast for a time horizon of
69 one hour. As the bootstrap method tends to reproduce the distribution of the climatology observed under
70 specific conditions, the predictions are not sharp (see [Subsection 4.2](#) for the description of the sharpness
71 property). [Chu and Coimbra \(2017\)](#) developed an ensemble model based on the k-nearest-neighbors (kNN)
72 algorithm for generating very short-term direct normal irradiance (DNI) probabilistic forecasts from 5 to 20
73 minutes ahead. This forecast proposal is based on the time-lagged DNI measurements, diffuse irradiance
74 measurements and three features extracted from all-sky camera images. [David et al. \(2016\)](#) used a para-
75 metric approach based on a combination of two linear time series model (ARMA and GARCH) to generate
76 10 minutes to 6 hours ahead GHI probabilistic forecasts using only past ground data. Recently, [David et al.](#)
77 [\(2018\)](#) made a comparison of 20 intra-day solar probabilistic models that used only solar irradiance ground
78 measurements and found that models based on linear quantile regression were one of the best performers.
79 The authors also conclude that future improvements could be obtained by including relevant exogenous vari-
80 ables in the models. Also, [Lauret et al. \(2017\)](#) showed that the quality of intra-day probabilistic forecasts
81 can be improved by considering both past ground measurements and day-ahead NWP forecasts (provided
82 by the European Center for Medium-Range Weather Forecast, ECMWF) in the set of explanatory variables.
83 [Pedro et al. \(2018\)](#) used a machine learning approach to generate very short-term probabilistic forecasts of
84 GHI and DNI. The input features of the proposed models were derived from the irradiance time series and
85 sky images. In particular, a variability index was calculated from the GHI time series but the impact of this
86 predictor was not assessed on the quality of the forecasts.

87 To the best of our knowledge, no works have been published regarding the use of satellite data and/or
88 short-term variability information for intra-day solar probabilistic forecasts. Therefore, this work investigates
89 if a combination of ground telemetry with variability and satellite information could improve the quality

90 of the probabilistic forecasts. As the goal here is to assess the added value of such a combination, we
91 restrict ourselves to a simple probabilistic statistical technique to build the different models, namely the
92 linear quantile regression (Koenker and Bassett, 1978). More precisely, we build three probabilistic models
93 that generate three hours ahead probabilistic forecasts using a 10 minutes granularity. The first model
94 makes use of actual and past solar ground measurements. The second one adds the solar variability as an
95 additional predictor to the past ground measurements while the third one enriched the second by adding
96 spatially-averaged satellite albedo as an exogenous input. In order to evaluate the quality of the probabilistic
97 forecasts, we use the evaluation framework provided by Pinson et al. (2007), recently adapted by Lauret
98 et al. (2019) for the solar irradiance forecasting field.

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

103 2. Data

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

110 2.1. Ground measurements

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

121 It is common practice in the field to work with the clear sky index, k_c , rather than the GHI data. The
122 GHI time series, G_h , have a diurnal and seasonal geometrical behavior that result from the Sun’s apparent

123 movement and introduces a deterministic component into the times series that can be removed using accurate
124 estimations of the clear sky irradiance. Indeed, an estimation of the clear sky GHI, G_h^{csk} , can be used to
125 calculate k_c as defined in Eq. (1), removing the geometrical behavior and the cloudless atmosphere's slow
126 variations, isolating the fluctuations only due to cloudiness;

$$k_c(t) = \frac{G_h(t)}{G_h^{\text{csk}}(t)}. \quad (1)$$

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

150 State-of-the-art forecasting methods described in Section 1 generally make use of the clear sky index,
151 k_c , rather than the clearness index, k_t , being the latter the normalization of the GHI by its corresponding
152 extraterrestrial irradiance at horizontal plane. Although k_t is a popular variable for decomposition and
153 transposition irradiance models, mostly due to its simple calculation, it exhibits significant correlations with
154 the solar zenith angle and the air mass, resulting in artifacts during early mornings and late afternoons. In

155 fact, k_t does not provide the well-behaved deseasonalized signal that is required for forecasting methods.
 156 The clear sky index overcomes this issue by providing an almost stationary time series that is representative
 157 of the cloudiness only. Hence, we prefer here the use of k_c instead of k_t , which is commonly admitted in
 158 the solar forecasting community (Inman et al., 2013; Diagne et al., 2013; Sengupta et al., 2015; Yang et al.,
 159 2018) and usually preferred by solar forecasting researchers (Yang, 2020).

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

167 Table 1 summarizes the main characteristics of the different sites considered in this work, including its
 168 location, provider and the site’s climate type. The code (first row) will be used as a short reference to the
 169 station. The short-term variability of the time series is also provided in Table 1, and it is derived from each
 170 station filtered data as the standard deviation of the 10 minutes clear sky changes: $\sigma_c = \text{std}\{\Delta k_c\}$. A site
 171 with variability $\sigma_c > 0.2$ is considered as experiencing very unstable GHI conditions (Hoff and Perez, 2012).
 172 As shown in Table 1, the solar irradiance variability of the Oahu, Fouillole and Tampon stations are around
 173 or above this threshold, which is typical of insular sites where partly cloudy conditions prevail. On the other
 174 hand, Desert Rock is a site where clear sky days prevail, thus it has a small GHI variability. The rest of the
 175 sites have intermediate conditions in terms of GHI variability (see Table 1), with a balance mix of clear sky,
 176 partly cloudy and cloudy days. Finally, for each station, two years of data, as shown in the last two rows
 177 of Table 1, are used as follow: the oldest year is used to train the quantile regression models and the most
 178 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

179 *2.2. Local short-term variability*

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

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

184 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
185 considered in the calculation of the standard deviation. We use $l_v = 5$, the same quantity of lags that will
186 be used for building the quantile regression models, so this local short-term variability accounts for the past
187 one hour from time (t). At the beginning of the day this calculation may not be indicative of the current
188 sky instability, as the moving window would consider samples from the present day mixed with samples of
189 the previous day (night-time values are removed by the solar altitude filter), which may have different sky
190 conditions. To solve this, the computation of $\sigma_c^{l_v}$ only takes into account the non-filtered daylight samples
191 of the current day, so the time window that is used for its calculation at the beginning of the day is adapted
192 to meet this requirement. The variability for the first sample of the day (after filtering, it will be a daylight
193 sample) is set to zero as a default value.

194 *2.3. Satellite information*

195 Visible channel satellite images measure the solar radiance reflected at each pixel of the Earth surface.
196 By knowing the extraterrestrial irradiance at the top of the atmosphere, it is possible to derive the Earth
197 surface reflectance, known as albedo, ρ_p . This information provided by geostationary satellites is the main
198 input for satellite based models for intra-hour solar irradiance assessment (Perez et al., 2002; Ceballos et al.,
199 2004; Rigollier et al., 2004; Alonso-Suárez et al., 2014; Qu et al., 2017). In this work, rather than using
200 the solar satellite estimates, we use directly the satellite albedo as an input for the probabilistic models.
201 This seeks to quantify the impact of satellite cloudiness in the forecast performance without using a solar
202 irradiance model that may be adding uncertainty to the problem. In absence of snow or any other kind of
203 high albedo terrain, a low albedo represents the ground surface and a high albedo represents the presence of
204 clouds, as they reflect more radiation to the outer space than the ground. Such is the case of the two South
205 American sites where satellite information is considered, whose ground albedo is around 5-15 % during all
206 the year. Albedo samples ranging from 20 % to 100 % represent the presence of cloudiness, in its various
207 types.

208 For the two considered sites, LE and MS, satellite information is used by averaging the satellite albedo
209 pixels in a 10 min \times 10 min latitude-longitude cell around the site. This cell size is known to be the optimal to

210 reduce the uncertainty of hourly irradiation assessments in the region (Laguarda et al., 2020; Alonso-Suárez,
 211 2017) and it is intended to represent the average hourly cloudiness using an ergodic assumption. As the
 212 satellite cloudiness is spatially averaged, information of the near future cloudiness over the target site is
 213 directly included. This is a very simple way to introduce satellite derived data as exogenous variable into
 214 solar forecasts. In this way, both satellite information and short-term variability lags represent approximately
 215 one hour. Finally, it has to be noted that satellite images for South America during the considered period
 216 (till year 2017) have a rate of two per hour, so the required 10 minutes granularity was obtained by linear
 217 interpolation using the available albedo samples. The images were downloaded in raw format from the
 218 NOAA CLASS website (<https://www.avl.class.noaa.gov>) and the calibration procedures for obtaining the
 219 corrected albedo were applied as recommended by NOAA (Wu and Sun, 2005).

220 3. Probabilistic forecast

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

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

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

233 where F is the CDF of a random variable Y , defined as $F(y) = \Pr(Y \leq y)$. In other words, a quantile
 234 q_τ indicates that there is a τ probability that the observation falls below the quantile q_τ . In this work,
 235 these quantiles are estimated for the clear sky index (k_c) and each time horizon, for the nominal proportions
 236 $\tau = [0.1, 0.2, \dots, 0.9]$, resulting in different sets of k_c quantiles $q_\tau = [q_1, q_2, \dots, q_9]$ that define each CDF.
 237 More precisely, the output (at time t) of the different probabilistic models that are generated in this work is
 238 the ensemble of nine clear sky index quantiles $q_\tau(t + \Delta t)$ for each forecasting time horizon $\Delta t = 10, 20, \dots, 180$
 239 minutes. For the sake of completeness and proper calculation of the performance metrics, the quantiles q_0
 240 and q_{10} that correspond to the nominal proportions $\tau = 0.0$ and $\tau = 1.0$ where set to zero and to the

241 maximum k_c value in the time series, respectively. Then, using Eq. (1), the k_c quantiles are converted to the
 242 required GHI quantiles. Prediction intervals with different nominal coverage rates can be inferred from the
 243 set of quantiles. Prediction intervals give a range of values in which the true value of GHI is expected to lie
 244 with a certain probability, namely, the nominal coverage rate. In this study, we chose prediction intervals
 245 with a nominal coverage rate of 80% as it leads to a good compromise as stated by (Pinson et al., 2007).
 246 Hence, the $(1 - \alpha) \times 100\%$ central prediction interval is generated by taking the $\frac{\alpha}{2}$ quantile as the lower
 247 bound and the $1 - \frac{\alpha}{2}$ quantile as the upper bound. In this case, the 80% confident interval ($\alpha = 0.2$) is
 248 obtained by considering the q_1 and q_9 GHI quantiles.

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

262 3.1. The Linear Quantile Regression (LQR) method

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

$$q_\tau = \beta_\tau X + \epsilon, \quad (4)$$

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

272 Following (Koenker and Bassett, 1978), and based on a training set of N samples, the quantiles $q_\tau =$
 273 $F^{-1}(\tau)$ can be estimated as the solution of the following optimization problem,

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

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

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

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

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

278 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
 279 of observed predictands and the set of predictors (Y_i and X_i) for the adjustment of $\hat{\beta}_\tau$ are taken from the
 280 training set. It must be noted that the quantile regression method estimates each quantile separately (i.e.
 281 the minimization of the quantile loss function is made for each τ separately). As a consequence, one can
 282 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
 283 used the rearrangement method described by (Chernozhukov et al., 2010).

284 3.2. Implemented regression models

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

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

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

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

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

294 The third model (L5-S) takes the current time satellite albedo $\rho_p(t)$ averaged over a $10 \text{ min} \times 10 \text{ min}$
 295 latitude-longitude cell around the specific forecast site, as explained in [Subsection 2.3](#). Its input matrix is
 296 expressed as:

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

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

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

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

301 3.3. Persistence ensemble

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

315 4. Performance metrics

316 Probabilistic forecast cannot be evaluated by using the standard deterministic performance metrics,
 317 such as the Mean Bias Error (MBE), Root Mean Squared Error (RMSE) and Forecasting Skill (FS), among
 318 other proposals. The computation of these quantities requires to compare two deterministic samples: a

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

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

352 *4.1. Reliability Property*

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

362 *4.2. Sharpness*

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

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

369 where Δt is the forecast horizon and α is the confidence level that corresponds to the nominal coverage rate
 370 $(1 - \alpha) \times 100\%$. $\hat{G}_{h,i|1-\frac{\alpha}{2}}(t + \Delta t)$ and $\hat{G}_{h,i|\frac{\alpha}{2}}(t + \Delta t)$ are the time series of the superior and inferior predicted
 371 GHI quantiles for the confidence level α . As stated before, PINAW is assessed in this work for a confidence
 372 level $\alpha = 0.2$ which correspond to a 80% nominal coverage. $G_h(t + \Delta t)$ is the actual GHI measurement at
 373 the prediction time. Notice that the N value required for calculating both averages is canceled in Eq. (12).

374 *4.3. CRPS*

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

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

378 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
 379 CDF step function that jumps from zero to one at the point where x equals the future observation x_0

380 $(F_{x_0}(x) = 1|_{\{x \geq x_0\}})$. The integrated squared difference between the two CDFs is averaged over the N pairs
 381 of forecast/observation data. A different value of CRPS is computed for each time horizon Δt .

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

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

$$\text{CRPSS} = 100 \times \left(1 - \frac{\text{CRPS}_m}{\text{CRPS}_o} \right), \quad (14)$$

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

395 5. Results and discussion

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

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

408 5.1. Reliability assessment

409 The reliability diagrams provide a visual inspection between estimated and nominal proportions. One
 410 form of visualization is to plot the deviations, i.e. the difference between both proportions, as a function of

411 the nominal proportion. In such case, the perfect reliability corresponds to a zero line plot. Fig. 1 shows
412 the deviations in the quantile forecasts for the four selected sites. The quantitative values of the plots are
413 given in the Appendix A in Table A.2 and Table A.3, including the other sites.

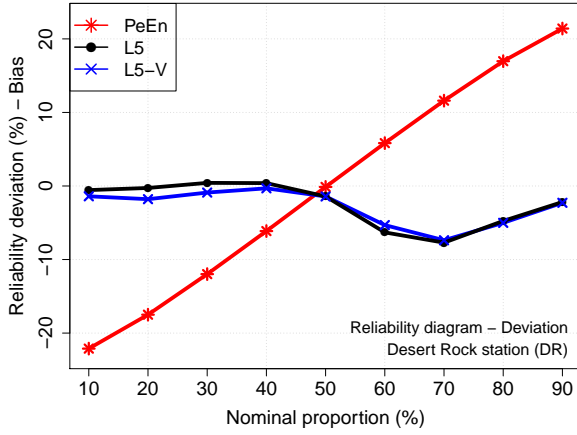
414 From the reliability diagrams, one can clearly argue that the Persistence (PeEn) model is not reliable.
415 Indeed, the pattern exhibited by the PeEn model (that overforecast quantiles for nominal proportion above
416 $\simeq 50\%$ and underforecast quantiles for nominal proportion below $\simeq 50\%$) is typical of models that generate
417 too narrow predictive distributions. As the persistence is not reliable, we should normally discard this model
418 at this point of the evaluation framework.

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

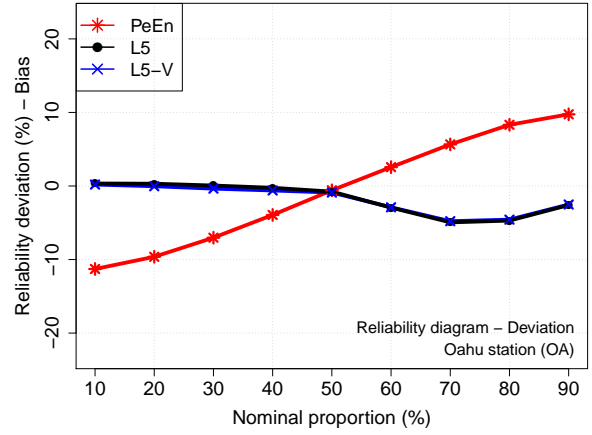
428 5.2. Sharpness assessment

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

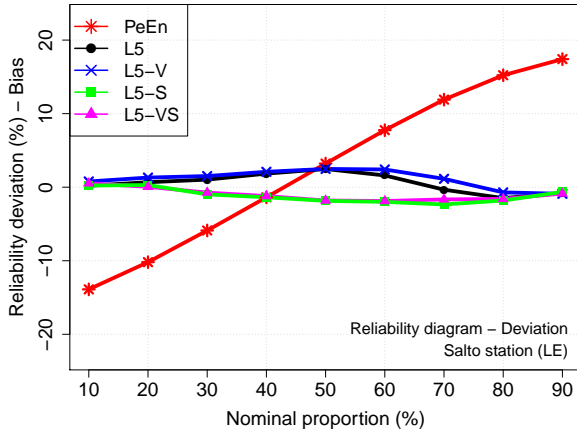
443 The sharpness evaluation indicates that short term variability reduces the interval forecasts (blue line)
444 albeit the picture is less clear for the Oahu site. In fact, in Subsection 5.4 it is shown that short-term



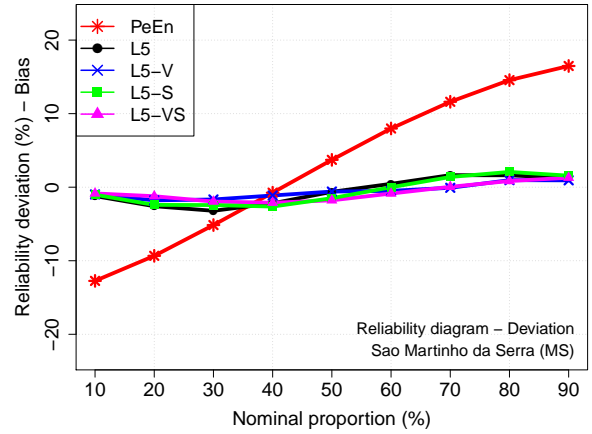
(a) Desert Rock station (DR).



(b) Oahu station (OA).



(c) Salto station (LE).

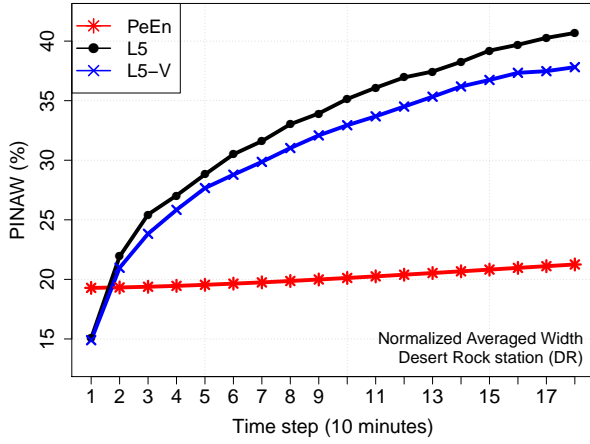


(d) São Martinho da Serra station (MS).

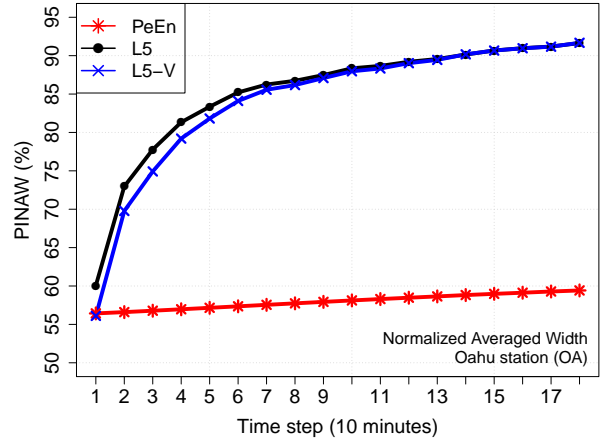
Figure 1: Reliability deviations of the inspected models in the four selected sites.

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

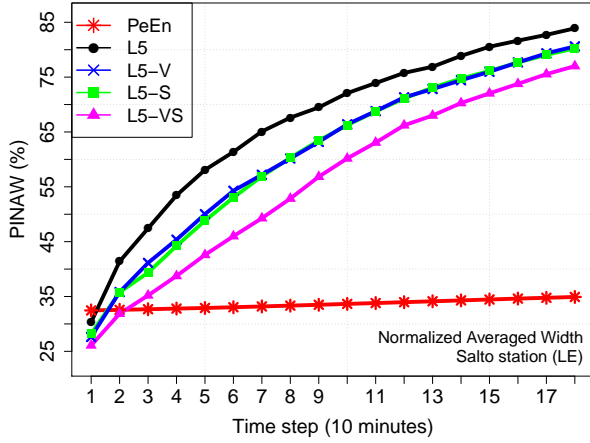
449 The spatially averaged satellite albedo tested for the two intermediate variability sites, LE and MS, also
 450 shows an improvement in the PINAW metric (see Figs. 2c and 2d). The models that only include satellite
 451 information obtain a PINAW trend similar to that of models that only include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



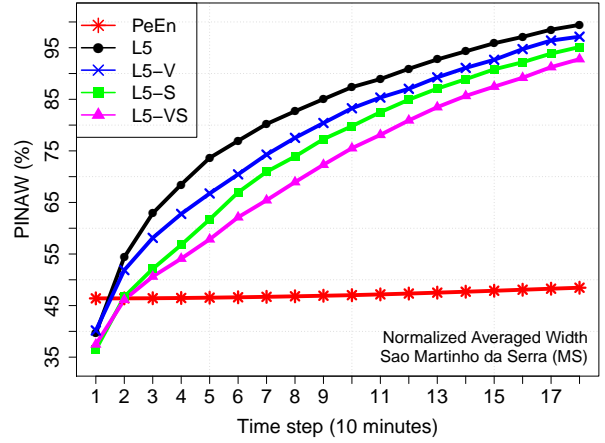
(a) Desert Rock station (DR).



(b) Oahu station (OA).



(c) Salto station (LE).



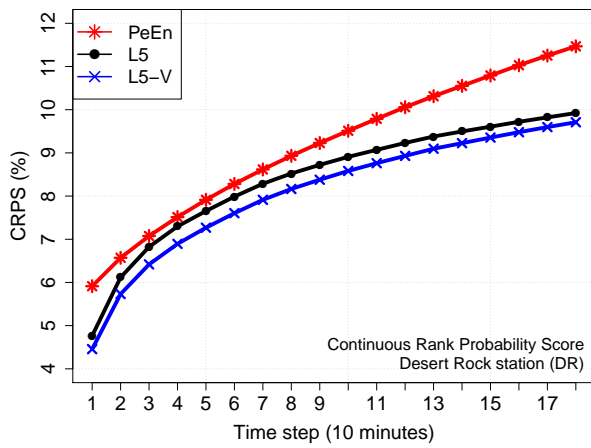
(d) São Martinho da Serra station (MS).

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

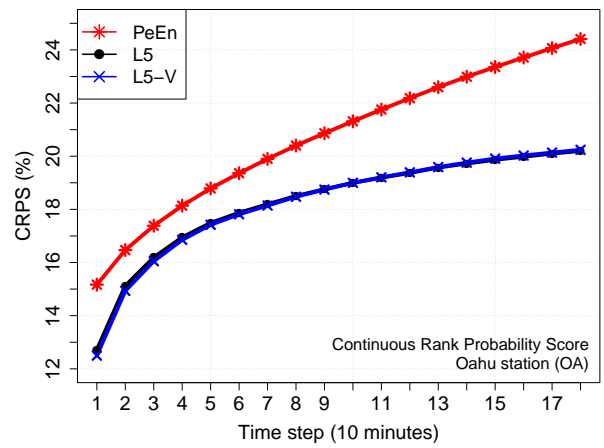
455 5.3. Overall skill with the CRPS

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

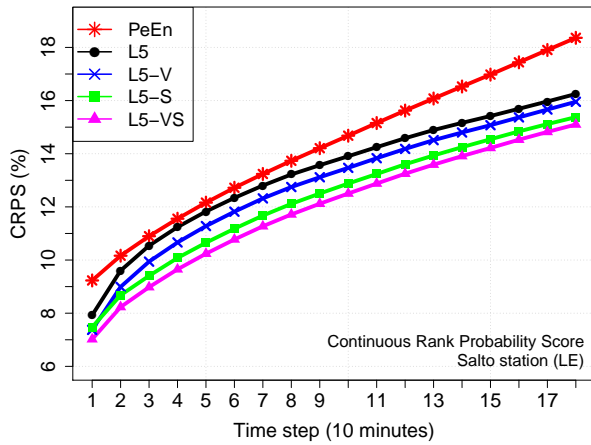
465 To better visualize the gain of including the exogenous variables in the LQR model, Fig. 5 provides the
 466 CRPSS difference between the L5-V, L5-S and L5-VS models with respect to the L5 model. This allows to
 467 isolate the improvement of adding the different variables to the baseline lagged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



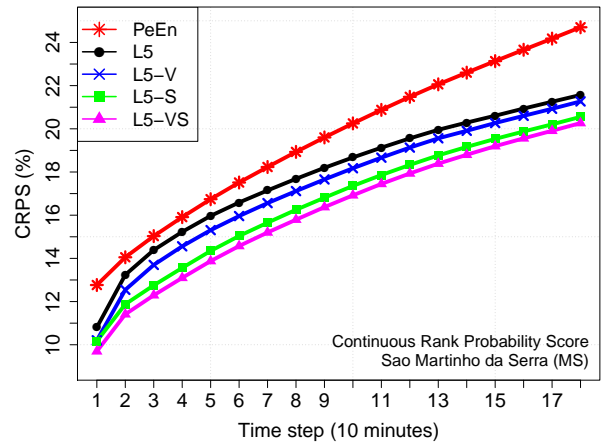
(a) Desert Rock station (DR).



(b) Oahu station (OA).

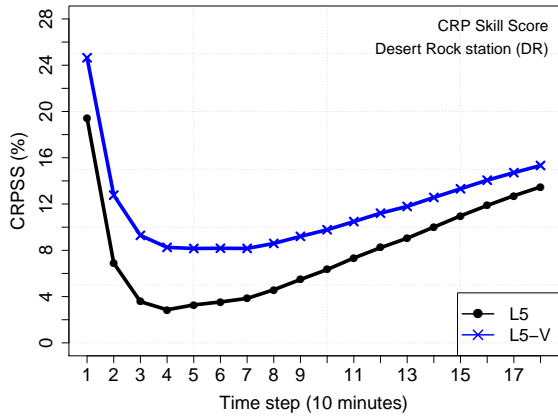


(c) Salto station (LE).

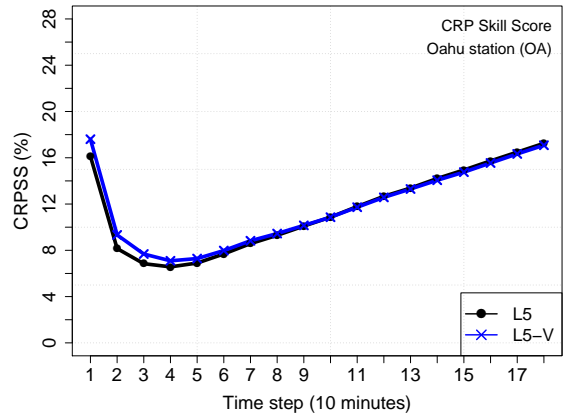


(d) São Martinho da Serra station (MS).

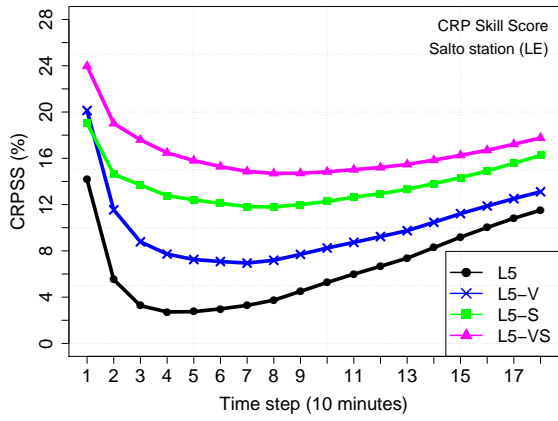
Figure 3: CRPS of the inspected models in the different solar variability sites.



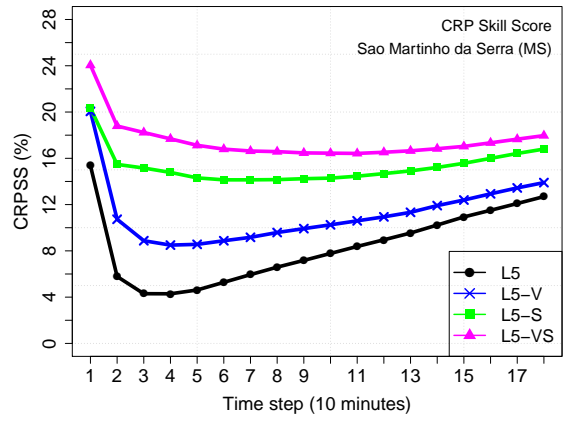
(a) Desert Rock station (DR).



(b) Oahu station (OA).

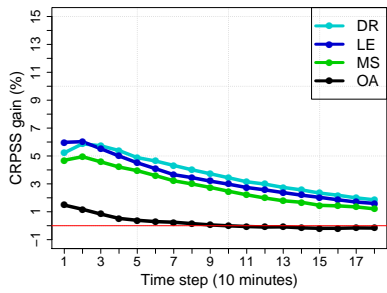


(c) Salto station (LE).

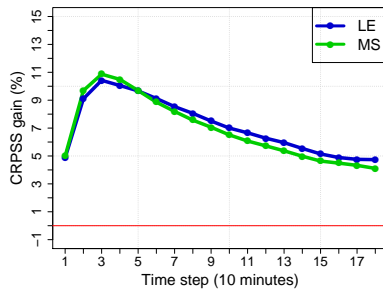


(d) São Martinho da Serra station (MS).

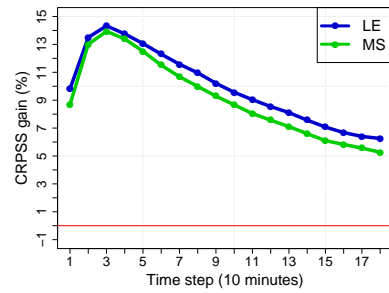
Figure 4: CRPSS of the inspected models in the different solar variability sites.



(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.

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

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

489 The best results are obtained by including both, short-term variability and satellite albedo, as shown in
490 Fig. 3, Fig. 4 and Fig. 5c. The CRPSS gain reaches $\simeq 14\%$, again, at $\Delta t = 30$ minutes, but is increased
491 for the first lead time (9-10% gain) in comparison with the last lead time (6-7% gain), as can be seen in
492 Fig. 5c. In this case, a slightly better improvement is observed for LE (see Fig. 5c). This is consistent with
493 the improvement observed for the L5-V model (where only short-term variability is added), for which the
494 CRPSS improvement is also slightly better for the LE site than for the MS site (see Fig. 5a). To sum up,
495 satellite albedo has a similar impact in both sites, but short-term variability is slightly more useful in the
496 LE site, and the plots for the L5-V, L5-S and L5-VS are consistent with each other. Further, as it can
497 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
498 L5-VS model, the performance of the ‘complete’ L5-VS model can be understood as roughly adding the
499 performance gain of each variable. This is also approximately verified for the first and largest lead times.

500 5.4. Analysis by sky condition

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

508 For this analysis, it is required to classify the samples into clear sky and non clear sky. This was done by

509 using the clear sky index and the short-term variability index, identifying the region in the two-dimensional
510 plot in where the clear sky samples lay. The region was defined tight enough to ensure that only clear sky
511 samples were selected for the clear sky data set. Visual inspection of the time-series was carried out to
512 check that no cloudy-contaminated samples were included in the clear sky data set. Thus, the non clear sky
513 data set (samples which are not in the clear sky data set) corresponds to partly cloudy and overcast sky
514 conditions.

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

519 *5.4.1. Analysis for L5-V*

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

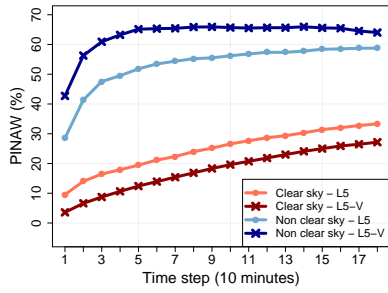
543 center panels), therefore the variable has a significant impact in these cases (even for the high variability
544 site of OA).

545 *5.4.2. Analysis for L5-S*

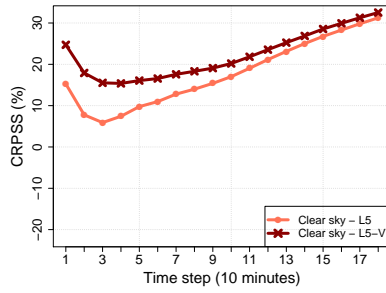
546 The performance comparison between the L5 baseline model and the L5-S model is provided in Fig. 7.
547 The inclusion of spatially averaged satellite albedo is only tested here for intermediate variability sites
548 (LE and MS, one in each row of Fig. 7). In these sites, it significantly improves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

561 *5.4.3. Analysis for L5-VS*

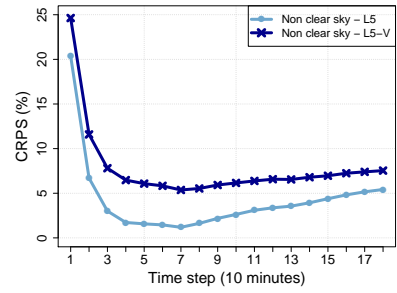
562 In the ‘complete’ L5-VS model, the inclusion of both variables compensates the drawbacks of each other.
563 The performance comparison with the L5 baseline model is presented in Fig. 8. As it can be seen in the
564 center and right panels, the L5-VS model outperforms the CDF prediction of the L5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



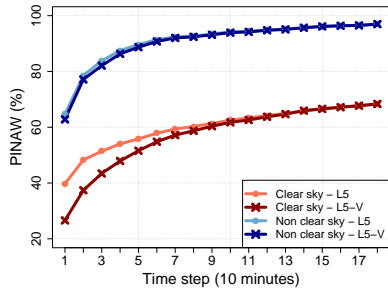
(a) DR site: PINAW metric.



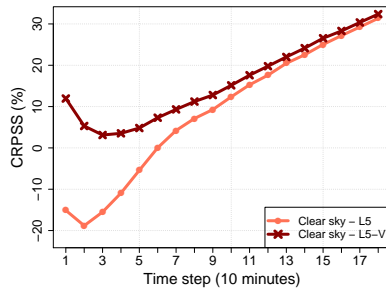
(b) DR site: clear sky CRPSS.



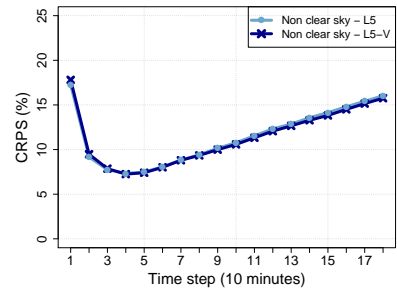
(c) DR site: non clear sky CRPSS.



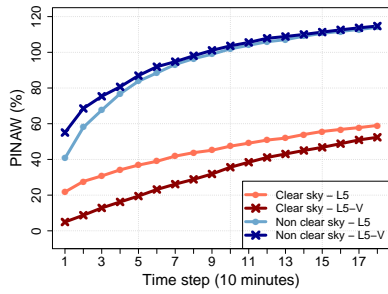
(d) OA site: PINAW metric.



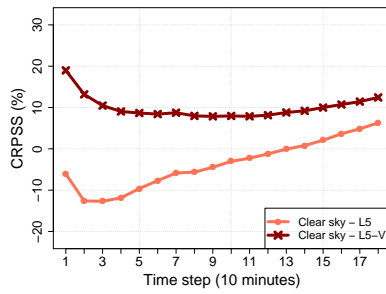
(e) OA site: clear sky CRPSS.



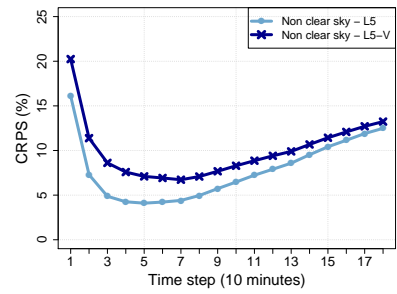
(f) OA site: non clear sky CRPSS.



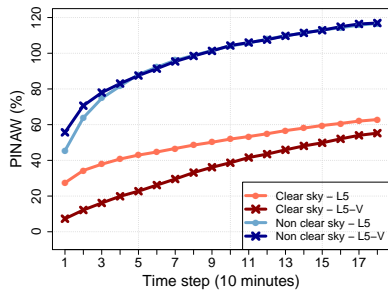
(g) LE site: PINAW metric.



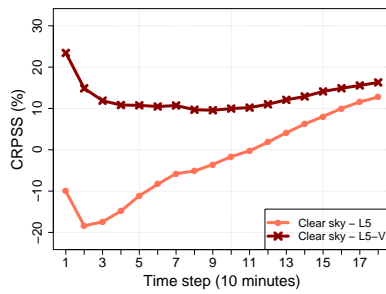
(h) LE site: clear sky CRPSS.



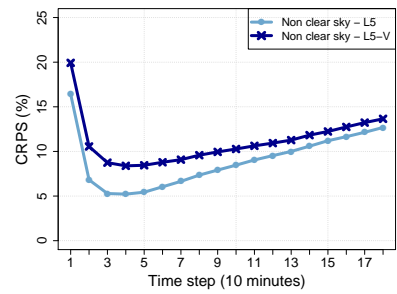
(i) LE site: non clear sky CRPSS.



(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.

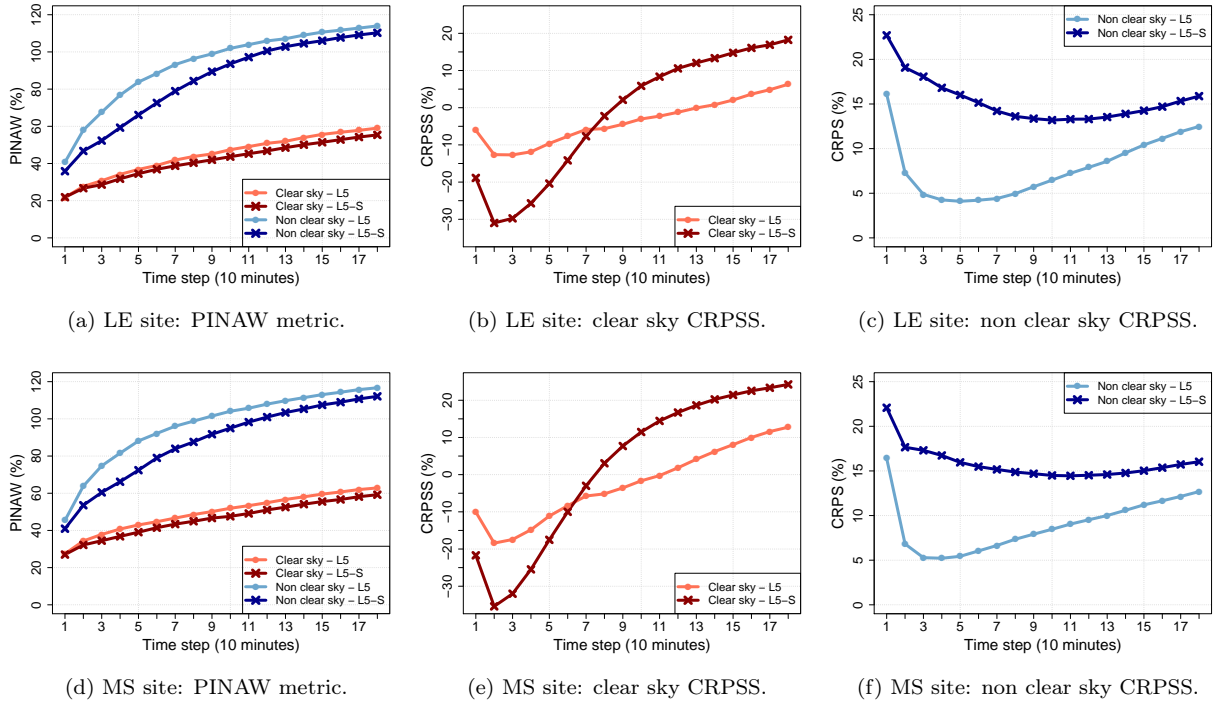


Figure 7: Performance assessment for the L5-S model discriminated by sky condition. The figures show the impact of including spatially averaged satellite albedo to the baseline L5 model for the two intermediate solar variability sites under study.

573 6. Conclusions

574 In this work, we quantified the impact of adding two exogenous variables to intra-day linear solar proba-
575 bilistic forecast: (a) local short-term solar irradiance variability and (b) spatially averaged satellite albedo.
576 Evidence that these information improve the forecast in both cases is provided. The Linear Quantile Re-
577 gression method was used with lagged past measurements and the extra variables to forecast the global
578 horizontal irradiance up to 3 hours ahead with a 10 minutes granularity. Short-term variability is tested in
579 7 sites, accounting for different climates that exhibit high, intermediate and low solar variability conditions.
580 Satellite albedo is tested in two intermediate variability sites where we had satellite images availability. A
581 detailed performance assessment was provided and the added value of each variable was quantified, both
582 separately and jointly. Furthermore, an analysis depending on the sky condition was presented, allowing us
583 to visualize the impact of each variable in the probabilistic forecast at clear and non clear sky conditions.

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

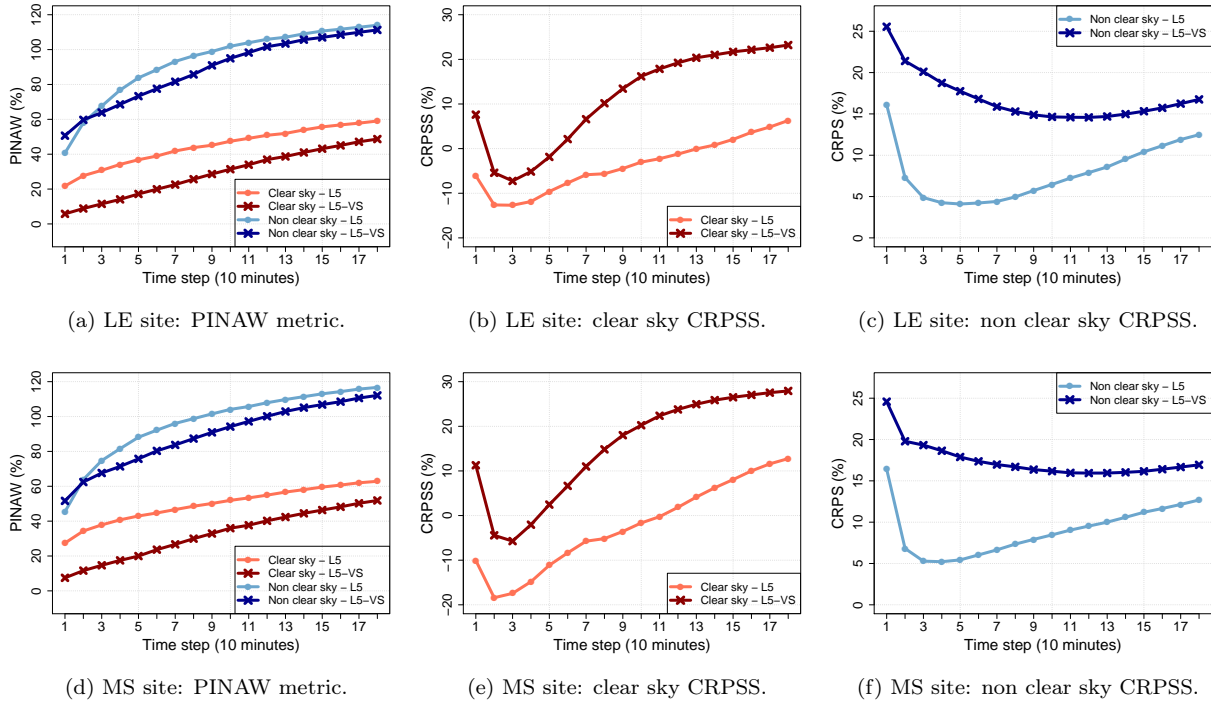


Figure 8: Performance assessment for the ‘complete’ L5-VS model discriminated by sky condition. The figures show the impact of including both inspected variables to the baseline L5 model for the two intermediate variability sites under study.

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

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

598 The inclusion of the satellite albedo to the L5 baseline model (L5-S) showed a very important im-
599 provement in the CDFs prediction accuracy, outperforming the overall value of short-term variability for
600 intermediate variability sites. There is also an improvement in the overall average interval width, but at a
601 less extent. When discriminating the impact into clear and non clear skies, it is observed that the satellite
602 predictor has an important effect on increasing the CDFs prediction accuracy under cloudy sky while at the
603 same time reducing its average width. For clear sky the average width is also reduced, but slightly, and
604 the effect on the CDFs accuracy depends on the time horizon. The drawbacks observed when adding each

605 variable separately, namely, higher cloudy sky PINAW when adding short-term variability and lower clear
606 sky CPRSS at shorter lead times when adding satellite information, are compensated. In fact, it is possible
607 to roughly understand the PINAW and CRPSS behavior of the L5-VS model by adding the value of each
608 variable with respect to the L5 model.

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

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

634 Acknowledgments

635 The authors gratefully acknowledge the laboratory PIMENT from the University of La Réunion, the
636 laboratory LARGE from the University of Guadeloupe, the National Renewable Energy Laboratory (NREL),
637 and the NOAA's SURFRAD meteorological network for providing part of the ground measurements used

638 in this work. The authors acknowledge the financial support given by Uruguay’s National Research and
639 Innovation Agency (ANII) under the FSE-ANII-2016-131799 grant.

640 **A. Performance detailed evaluation**

641 This Appendix provides the quantitative values of the performance assessment for all inspected sites.
642 Tables A.2 and A.3 present the bias (reliability deviations) and Tables A.4 and A.5 the PINAW metric.
643 Tables from Table A.6 to Table A.9 present the CRPS and CRPSS. This Appendix also presents the overall
644 metrics’ plots for the FP, FO and TM sites (see Fig. A.9), which were not included in the general analysis
645 for the sake of simplicity. FP is an intermediate variability site and hence its results are similar to those of
646 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	
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	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	
Δt_i	L5-S	L5-VS	L5-S	L5-VS	Δt_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

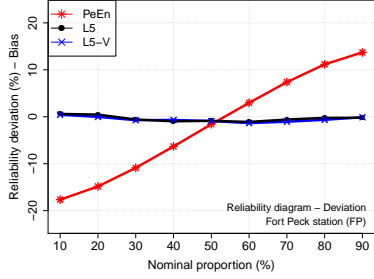
Table A.7: CRPS of L5S and L5VS models

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

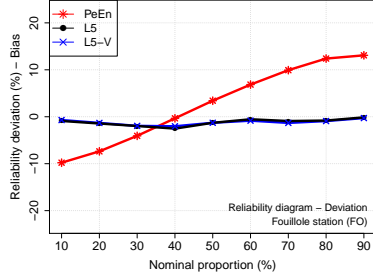
Table A.8: CRPSS of L5 and L5V models

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

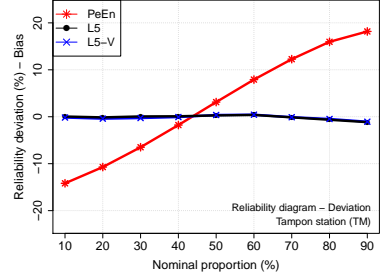
Table A.9: CRPSS of L5S and L5VS models



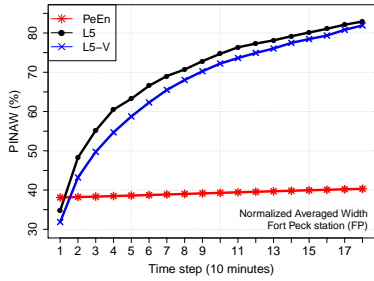
(a) FP site: reliability diagram.



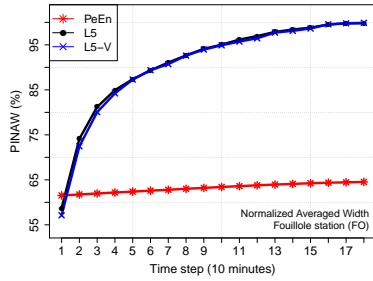
(b) FO site: reliability diagram.



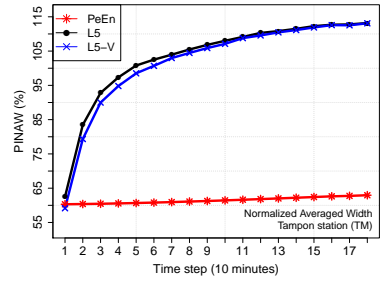
(c) TM site: reliability diagram.



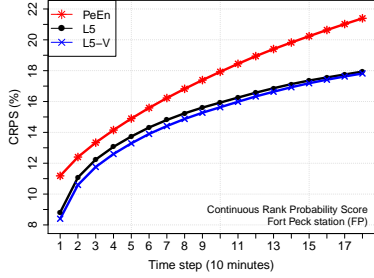
(d) FP site: PINAW metric.



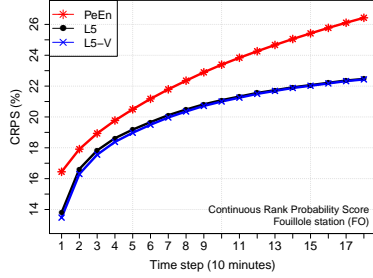
(e) FO site: PINAW metric.



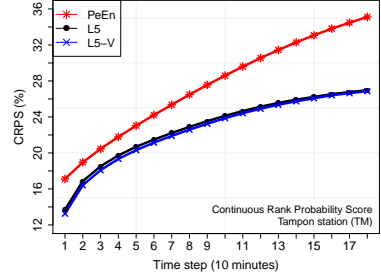
(f) TM site: PINAW metric.



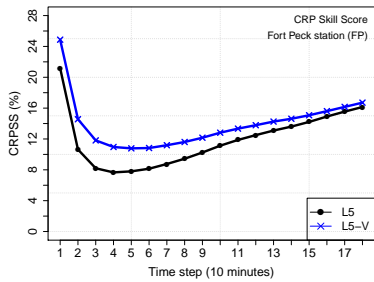
(g) FP site: CRPS.



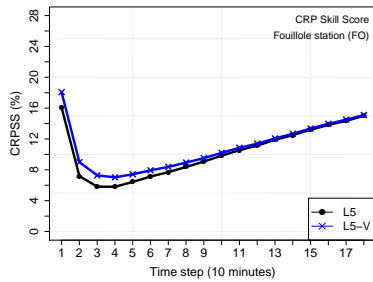
(h) FO site: CRPS.



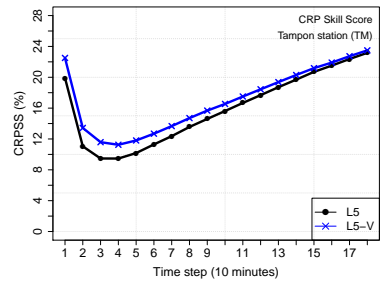
(i) TM site: CRPS.



(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).

647 References

- 648 Aguiar, L. M., Pereira, B., David, M., Díaz, F., and Lauret, P. (2015). Use of satellite data to improve solar radiation forecasting
649 with bayesian artificial neural networks. *Solar Energy*, 122:1309–1324.
- 650 Alessandrini, S., Monache, L. D., Sperati, S., and Cervone, G. (2015). An analog ensemble for short-term probabilistic solar
651 power forecast. *Applied Energy*, 157:95–110.
- 652 Alonso-Suárez, R. (2017). *Estimación del recurso solar en Uruguay mediante imágenes satelitales*. PhD thesis, Facultad de
653 Ingeniería, Universidad de la República. Available at: <https://hdl.handle.net/20.500.12008/20200>.
- 654 Alonso-Suárez, R., Abal, G., Musé, P., and Siri, R. (2014). Satellite-derived solar irradiation map for Uruguay. In *Elsevier*
655 *Energy Procedia*, volume 57, pages 1237–1246.
- 656 Alonso-Suárez, R., Abal, G., Siri, R., and Musé, P. (2012). Brightness-dependent Tarpley model for global solar radiation
657 estimation using GOES satellite images: application to Uruguay. *Solar Energy*, 86(11):3205–3215.
- 658 Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., de Pison, F. M., and Antonanzas-Torres, F. (2016). Review of photovoltaic
659 power forecasting. *Solar Energy*, 136:78–111.
- 660 Antonanzas-Torres, F., Urraca, R., Polo, J., Perpiñán-Lamigueiro, O., and Escobar, R. (2019). Clear sky solar irradiance
661 models: A review of seventy models. *Renewable and Sustainable Energy Reviews*, 107:374–387.
- 662 Bacher, P., Madsen, H., and Nielsen, H. A. (2009). Online short-term solar power forecasting. *Solar Energy*, 83(10):1772–1783.
- 663 Botterud, A. (2017). Chapter 10 - forecasting renewable energy for grid operations. In Jones, L. E., editor, *Renewable Energy*
664 *Integration (Second Edition)*, pages 133–143. Academic Press, Boston, second edition edition.
- 665 Ceballos, J. C., Bottino, M., and de Souza, J. (2004). A simplified physical model for assessing solar radiation over Brazil using
666 GOES 8 visible imagery. *Journal of Geophysical Research: Atmospheres*, 109(D2).
- 667 Chernozhukov, V., Fernández-Val, I., and Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica*,
668 78(3):1093–1125.
- 669 Chu, Y. and Coimbra, C. F. (2017). Short-term probabilistic forecasts for direct normal irradiance. *Renewable Energy*,
670 101:526–536.
- 671 Cros, S., Liandrat, O., Sébastien, N., Schmutz, N., and Voyant, C. (2013). Clear sky models assessment for an operational
672 PV production forecasting solution. In *Proceedings of the 28th European PV Solar Energy Conference and Exhibition*,
673 Villepinte, France.
- 674 Dambreville, R., Blanc, P., Chanussot, J., and Boldo, D. (2014). Very short term forecasting of the global horizontal irradiance
675 using a spatio-temporal autoregressive model. *Renewable Energy*, 72:291–300.
- 676 David, M., Luis, M. A., and Lauret, P. (2018). Comparison of intraday probabilistic forecasting of solar irradiance using only
677 endogenous data. *International Journal of Forecasting*, 34(3):529–547.
- 678 David, M., Ramahatana, F., Trombe, P., and Lauret, P. (2016). Probabilistic forecasting of the solar irradiance with recursive
679 ARMA and GARCH models. *Solar Energy*, 133:55–72.
- 680 Diagne, M., David, M., Lauret, P., Boland, J., and Schmutz, N. (2013). Review of solar irradiance forecasting methods and a
681 proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews*, 27:65–76.
- 682 Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the*
683 *Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268.
- 684 Golestaneh, F., Pinson, P., and Gooi, H. B. (2016). Very short-term nonparametric probabilistic forecasting of renewable
685 energy generation—with application to solar energy. *IEEE Transactions on Power Systems*, 31(5):3850–3863.
- 686 Grantham, A., Gel, Y. R., and Boland, J. (2016). Nonparametric short-term probabilistic forecasting for solar radiation. *Solar*
687 *Energy*, 133:465–475.
- 688 Gueymard, C. A. (2019). Chapter 5 - Clear-sky radiation models and aerosol effects. In Polo, J., Martín-Pomares, L., and
689 Sanfilippo, A., editors, *Solar Resources Mapping: Fundamentals and Applications*, pages 137–182. Springer, Cham.

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

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

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

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

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

700 Ineichen, P. (2016). Validation of models that estimate the clear sky global and beam solar irradiance. *Solar Energy*, 132:332–
701 344.

702 Ineichen, P. and Perez, R. (2002). A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy*,
703 73(3):151–157.

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

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

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

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

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

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

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

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

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

723 Lauret, P., David, M., and Pedro, H. T. C. (2017). Probabilistic solar forecasting using quantile regression models. *Energies*,
724 10(10).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

753 Reikard, G. (2009). Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Solar Energy*,
754 83(3):342–349.

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

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

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

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

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

766 Wu, X. and Sun, F. (2005). Post-launch calibration of GOES imager visible channel using MODIS. In *Proc. SPIE*, volume
767 5882.

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

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

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

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