Probabilistic solar forecasts evaluation
Part I: Ensemble Prediction Systems (EPS)

Mathieu David[1], Josselin Le Gal La Salle[1], Faly Ramahatana-Andriamasomanana[1], Philippe Lauret[1]

[1] PIMENT, Université de La Réunion, Saint-Pierre (Reunion)

Abstract

This pair of articles proposes a framework to perform a comprehensive testing procedure on solar probabilistic forecasts. The evaluation framework is based on graphical diagnosis tools and quantitative scores initially designed by the weather forecasts verification community. To illustrate the application of the proposed verification framework, two sites, which experience very different climatic conditions, have been selected. First, Desert Rock, situated in the continental US, has high occurrence of clear skies. And second, Tampon, situated on the Reunion tropical island, presents highly variable sky conditions.

Part I focuses on the assessment of ensemble forecasts commonly provided by meteorological utilities such as ECMWF or NCEP. The singular nature of this type of probabilistic forecasts requires to carefully indicate the assumptions used to define the associated CDF and to use suitable verification tools.

Keywords: solar forecasts, ensemble prediction systems, evaluation framework, diagnostic tools, scores

1. Introduction

Forecasts of solar energy production are necessary to efficiently integrate solar renewables into grids and also to decrease the associated costs. Indeed, power generation from photovoltaic (PV) or solar thermal plants is highly variable since weather dependent. Therefore, accurate knowledge of the future production of solar renewables is necessary to limit the needs of additional balancing services and potentially storage. Therefore, increasing the value of solar renewables generation through the improvement of solar forecasting models is of paramount importance. This work will use the global horizontal irradiance (GHI) to illustrate the application of the proposed evaluation framework.

Numerous works are dedicated to the development of models that generate deterministic forecasts and the solar forecasting community has already defined a set of metrics, which are commonly used to evaluate their quality (Coimbra et al., 2013, Hoff et al., 2013). However, a forecast is inherently uncertain and in a context of decision-making faced by the grid operator, a point forecast plus an uncertainty (or prediction) interval is a true added-value. Compared to the wind power community, where probabilistic forecasts are developed since many years (Iversen et al., 2015; Jung and Broadwater, 2014; Morales et al., 2014; Pinson et al., 2007), the development of probabilistic solar forecast is relatively new and most of the works on the topic are recent (Alessandrini et al., 2015; Ben Bouallègue, 2015; David et al., 2016; Grantham et al., 2016; Sperati et al., 2016; Zamo et al., 2014). Therefore, a framework to evaluate probabilistic prediction is now required.

Assessing the accuracy of probabilistic forecasts is harder than for deterministic ones (Pinson et al., 2007). Figure 1 gives an example of GHI probabilistic forecasts. Prediction intervals (PIs) enrich the deterministic prediction. The deviation between recorded GHI (black line) and deterministic forecasts (blue dashed line) can easily be assessed by a visual inspection. But it is almost impossible to evaluate the accuracy of the prediction intervals. A comprehensive assessment of probabilistic forecasts needs to use relevant diagnostic tools and scores.

doi:10.18086/swc.2019.43.02 Available at http://proceedings.ises.org

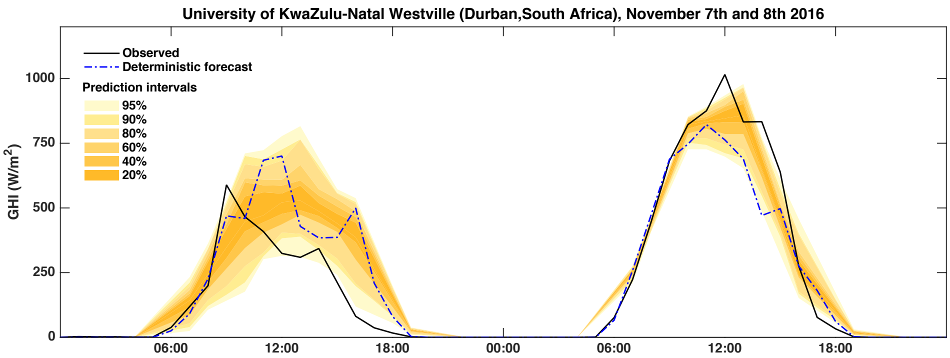

Fig. 1: 2 days ahead of hourly GHI forecasts (ECMWF-EPS) and observations at the University of KuwaZulu-Natal Westville, South Africa (Brooks et al., 2015)

(Murphy, 1993) defines three different characteristics of the goodness of weather forecasts: consistency, quality and value. Among these characteristics, this work concentrates on the assessment of the quality of the models. Several attributes characterize the quality of a probabilistic forecasting system (Wilks, 2009) but two main properties (i.e. reliability and resolution) are used to measure the quality of a forecasting system (Jolliffe and Stephenson, 2003). Reliability or calibration refers to the statistical consistency between the forecasts and the observations. Resolution measures the capacity of a forecasting model to issue forecasts that are case-dependent. The resolution property is commonly not considered by the solar forecasting community. A third widely used attribute, namely sharpness, characterizes how informative the forecasts are. Practically, sharpness refers to the concentration of the predictive distributions and does not provide any indication about the quality of the forecasts (Pinson et al., 2010, Gneiting and Raftery, 2007).

Numerous tools and metrics have already been used to assess the quality of solar probabilistic forecasts (Alessandrini et al., 2015; Chu and Coimbra, 2017; David et al., 2018; Golestaneh et al., 2016; Grantham et al., 2016; Sperati et al., 2016; Verbois et al., 2018; Zamo et al., 2014). Regarding the graphical ones, reliability diagrams and rank histograms, which are commonly used by the weather forecast community, were proposed to visually evaluate the reliability. Several metrics have also been used to assess the quality and properties of solar probabilistic forecasts. The review of the literature shows that the Continuous Rank Probability Score (CRPS) is the main score for both solar and weather forecasting communities. The CRPS assess simultaneously the reliability and resolution. Indeed, the CRPS corresponds to the sum of three components, namely reliability, resolution and uncertainty. This decomposition gives a detailed picture of the performance of the forecasting methods (Hersbach, 2000) and consequently may help in the ranking of the probabilistic forecasts. Other metrics were also proposed. The Prediction Interval Coverage Probability (PICP) and the Prediction Interval Normalized Average Width (PINAW) are respectively used to assess the reliability and the sharpness (Chu and Coimbra, 2017; Khosravi et al., 2013; Lauret et al., 2017). The coverage width-based criterion (CWC) was also proposed to assess the quality of the prediction intervals (Khosravi et al., 2013). Unfortunately, these metrics are not all relevantly or appropriately used. For example, (Lauret et al., 2017) and (David et al., 2018) applied the version originally dedicated to assess ensemble forecasts (Hersbach, 2000) to compute the CRPS of discrete quantile forecasts. Furthermore, as shown by (Pinson and Tastu, 2014), the CWC can lead to, possible misinterpretations. But some works use (Chu et al., 2015; Grantham et al., 2016) and cite (van der Meer et al., 2018; Yang et al., 2018) the CWC as a relevant metric.

That's why we think it's time to take stock on the evaluation tools used to test solar probabilistic forecasts. The objective of this work is to provide the forecasting solar community a comprehensive overview of diagnostic tools and scores that can be used to assess the performance of probabilistic forecasting methods. In particular, the proposed verification framework may help the user to consistently evaluate the quality of the models. In addition, we will propose a measure of resolution, as this attribute is not currently assessed in the literature.

Two types of GHI probabilistic forecasts exist. The first one is the ensemble forecast commonly provided by Ensemble Prediction Systems (EPS) of the Numerical Weather Predictions (NWP) of meteorological utilities. The second one is based on statistical methods and produces a set of quantiles spanning the unit interval. This 1 part of this pair of articles focuses on the assessment of EPS.

2. Assumption underlying the evaluation of EPS

An EPS gives the distribution of the future event as an ensemble of members that are not directly linked to the notion of probabilities. For example, in the case of a NWP, an ensemble forecast corresponds to a perturbed set of forecasts computed by slightly changing the initial conditions of the control run and of the modeling of unresolved phenomena (Leutbecher and Palmer, 2008). This EPS allows representing the uncertainties of the prediction scheme. Nevertheless, ensemble forecasts can be seen as discrete estimates of a CDF when they are sorted in ascending order. In the literature, different ways to associate these sorted members to define cumulative probabilities are proposed. Considering M sorted members of an ensemble $E = (e_1, \ldots, e_M)$ and the corresponding observation x_{obs}, the most common definition states that there is a probability of $1/M$ that the observation falls between two consecutive members e_j and e_j+1 (Anderson, 1996; Hersbach, 2000). If we assign a null probability for future events that fall outside the ensemble (i.e. $x_{obs} < e_1$ or $x_{obs} > e_M$), the predictive distribution can be seen as a piecewise constant function

$$\hat{F}(x) = \sum_{k=1}^{M} \alpha_k \, H(x - e_k). \qquad \text{(eq. 1)}$$

H is the Heaviside function. The weight $\alpha_k = 1/M$ corresponds to the jump of probability that happens when $x = e_k$. Figure 2 (left) shows the classical representation of an ensemble containing 4 members ($M = 4$). Considering a continuous variable, such as solar irradiance, the "scale" shape of this CDF is not realistic. Some works (Bröcker, 2012; Pinson et al., 2010; Roulston and Smith, 2002) proposed alternative definitions to face this issue. For instance, they permit defining a continuous predictive distribution and non-null probabilities outside the ensemble boundaries. We briefly present below two other ways to build a CDF from an ensemble forecast.

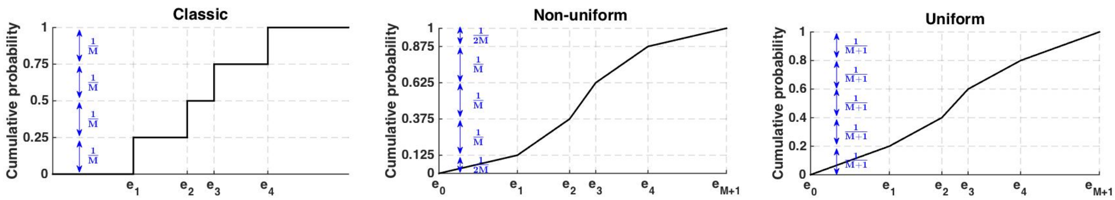

Fig. 2: Classic, non-uniform and uniform definitions of a CDF derived from an ensemble forecast

The first one is proposed by (Bröcker, 2012). A $1/M$ jump of probability separate two consecutive members and a probability mass of $1/2M$ is assigned to the events that fall outside of the ensemble. The resulting CDF is a non-uniform partition of the probability space [0; 1]. Figure 2 (middle) shows an example for an ensemble with 4 members ($M = 4$) and a linear interpolation between the members. The distribution is bounded by e_0 and e_{M+1}. These limits may be arbitrary chosen. For instance, (Roulston and Smith, 2002) set these boundaries using the minimum and the maximum of the climatology.

Regarding the second definition (Bröcker, 2012; Pinson et al., 2010), a probability mass of $1/(M+1)$ is assigned between two consecutive members and for the events that fall outside of the ensemble. This representation results in an uniform spacing of the cumulative probabilities. Figure 2 (right) shows the shape of this uniform CDF for an ensemble with 4 members and a linear interpolation between the members. Once again, the boundaries of the CDF, e_0 and e_{M+1}, have to be arbitrarily chosen.

The choice of the definition of the CDF will obviously affect the generation and the analysis of the evaluation tools. The classical definition generates a piecewise constant CDF using only the available members. The uniform and the non-uniform definitions require to arbitrary choose the boundaries of the CDF (i.e. e_0 and e_{M+1}). Currently, the classical definition is the most used to generate the evaluation tools.

3. Verification tools

Evaluation of probabilistic solar forecasts has to be done with tools and metrics that are relevant related to their nature. Considering EPS and its particularity to be defined by an ensemble of members, the classical definition of the corresponding CDF seems to be the most appropriate. Indeed, no additional assumption is required, such as setting the boundaries of the distribution. Furthermore, the weather forecasting community has already developed a set of tools and metrics that are suitable with this CDF definition. Several attributes characterize probabilistic forecasts and they do not have the same importance when evaluating their quality. The reliability is a primary requirement as non reliable forecasts would lead to a systematic bias in subsequent decision-making processes (Pinson et al., 2007). Thus, reliability is the first propriety to check.

3.1. Graphical tools

Several visual tools are available to assess the reliability. The most suitable for EPS is the rank histogram. Rank histograms are useful for determining the statistical consistency of the ensemble, that is, if the observation being predicted looks statistically just like another member of the forecast ensemble (Wilks, 2009). A necessary condition for ensemble consistency is an appropriate degree of ensemble dispersion leading to a flat rank histogram. Considering a finite number of observation/forecast pairs, a consistency band, as proposed by (Bröcker and Smith, 2007), defines the area where the ranks must fall. Regarding the evaluation of the resolution, no graphical tool exists. Finally, sharpness diagrams are useful to visually assess the concentration of the prediction intervals. One of the most used by the solar community is the Prediction Interval Average Width (PINAW) diagram (Chu and Coimbra, 2017; Khosravi et al., 2013; Lauret et al., 2017). As the boundaries of the PIs depend on the assumption done to define the CDF from the EPS, it is mandatory to clearly indicate which one was used (e.g. uniform or non-uniform spacing) when presenting a PINAW diagram.

3.2. Scores

Numerical scores provide summary measures for the evaluation of the quality of probabilistic forecasts (Gneiting and Raftery, 2007). Scores may help to rank competing probabilistic models. Several scores are available in the literature (e.g. Quantile Score, Interval Score and Ignorance Score). Among others, the Continuous Rank Probability Score (CRPS) (Hersbach, 2000) offers appealing characteristic: propriety, normalization, similarity with the Mean Absolute Error (MAE) and decomposition in reliability (REL), resolution (RES) and uncertainty (UNC) (eq. 1).

$$CRPS = \sum_{i=1}^{N} \int_{-\infty}^{+\infty} \left[\hat{F}_{fcst}^{i}(x) - F_{obs}^{i}(x)\right]^2 dx = REL + UNC - REL \qquad \text{(eq. 2)}$$

3.1 Proposed evaluation framework

Considering ensemble forecasts, we propose to use the rank histogram including consistency bars and the CRPS as defined by (Hersbach, 2000) to respectively qualify and quantify the performances of the EPS. Indeed, these two tools doe not require additional assumptions (i.e. to define the nature of the distribution and its boundaries) and they are already widely used. PINAW diagrams can complement the characterization of the forecasting methods. However, PINAW diagrams must be interpreted with care because they are relevant only if the associated forecasts are reliable. Finally, when PINAW diagrams are derived from ensemble forecasts, it is important to clearly indicate the assumption done to obtain the prediction intervals (e.g. classical, uniform or non-uniform spacing).

4. Application

To illustrate the use of the different evaluation tools, we will present in this section examples relative to two locations with different sky conditions. The first site, Desert Rock (USA), has an arid climate with a very sunny and stable sky. The second site, Tampon (Reunion), is located in a tropical island and experiences a very variable sky. The experimental dataset corresponds to two consecutive years of recorded and forecasted data of global horizontal irradiance (GHI), the first year of data (2012) as training set and the second year of data (2013) as testing set. The day-ahead ensemble predictions were provided by the Integrated Forecasting System (IFS) of the European Centre of Medium-Range Weather Forecasts (ECMWF-EPS). They consist in 50 perturbed members. The temporal resolution is of 3 hours and the spatial resolution is of 0.2° in both longitude and latitude. In addition, we also propose a post-processed version of the original ECMWF-EPS forecasts calibrated with the Variance Deficit (VD) method (Sperati et al., 2016).

4.1. Reliability assessment

As discussed above, the first requirement for a probabilistic forecast is the reliability. The rank histogram is a suitable tool to check this property of EPS because it permits to visually assess if the members of the ensemble are indistinguishable from the observations. Figures 3 and 4 shows the rank histograms of the ECMWF-EPS forecasts with and without calibration. The rank histogram of raw the ECMWF-EPS forecasts exhibits a U-shape indicating a lack of spread of the members. This is a common issue of EPS provided by NWP. The calibration spreads the members and the resulting rank histograms for both sites are flatter. However, a dissymmetry appears with the lower ranks (left side) more populated than the higher ones (right side). This specific shape indicates that the calibration create a bias in the forecasts. Finally, even if the rank histograms after the calibration are better than the ones corresponding to the original ECMWF-EPS, the post-processing does not produce a reliable EPS as a large number of ranks remain out of the consistency band.

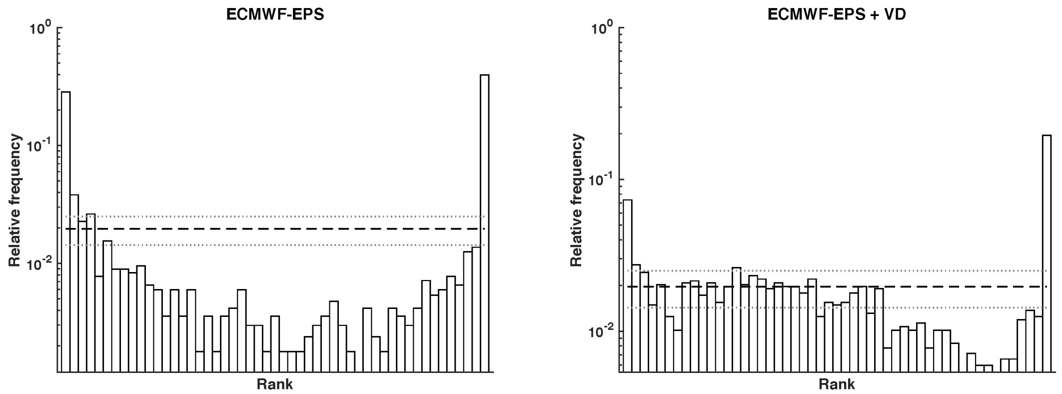

Fig. 3: Rank histograms of raw ECMWF-EPS (left) and ECMWF-EPS calibrated with variance deficit (right) for Desert Rock

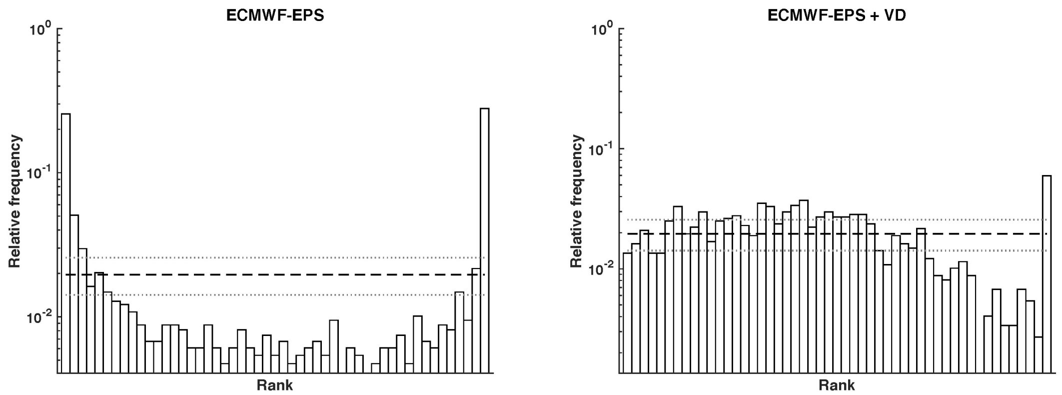

Fig. 4: Rank histograms of raw ECMWF-EPS (left) and ECMWF-EPS calibrated with variance deficit (right) for Tampon

4.2. Sharpness assessment

As the forecasts are not reliable, there is no need to lead further investigations about the sharpness of the prediction intervals as it could lead to a misinterpretation. However, we will evaluate the sharpness of the considered forecasts to better illustrate this issue. Figure 5 shows sharpness diagrams for coverage rates ranging from 0% to 100%, for the two sites and for the two considered ensemble forecasts. To compute the mean size of the central prediction intervals, we assume an uniform spacing of the quantiles (see section 2). PIs of original ECMWF-EPS forecasts are narrower than the calibrated ones. This is the consequence of the under-dispersion and therefore of the low reliability of the ECMWF-EPS forecasts. So, in this case, even if narrow PIs are preferred, sharpness diagrams should not be used as criteria to assess the quality of the forecasts.

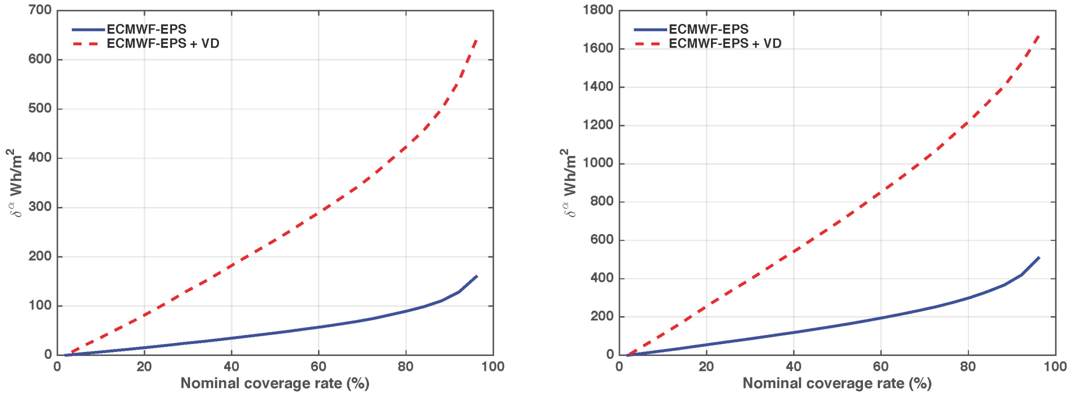

Fig. 5: PINAW diagrams of ECMWF-EPS and ECMWF-EPS + VD for Desert Rock (left) and for Tampon (right)

4.3. CRPS

To complete this analysis, Table 1 gives the CRPS and its decomposition. The calibration decreases the CRPS only for Tampon. The decomposition highlights that the VD method improves the reliability for both sites but decreases the resolution. But in the case of Desert Rock the increase in reliability, resulting from the calibration, does not counter-balance the reduction in resolution. Figure 6 illustrates for a specific day why the calibration results in a worst CRPS for Desert Rock. Indeed, the initial ECMWF-EPS forecast (blue line) already contains the observation (black line) and the associated CDF is very sharp. So, the ECMWF-EPS forecast is relatively good and results in a low CRPS. The VD method (red line) spreads the members and the resulting CDF covers a wider range of possible irradiances. As the CRPS takes into account the distance between the observation and the associated CDF range, the CRPS increases significantly. This example corresponds to a clear sky that has been forecasted and occurred. Numerous clear sky occurrences are forecasted and observed at Desert Rock. As a consequence, the results presented in the example can be extended to the whole year. So, the VD method spreads blindly the ECMWF forecasts, even when it is not necessary.

Tab. 1: CRPS normalized by the mean irradiance and its decomposition

	CRPS (%)	CRPS decomposition (%)		
		Reliability	Resolution	Uncertainty
		Desert Rock		
ECMWF-EPS	6.97	1.77	37.9	43.1
ECMWF-EPS + VD	7.37	0.97	36.7	43.1
		Tampon		
ECMWF-EPS	25.1	6.03	23.5	42.6
ECMWF-EPS + VD	23.1	2.41	21.9	42.6

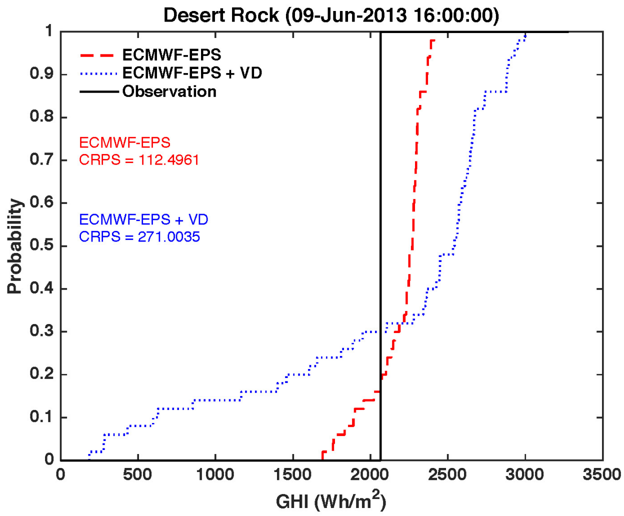

Fig. 6: CDF of the observation and of calibrated (ECMWF-EPS+VD) and the raw ECMWF-EPS forecasts for the site of Desert Rock June 9th 2013 at 16:00

5. Conclusion

In this work, a framework for evaluating solar probabilistic forecasts is proposed. This framework is based on visual tools and scoring rules originally designed by the weather forecast verification community. However, no methodology to use them relevantly exists in the realm of solar forecasting. When dealing with ensemble forecasts, dedicated verification tools, such as rank histograms and the CRPS as proposed by Hersbach, can be used without any additional assumption. Indeed, they assume a classical definition of the underlying CDF and one does not require defining the CDF boundaries. However, care must be taken while deriving quantiles, prediction intervals and associated metrics (i.e. QS, PINAW, etc.) from ensembles. As several possibilities are available, it is important to clearly state which one is used (e.g. uniform or non-uniform spacing). The authors have a preference for the uniform spacing because it performs quantile in a similar manner to discrete predictive distributions. This work focused on the forecasting of the solar irradiance. However, the proposed methodology and associated tools can be extended to the energy generation of solar renewables.

6. References

Alessandrini, S., Delle Monache, L., Sperati, S., Cervone, G., 2015. An analog ensemble for short-term probabilistic solar power forecast. Appl. Energy 157, 95–110. https://doi.org/10.1016/j.apenergy.2015.08.011

Anderson, J.L., 1996. A Method for Producing and Evaluating Probabilistic Forecasts from Ensemble Model Integrations. J. Clim. 9, 1518–1530. https://doi.org/10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2

Ben Bouallègue, Z., 2015. Statistical postprocessing of ensemble global radiation forecasts with penalized quantile regression. Meteorol. Z. https://doi.org/10.1127/metz/2016/0748

Bröcker, J., 2012. Evaluating raw ensembles with the continuous ranked probability score. Q. J. R. Meteorol. Soc. 138, 1611–1617. https://doi.org/10.1002/qj.1891

Bröcker, J., Smith, L.A., 2007. Increasing the Reliability of Reliability Diagrams. Weather Forecast. 22, 651–661. https://doi.org/10.1175/WAF993.1

Brooks, M.J., Du Clou, S., Niekerk, V., L, W., Gauche, P., Leonard, C., Mouzouris, M.J., Meyer, R., Van der Westhuizen, N., Dyk, V., E, E., Vorster, F.J., 2015. SAURAN : a new resource for solar radiometric data in Southern Africa. J. Energy South. Afr. 26.

Chu, Y., Coimbra, C.F.M., 2017. Short-term probabilistic forecasts for Direct Normal Irradiance. Renew. Energy 101, 526–536. https://doi.org/10.1016/j.renene.2016.09.012

Chu, Y., Li, M., Pedro, H.T.C., Coimbra, C.F.M., 2015. Real-time prediction intervals for intra-hour DNI forecasts. Renew. Energy 83, 234–244. https://doi.org/10.1016/j.renene.2015.04.022

Coimbra, C.F.M., Kleissl, J., Marquez, R., 2013. Overview of Solar-Forecasting Methods and a Metric for Accuracy Evaluation, in: Solar Energy Forecasting and Resource Assessment. Elsevier, pp. 171–194. https://doi.org/10.1016/B978-0-12-397177-7.00008-5

David, M., Luis, M.A., Lauret, P., 2018. Comparison of intraday probabilistic forecasting of solar irradiance using only endogenous data. Int. J. Forecast. 34, 529–547. https://doi.org/10.1016/j.ijforecast.2018.02.003

David, M., Ramahatana, F., Trombe, P.J., Lauret, P., 2016. Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models. Sol. Energy 133, 55–72. https://doi.org/10.1016/j.solener.2016.03.064

Gneiting, T., Raftery, A.E., 2007. Strictly Proper Scoring Rules, Prediction, and Estimation. J. Am. Stat. Assoc. 102, 359–378. https://doi.org/10.1198/016214506000001437

Golestaneh, F., Pinson, P., Gooi, H.B., 2016. Very Short-Term Nonparametric Probabilistic Forecasting of Renewable Energy Generation— With Application to Solar Energy. IEEE Trans. Power Syst. 31, 3850–3863. https://doi.org/10.1109/TPWRS.2015.2502423

Grantham, A., Gel, Y.R., Boland, J., 2016. Nonparametric short-term probabilistic forecasting for solar radiation. Sol. Energy 133, 465–475. https://doi.org/10.1016/j.solener.2016.04.011

Hersbach, H., 2000. Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. Weather Forecast. 15, 559–570. https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2

Hoff, T.E., Perez, R., Kleissl, J., Renne, D., Stein, J., 2013. Reporting of irradiance modeling relative prediction errors: Reporting of irradiance modeling relative prediction errors. Prog. Photovolt. Res. Appl. 21, 1514–1519. https://doi.org/10.1002/pip.2225

Iversen, E.B., Morales, J.M., Møller, J.K., Madsen, H., 2015. Short-term probabilistic forecasting of wind speed using stochastic differential equations. Int. J. Forecast. https://doi.org/10.1016/j.ijforecast.2015.03.001

Jolliffe, I.T., Stephenson, D.B., 2003. Forecast Verification. A practitioner's guide in atmospheric science. Wiley, Chichester, England.

Jung, J., Broadwater, R.P., 2014. Current status and future advances for wind speed and power forecasting. Renew. Sustain. Energy Rev. 31, 762–777. https://doi.org/10.1016/j.rser.2013.12.054

Khosravi, A., Nahavandi, S., Creighton, D., 2013. Prediction Intervals for Short-Term Wind Farm Power Generation Forecasts. IEEE Trans. Sustain. Energy 4, 602–610. https://doi.org/10.1109/TSTE.2012.2232944

Lauret, P., David, M., Pedro, H., 2017. Probabilistic Solar Forecasting Using Quantile Regression Models. Energies 10, 1591. https://doi.org/10.3390/en10101591

Leutbecher, M., Palmer, T.N., 2008. Ensemble forecasting. J. Comput. Phys. 227, 3515–3539. https://doi.org/10.1016/j.jcp.2007.02.014

Morales, J.M., Conejo, A.J., Madsen, H., Pinson, P., Zugno, M., 2014. Integrating Renewables in Electricity Markets, International Series in Operations Research & Management Science. Springer US, Boston, MA. https://doi.org/10.1007/978-1-4614-9411-9

Murphy, A.H., 1993. What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. Weather Forecast. 8, 281–293. https://doi.org/10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2

Pinson, P., McSharry, P., Madsen, H., 2010. Reliability diagrams for non-parametric density forecasts of continuous variables: Accounting for serial correlation. Q. J. R. Meteorol. Soc. 136, 77–90. https://doi.org/10.1002/qj.559

Pinson, P., Nielsen, H.Aa., Møller, J.K., Madsen, H., Kariniotakis, G.N., 2007. Non-parametric probabilistic forecasts of wind power: required properties and evaluation. Wind Energy 10, 497–516. https://doi.org/10.1002/we.230

Pinson, P., Tastu, J., 2014. Discussion of "Prediction Intervals for Short-Term Wind Farm Generation Forecasts" and "Combined Nonparametric Prediction Intervals for Wind Power Generation." IEEE Trans. Sustain. Energy 5, 1019–1020. https://doi.org/10.1109/TSTE.2014.2323851

Roulston, M.S., Smith, L.A., 2002. Evaluating Probabilistic Forecasts Using Information Theory. Mon. Weather Rev. 130, 1653–1660.

Sperati, S., Alessandrini, S., Delle Monache, L., 2016. An application of the ECMWF Ensemble Prediction System for short-term solar power forecasting. Sol. Energy 133, 437–450. https://doi.org/10.1016/j.solener.2016.04.016

van der Meer, D.W., Widén, J., Munkhammar, J., 2018. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. Renew. Sustain. Energy Rev. 81, 1484–1512. https://doi.org/10.1016/j.rser.2017.05.212

Verbois, H., Rusydi, A., Thiery, A., 2018. Probabilistic forecasting of day-ahead solar irradiance using quantile gradient boosting. Sol. Energy 173, 313–327. https://doi.org/10.1016/j.solener.2018.07.071

Wilks, D.S., 2009. Statistical methods in the atmospheric sciences, 2. ed., [Nachdr.]. ed, International geophysics series. Elsevier [u.a.], Amsterdam.

Yang, D., Kleissl, J., Gueymard, C.A., Pedro, H.T.C., Coimbra, C.F.M., 2018. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. Sol. Energy 168, 60–101. https://doi.org/10.1016/j.solener.2017.11.023

Zamo, M., Mestre, O., Arbogast, P., Pannekoucke, O., 2014. A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. Part II: Probabilistic forecast of daily production. Sol. Energy 105, 804–816. https://doi.org/10.1016/j.solener.2014.03.026