Complex-Valued Time Series Based Solar Irradiance Forecast

Cyril Voyant, Philippe Lauret, Gilles Notton, Jean-Laurent Duchaud, Luis Garcia-Gutierrez, Ghjuvan Antone Faggianelli

Complex-Valued Time Series Based Solar Irradiance Forecast

Cyril Voyant,[1, a)] Philippe Lauret,[2] Gilles Notton,[1] Jean-Laurent Duchaud,[1] Luis Garcia-Gutierrez,[3] and Ghjuvan Antone Faggianelli[1]

[1)] *University of Corsica, SPE Laboratory-Georges Peri'centre - UMR6134, Ajaccio (France)*

[2)] *University of Reunion, PIMENT Laboratory, Saint-Pierre (France)*

[3)] *University of Lorraine, LMOPS Laboratory , Metz (France)*

(Dated: 3 December 2022)

A new method for short-term probabilistic forecasting of global solar irradiance from complex-valued time series is explored. Measurement defines the real part of the time series while the estimate of the volatility is the imaginary part. A complex autoregressive model (capable to capture quick fluctuations) is then applied with data gathered on Corsica island (France). Results show that even if this approach is easy to implement and requires very little resource and data, both deterministic and probabilistic forecasts generated by this model are in agreement with experimental data (root mean square error ranging from 0.196 to 0.325 considering all studied horizons). In addition, it exhibits sometimes a better accuracy than classical models such as the Gaussian process, bootstrap methodology, or even more sophisticated models such as quantile regression. Many studies and many fields of physics could benefit from this methodology and from the many models that could result from it.

Keywords: Probabilistic; Forecasting; Univariate; Interval

I. INTRODUCTION

Nowadays, it is acknowledged that to limit the impact of the random and variable nature of the solar resource and thus to facilitate its integration, developments are necessary. They concern the energy storage means, the smart grid energy management, and the forecasting methods for both power generation and user's consumption[1]. The topic of this paper falls within the development of a forecasting method for Photovoltaic (PV) power generation and concerns nowcasting. Numerous machine learning methods benchmarks have been published in the literature and most of them compare the models in terms of accuracy[2,3] with regard to time horizons. These methods capture often the general trend and fail to capture the quick fluctuations, while advanced nonparametric approaches that attempt to do so may be prone to overfitting or too complicated for practical applications (lack of data, acquisition system failures, process execution time, etc.). Many grid managers prefer to use the simplest and the most robust ones, sometimes at the expense of their performance[4]. In agreement with the "No Free Lunch theorem" of Wolpert and Macready [5], which explains that no learning algorithm is the most suitable in all scenarios[6], we propose a new data mining based non-parametric probabilistic method, easy to implement, with good accuracy and based on a new theoretical basis integrating trend but also rapid fluctuations predictions. A univariate methodology based on a complex number generation is applied to predict simultaneously the hourly solar global horizontal irradiance (GHI) and an estimate of its volatility from previous ground measurements.

[a)] Electronic mail: voyant_c@univ-corse.fr

II. DATA

As detailed by Yang [7], an adequate analysis and modeling are essential to issue good forecasts when a time series exhibits seasonal or cyclic behavior as it is the case for GHI with its two seasonal periods (yearly and diurnal cycles). Since 1961 and the first works about stationary processes with a finite second-moment[8] and periodic correlation (or covariance)[9], the scientists know that it is important to pay attention to trends when time series is used. Box and Jenkins' first formalism[10] clarified this aspect by proposing a decomposition, especially when seasonality is easily quantifiable. Usually, a multiplicative scheme is chosen, and a classical ratio between GHI in clear sky condition (denoted GHI_{CS}) and GHI is operated. This parameter (considered "sufficiently" stationary or at least locally stationary as demonstrated by Yang et al[7]) results in a normalized quantity (theoretically comprised between 0 and 1 as long as the over-irradiance phenomenon is neglected) known as $\kappa(t)$ the clear-sky index[11],

$$\kappa(t) = GHI(t)/GHI_{CS}(t) \in [0,1] \quad (1)$$

Thus, most solar forecasters build their forecasting models on κ, rather than on GHI itself. As a part of this study, several rules and explanations must be given to improve the objectivity of conclusions:

- GHI time series is measured in Ajaccio (Corsica, France, 41.92N-8.74E, 5m above sea level) endowed with a warm Mediterranean climate (Csa Köppen climate classified) and yearly solar irradiation of 1642 kWh.m^{-2},
- Models are evaluated during only daytime irradiance values, filtering the checked data (less than 1% are left according to quality control[12]) on solar

zenith angle ($GHI = \emptyset$ if $\theta_Z > 85°$),

↬ GHI_{CS} is computed with the `Solis` model which proposes an atmospheric scheme based on radiative transfer calculations and the Lambert-Beer relation[13].

This paper is dedicated to volatility prediction which is used to generate GHI prediction intervals with respect to the prediction horizon from $1h$ to $6h$ with $1h$ time granularity (training during the years 2008-2017 and testing during the year 2018). Several methods such as autoregressive conditional heteroskedasticity ($ARCH$) models are devoted to this task (volatility modeling) and were extensively studied in econometrics. However, concerning the GHI prediction and its applications in energy management for PV systems, this kind of method has never been used, probably due to its complexity[14], the restrictive assumptions[15] or the quality of its results which seemed even so promising[16]. An important conclusion of Dimson and Marsh [14] concerning the $ARCH$ family predictors is another form of the Occam's razor principle and implies "that for those who are interested in forecasts with reasonable predictive accuracy, the best forecasting models might well be the simplest ones".

III. METHODOLOGY

The method exposed in this paper concerns a new formalism for:

↬ The prediction of the conditional volatility using parameters like the *return* and its *standard deviation* (see definition in Eq.2),

↬ The generation of GHI prediction intervals.

From the computed κ time series (Eq.1), another series reporting on its intrinsic variability (or volatility $\sigma_\tau(t)$) and highlighting the concept of predictive risk is built. To this end, we suggest to use the standard deviation of the κ return ($r(t) = \kappa(t) - \kappa(t-1)$) computed over the τ-sliding windows ($\tau \in \mathbb{N}_{>1}$) (Eq.2),

$$\sigma_\tau(t) = \sqrt{\frac{1}{\tau}\sum_{i=0}^{\tau-1}\left(r(t-i) - \frac{1}{\tau}\sum_{n=0}^{\tau-1} r(t-n)\right)^2} \qquad (2)$$

where $\tau = 30$, because for Ajaccio a $30h$ window provides the best results. In the literature, other definitions can be found for the volatility[17,18] using in particular the *logarithm* or the *absolute-value* norm. However, here, the given definition yields the best results and constitutes the simplest way to establish the volatility. Fig.1 shows that the trend of the centered σ_τ (i.e. volatility minus its mean) distribution for Ajaccio can be considered as normal shape with a slight platykurtic tendency (confirmed with the Jarques-Bera test at the 10% significance level). Rather than working separately on κ and σ_τ, we propose

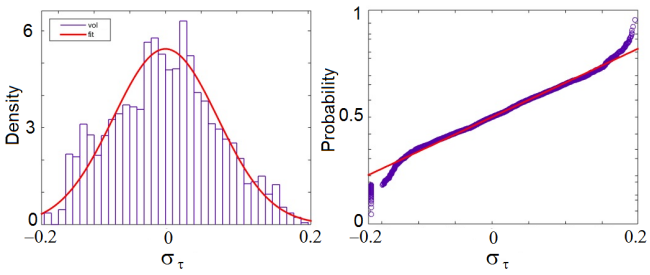

FIG. 1. Graphical method for comparing the centered σ_τ probability distribution in comparison with a normal distribution (probability density function (PDF) and probability plot).

to build $z = \{\kappa(t) + j\sigma_\tau(t), t \in \mathbb{Z}, j^2 = -1\}$, a scalar complex-valued time series and to model this discrete stochastic process with an autoregressive process of order p ($AR(p)$[10]). This model involves random variables defined on the same sample and event spaces and with the same probability measure (that makes it possible to define distribution function F). This method is to be compared to that exposed by Ivan Svetunkov concerning the logic of Brown's exponential smoothing methods and the complex-valued time series used to forecast two-time series simultaneously[19], with the difference that there is no volatility issue. From now on, only mean-centered variables will be considered, but will not be introduced in the following equations for readability reasons.

The complex-valued transform replaces a system of equations related to the prediction of κ and its volatility σ_τ (Eq.3a with $\widehat{(.)}$ for predicted values) by a single regression equation (Eq.3b, the proof is obvious setting $\boldsymbol{\omega} = \boldsymbol{\xi} + j\boldsymbol{\zeta}$, $\boldsymbol{\omega} \in \mathbb{C}$ and $\boldsymbol{\xi}, \boldsymbol{\zeta} \in \mathbb{R}$).

$$\begin{cases} \widehat{\kappa}(t+1) = \sum_{i=0}^{p-1} \kappa(t-i)\xi_i - \sum_{i=0}^{p-1} \sigma_\tau(t-i)\zeta_i \\ \widehat{\sigma}_\tau(t+1) = \sum_{i=0}^{p-1} \kappa(t-i)\zeta_i + \sum_{i=0}^{p-1} \sigma_\tau(t-i)\xi_i \end{cases} \qquad (3a)$$

$$\widehat{z}(t+1) = \sum_{i=0}^{p-1} z(t-i)\omega_i \qquad (3b)$$

Before using an autoregressive (AR) model, it is important to deal with the model identification (the choice of the parameter or order p in Eq.3b); a classical tool widely studied in regression analysis is employed. It handles with the interpretation of the partial autocorrelation factor (β[10]) according to real and imaginary parts of z (respectively $\Re(z)$ and $\Im(z)$) as described in,

$$\exists\, p \mid \beta(t, t-p) \neq 0 \text{ and } \beta(t, t-p-1) = 0\ \forall\{t > p\} \in \mathbb{Z}^{20,21} \qquad (4)$$

where $p_\Re$, $p_\Im$ and p denote AR orders having connections

with $\Re(z)$, $\Im(z)$ and z. We are setting $p = \max(p_{\Re}, p_{\Im})$ to make the problem easier. It is better to benefit from an excess than from a lack of information while referring to the bias-variance trade-off and being aware that the number of inputs should not be too large ($p \in [2, 6]$ for all the horizons concerning the studied site). By setting this rule, we have neglected the particular algebra imposed by the use of complex numbers, but there is, to our knowledge, no other way of doing it that would be as simple. Note that dealing with deseasonalized series κ, p does not have to be greater than $24h$ to capture the daily seasonal cycle. There are other identification methods, as for example the complex autocorrelation factor[22], however, it seems that this method has a worse ratio complexity-efficiency. The next step is the model estimation ($\boldsymbol{\omega}$) by transposing what has been done for many years in the real-valued case (least square optimization) to the complex-valued case. Considering an input matrix $\mathbf{I} \in \mathbb{C}^{D\times p}$ (Eq.5) and an output column vector $\mathbf{o} \in \mathbb{C}^{D\times 1}$ (Eq.6), the solution of the $AR(p)$ least squares problem consists in determining unknown parameters ($\boldsymbol{\omega} \in \mathbb{C}^{p\times 1}$ in Eq.6). By the way, the problem, already raised and well detailed in the paper of Adrian et al.[23] for spatial data-based model is resumed as in the classical real-valued case by $\mathbf{I}\boldsymbol{\omega} = \mathbf{o}$ and can be solved from the formulation of the mean square error estimation $\mathbb{E}[\mathbf{e}^2] = \|\mathbf{I}\boldsymbol{\omega} - \mathbf{o}\|^2$[24]. Note that the authors of this paper do not venture to state that the least squares method provides the best solution to the problem. To be able to assert it, one has to prove that there is equivalence with the maximum likelihood and have to formulate hypotheses on the complex residual distribution.

$$\mathbf{I} = \begin{pmatrix} z(t-1) & z(t-2) & \cdots & z(t-p) \\ z(t-2) & z(t-3) & \cdots & z(t-p-1) \\ \vdots & \vdots & & \vdots \\ z(t-D) & z(t-D-1) & \cdots & z(t-D-p+1) \end{pmatrix}_{D\times p} \tag{5}$$

$$\boldsymbol{\omega} = \big(\omega_1, \ldots, \omega_p\big)'_{p\times 1}, \mathbf{o} = \big(z(t), \ldots, z(t-D+1)\big)'_{D\times 1} \tag{6}$$

The complex-valued case differs from the real-valued one, replacing the L_2-norm by the Frobenius norm introducing the Frobenius inner product[25] on $\mathbb{C}^D$ ($\mathbb{E}[\mathbf{e}^2] =< (\mathbf{I}\boldsymbol{\omega} - \mathbf{o}), (\mathbf{I}\omega - \mathbf{o}) >_F$). Classically, the minimum of the squared expected value ($\text{argmin}(\mathbb{E}[\mathbf{e}^2]) := \{\boldsymbol{\omega} \in \mathbb{C}^p | \forall \boldsymbol{\omega}^* \in \mathbb{C}^p : \mathbb{E}[\mathbf{e}^2(\boldsymbol{\omega}^*)] \geq \mathbb{E}[\mathbf{e}^2(\boldsymbol{\omega})]$) is carried out computing its differentiating (Eq.7) and by letting $\partial\mathbb{E}[\mathbf{e}^2]/\partial(\boldsymbol{\omega}^H) = 0$ where $(.)^H$ defines conjugate transpose.

$$\begin{aligned} \frac{\partial\mathbb{E}[\mathbf{e}^2]}{\partial(\boldsymbol{\omega}^H)} &= \frac{\partial(I\boldsymbol{\omega} - \mathbf{o})^H(I\boldsymbol{\omega} - \mathbf{o})}{\partial(\boldsymbol{\omega}^H)} \\ &= \frac{\partial(\boldsymbol{\omega}^H I^H - \mathbf{o}^H)(I\boldsymbol{\omega} - \mathbf{o})}{\partial(\boldsymbol{\omega}^H)} \end{aligned} \tag{7}$$

It must be emphasised that $\partial\mathbb{E}[\mathbf{e}^2]/\partial(\boldsymbol{\omega}^H)$ is the complex conjugate transpose of $\partial\mathbb{E}[\mathbf{e}^2]/\partial\boldsymbol{\omega}$, thus, setting one to zero also sets the other to zero. The normal equation, in this complex-valued case, becomes,

$$\mathbf{I}^H(\mathbf{I}\boldsymbol{\omega} - \mathbf{o}) = 0 \tag{8}$$

Furthermore, the solution of this matrix equation corresponds to a regression coefficients[26] estimated by $\tilde{\boldsymbol{\omega}} = (\mathbf{I}^H\mathbf{I})^{-1}\mathbf{I}^H\mathbf{o}$. Therefore, differentiating by a complex-valued vector is an abstract concept, but it yields the same set of equations as differentiating separately each scalar component (real and imaginary) and is a more concise form[27]. On top of that, to improve the condition number of the problem, one can introduce a constrained minimization with $\|\boldsymbol{\omega}\|^2 < r(\lambda)$ where r is a bijective function and λ is the Lagrange multiplier of the constraint ($\mathbf{I}^H\mathbf{I} + \lambda^H\lambda$) defined positive and so invertible. This approach denoted `Ridge` approach[28] can also be used in the complex-value case, in the form $\tilde{\boldsymbol{\omega}}_\lambda = (\mathbf{I}^H\mathbf{I} + \lambda\mathbb{1})^{-1}\mathbf{I}^H\mathbf{o}$. A machine learning-like approach consists in performing cross-validation and selecting the λ value that minimizes the out-sample sum of squared residuals; in our experimental setup, $\lambda = 3.74$ is the best choice (i.e. inducing the lowest prediction errors). Now the identification and optimization problems have been analyzed (the outcome of the experiment is detailed in Fig.2), it is required to theoretically validate the use of predictions of both the GHI and its volatility in the case of the probabilistic forecasting[29,30]. The goal is to show that the volatility detailed previously could be used to capture the idea of unpredictable and quick fluctuations. Thereby, considering the ℓ-step head prediction, it would be possible given a well-chosen $\mu_{t+\ell}$ parameter, to bound the prediction considering that $GHI(t+\ell)$ measurement is included in the interval Λ satisfying the Eq.9. The following is dedicated to answering the question: can we theoretically explain $\mu_{t+\ell}$?

$$\begin{aligned} \Lambda = [&\widehat{GHI}(t+\ell) - \mu_{t+\ell}GHI_{CS}(t+\ell)\widehat{\sigma}_\tau(t+\ell), \\ &\widehat{GHI}(t+\ell) + \mu_{t+\ell}GHI_{CS}(t+\ell)\widehat{\sigma}_\tau(t+\ell)] \end{aligned} \tag{9}$$

First of all, it is important to consider this study within non-standard analysis framework with S-integrable time series (additive decomposition[31]) and by referring to the Cartier-Perrin theorem[32], as suggested in many papers from Fliess[33,34]. In this context, we can explore the fact that our prediction method (and more generally all the machine learning approaches) only predicts the trend of the GHI but certainly not fast fluctuations. From Eq.2, it is conceivable to interpret the volatility thanks to the Backshift operator $\boldsymbol{B}$; it holds,

$$\begin{aligned} \sigma^2_\tau(t) &= \mathbb{E}[r^2] - \mathbb{E}[r]^2 \simeq \mathbb{E}[r^2] = \mathbb{E}[(\kappa - \boldsymbol{B}^1\kappa)^2] \\ &= \frac{1}{\tau}\sum_{i=0}^{\tau-1}\Big(\kappa(t-i) - \kappa(t-i-1)\Big)^2 \end{aligned} \tag{10}$$

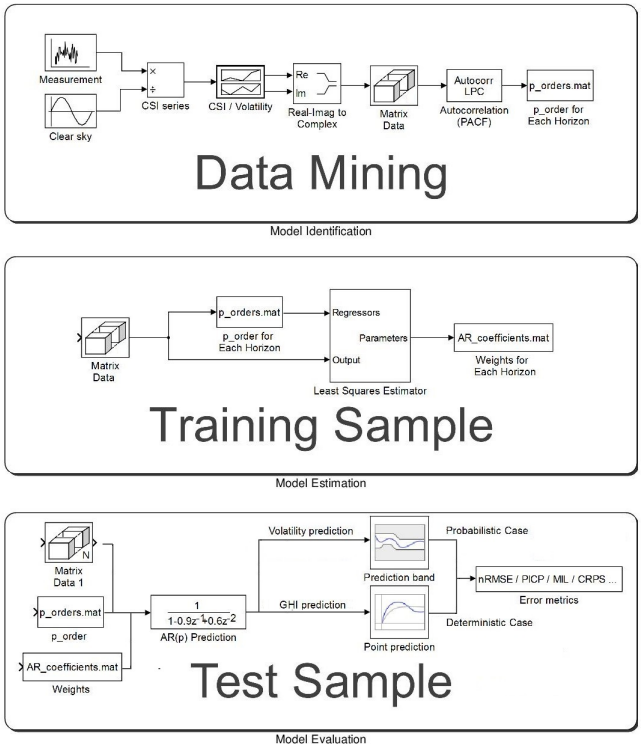

FIG. 2. Conceptual diagram of the experiment.

where over a sufficiently large interval, $\kappa(t)$ oscillates around a constant mean value, making the average of the return close to 0 ($\mathbb{E}[r]^2 \underset{\tau\to+\infty}{\longrightarrow} 0$) and $\boldsymbol{B}^{\ell}\kappa(t) = \kappa(t-\ell)$. This equation is not unlike the classical formulation of the variance in which the mean of κ is replaced by the $\kappa(t-i-1)$. Of course, it sounds appealing to propose probabilistic prediction using the variance of residual; but in GHI prediction, the Gaussian hypothesis is never verified[35,36]. It is also known that the prediction intervals are too wide and become quickly unusable with the hypothesis of the persistence of the variance[33]. Furthermore, the option of proposing increasingly complex non-parametric methods is satisfactory from a theoretical point of view but it is not very advantageous in practice. There are a lot of interpretations of Eq.10 and the attentive reader will recognize the formulation of the mean square error with respect to a persistence model or the fact that $\sigma_\tau^2(t) = \mathbb{E}[\left(\frac{\partial\kappa}{\partial t}\right)^2]$. With this equation, it is tempting to believe that we have stumbled upon a deep difficulty, but a κ breakdown into a trend (T) and fast fluctuations terms (ϵ) transforms Eq.10 into,

$$\sigma_\tau^2(t) = \frac{1}{\tau}\sum_{i=0}^{\tau-1}\Big((T(t-i)-\mathbb{E}[T]) - (\boldsymbol{B}^1 T(t-i)-\mathbb{E}[T]) + (\epsilon(t-i)-\mathbb{E}[\epsilon]) - (\boldsymbol{B}^1\epsilon(t-i)-\mathbb{E}[\epsilon])\Big)^2 \tag{11}$$

considering that $\kappa(t) = T(t) + \epsilon(t)$. To go further, we shall consider the co-variance $\sigma_\epsilon(t,t-1)$ and the partial autocorrelation functions which is identical to the autocorrelation function for the lag 1 $\beta_\epsilon(t,t-1)$ for dependency between ϵ and himself 1 lag delayed. By contrast with the standard analysis, here the high-frequency term (ϵ) has no mean and co-variance functions tending to 0 (β_T is a function close to 1 and $\sigma_T^2 << \sigma_\epsilon^2$) which means that Eq.11 can be replaced by,

$$\begin{aligned}\sigma_\tau^2(t) &= 2\sigma_\epsilon^2(t) - 2\sigma_\epsilon(t,t-1) + 2\sigma_T^2(t) - 2\sigma_T(t,t-1)\\ &= 2\sigma_\epsilon^2(t)\big(1-\beta_\epsilon(t,t-1)\big) + 2\sigma_T^2(t)\big(1-\beta_T(t,t-1)\big)\end{aligned} \tag{12}$$

Bearing in mind the above, this equation shows that there is a link between the volatility as described in Eq.2 and the variance of the high-frequency component (ϵ). By means of the König-Huygens' theorem and characteristics of linear correlation coefficient of Bravais-Pearson with the fact that the covariances between T and ϵ are close to 0 (considering $T \perp \epsilon$ with $\mathbb{E}[\epsilon] = 0$), we may show that Eq.12 could be replaced by the more practical expression $\sigma_\tau^2(t) \simeq 2\sigma_\epsilon^2(t)\big(1-\beta_\epsilon(t,t-1)\big)$. It is possible to settle T with a classical moving average defined by a (2n+1)-point mean values: $T(t) = \mathbb{E}[\kappa(t-n:t+n)]$ and ϵ with $\epsilon(t) = \kappa(t) - T(t)$. With the daytime filtering process (Section II) $n = 5$ provides a daily average. It will be needful to introduce the transform $\Gamma : \sigma_\tau \in [0,1] \to \Gamma(\sigma_\tau) \in \mathbb{R}^+$ in order to handle with σ_ϵ^2 (Eq.13 with $\beta < 1$, $\sqrt{(2)}\sigma_\epsilon = \Gamma(\sigma_\tau)$).

$$\Gamma(\sigma_\tau) = \frac{\sigma_\tau}{\sqrt{1-\beta_\epsilon(t,t-1)}} \tag{13}$$

Once the prediction $\widehat{z}(t+1)$ is obtained, it is easy to compute next value of the $\widehat{\kappa}(t+1) = \Re[\widehat{z}(t+1)]$ and the associated volatility $\widehat{\sigma}_\tau(t+1) = \Im[\widehat{z}(t+1)]$. From here, we propose to build an estimate of the probabilistic GHI prediction based on the point prediction ($\widehat{GHI}(t+1) = \widehat{\kappa}(t+1)GHI_{CS}(t+1)$) and the cumulative distribution function (F_κ) computed from the conditional volatility: $F_\kappa(x) = \mathbb{P}(\kappa < x)$[37]. This last term corresponds to the probability that the random κ variable takes on a value less than or equal to x. The probability that κ lies in the semi-closed interval $(a,b]$, is therefore $\mathbb{P}(a < \kappa \leq b) = F_\kappa(b) - F_\kappa(a)$. In the Gaussian case assumed here (the quantities *mean*, *expectation*, *median* and *mode* of the distribution are identical[38]), F_κ and his inverse F_κ^{-1} are defined from the error function (erf) as described respectively in,

$$\widehat{F}_\kappa(x) = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{x-\widehat{\kappa}}{\Gamma(\widehat{\sigma}_\tau)}\right) \tag{14}$$

where $x \in \mathbb{R}$ and,

$$\widehat{F}_\kappa^{-1}(q) = \widehat{\kappa} + \Gamma(\widehat{\sigma}_\tau)\,\mathrm{erf}^{-1}(2q-1) \tag{15}$$

where $0 < q < 1$. Probabilistic forecasting is more powerful than the deterministic one and allows us to bound

the prediction proposing that is called prediction interval from quantiles estimation at probability level $q \in [0,1]$ $\widehat{Q}(q) = inf\{x \in \mathbb{R} : \widehat{F}_\kappa(x) \geq q\}$. Consider that, if the the function $\widehat{F}$ is continuous and strictly monotonically increasing, we have the quantile function defined by $\widehat{Q}(q) = \widehat{F}^{-1}(q)$ (denoted `probit` function in the Gaussian case)[39]. In that instance of a central prediction interval (the most common way is to center the prediction interval on the median considering there is the same probability of risk below and above the median[40]). with a nominal coverage rate of $(1-\alpha)100\%$, the lower bound ($\underline{GHI}$) is estimated by using the $\alpha/2$ quantile and the upper bound ($\overline{GHI}$) using the $1-\alpha/2$ quantile as described in Eq.16 with an example quantile function estimation in the normal distribution case (erf^{-1} is an odd function).

$$\begin{cases} \underline{GHI} = \widehat{Q}(\alpha/2) = \widehat{GHI} - \text{erf}^{-1}(1-\alpha) GHI_{CS} \Gamma(\widehat{\sigma}_\tau) \\ \overline{GHI} = \widehat{Q}(1-\alpha/2) = \widehat{GHI} + \text{erf}^{-1}(1-\alpha) GHI_{CS} \Gamma(\widehat{\sigma}_\tau) \end{cases} \tag{16}$$

Point out that as this is very frequently done in solar irradiance prediction, these interval limits can in turn be limited by considering that the upper limit is necessarily lower than GHI_{CS} and that the lower limit shall be higher than the diffuse component of the GHI_{CS} (this quantity is easily obtained with the `Solis` modeling)[34]. We previously treated the $t+1$ case, nevertheless the reasoning for $\mu_{t+\ell}$ is rather similar replacing (in Eq.3b) $\widehat{z}(t+1)$ by $\widehat{z}(t+\ell)$. Assuming all the approximations made so far (normal assumption of σ_τ in Eqs.14 and 15, the arbitrary choice of τ in Eq.2 and the hypothesis on $\mathbb{E}[r]^2$ in Eq.10), it is doable and advisable to calibrate the $\mu_{t+\ell}$ value in Eq.9 according to nominal coverage rate $(1-\alpha)100\%$ by performing simulations on the training space (link between $(1-\alpha)100\%$ and $\mu_{t+\ell}$ values). Furthermore, in Table I, it is shown that the two approaches lead to quite different results and that Eq.16 shall only be considered as a first approximation requiring data-driven corrections (see Annex for details).

TABLE I. α_{t+1} estimations from Eq.16 and data guided approach performing simulations on the training space (data driven correction for $1h$ horizon and a 11-point mean values corresponding to $n=5$ and $\beta_\epsilon(t,t-1)=0.38$).

	$\boldsymbol{\alpha=0.2}$	$\boldsymbol{\alpha=0.4}$	$\boldsymbol{\alpha=0.6}$	$\boldsymbol{\alpha=0.8}$
α_{t+1} (Eq.16)[a]	1.15	0.76	0.47	0.23
Data driven[b]	1.04	0.57	0.31	0.17

[a] coupling Eqs13 and 16 α_{t+1} can be found like equal to $(1-\beta_\epsilon(t,t-1))^{-1/2}\,\text{erf}^{-1}(1-\alpha)$

[b] these data can be fitted with an exponential decay ($R^2=0.999$) according to $\alpha_{t+1} = 1.916e^{-3.034\alpha}$

IV. RESULTS

Despite the fact that the purpose of this paper is to elaborate a new way to propose GHI probabilistic forecast (complex-valued method denoted $Compl$), it is important to compare results with some classical tools, like a Gaussian parametric process (denoted $Gauss$ and based on the variance of the residual[36]), a non-parametric bootstrapped process (denoted $Boot$[41]) and a Ridge quantile regression model (denoted $Quant$[42–44]). The used error metrics for the comparison in the deterministic case is normalized root mean square error ($nRMSE$ [45]) while in the probabilistic case, we choose normalized mean interval length (MIL sometime denoted $PINAW$ for prediction interval normalized average width), percentage interval coverage probability ($PICP$), continuous rank probability score ($CRPS$) and mean scaled interval score ($MSIS$). All these metrics are described in van der Meer, Widén, and Munkhammar [43], Lauret, David, and Pinson [46], Hyndman and Koehler [47] and references therein. In Table II is shown the comparison between all the prediction interval methodologies and is proved that the complex approach is equivalent in terms of deterministic prediction ($nRMSE$ nearly identical for all five methods) but grants, considering a nominal coverage rate of 80%, a significant MIL decrease that is worthwhile for a grid operator who seeks a predictive methodology offers the lowest conceivable MIL for a given $PICP$. Another interesting element is the fact that for both $Quant$ and $Compl$ methods, $\alpha = 0.2$ (nominal covering rate of $100\%(1-0.2) = 80\%$) effectively corresponds to a $PICP$ close to 80% unlike the two other cases.

TABLE II. Models comparison for a nominal coverage probability of 80% ($\alpha = 0.2$)

Horizons	**Metrics**	**Gauss**	**Boot**	**Quant**	**Compl**
1h	*nRMSE*	0.197	0.203	0.201	0.196
	PICP(%)	83.81	75.21	79.65	80.01
	MIL(%)	51.24	40.36	42.24	41.24
2h	*nRMSE*	0.251	0.267	0.258	0.252
	PICP(%)	81.39	76.67	79.90	80.74
	MIL(%)	64.57	55.47	58.72	56.72
3h	*nRMSE*	0.282	0.304	0.289	0.282
	PICP(%)	80.68	73.27	80.07	80.06
	MIL(%)	70.97	58.22	67.32	63.09
4h	*nRMSE*	0.302	0.327	0.312	0.303
	PICP(%)	80.42	81.69	80.64	80.02
	MIL(%)	75.39	76.82	74.47	67.10
5h	*nRMSE*	0.316	0.362	0.328	0.317
	PICP(%)	80.99	74.44	80.81	79.49
	MIL(%)	78.73	64.45	79.08	69.32
6h	*nRMSE*	0.324	0.358	0.339	0.325
	PICP(%)	81.40	78.69	81.43	79.66
	MIL(%)	81.41	73.79	82.73	70.05

In Fig.3, one can observe how *Compl* forecast intervals are distributed considering $1h$ horizon. The main attraction of the method lies in the fact that the interval band is conditioned by the variability observed in the previous hours. Thus, for the days close to the 5150^{th} the prediction band (very small) is completely different from what is observed close to the 5350^{th} hours (very large). The probabilistic counterpart of the mean absolute error is the $CRPS$, making it possible to quantify the total error made with the predicted distributions as it is shown in Fig.4. Thence, it is a robust score that is designed in such a way that it measures both reliability and sharpness. An advantage of the $CRPS$ is that it reduces the absolute error if the forecast is deterministic, and allows the comparison between probabilistic and point forecasts[43]. We may note that even if the quantile regression is the best tool considering this metric, the errors observed by the complex-valued methodology are not prohibitive. This phenomenon is also visible by comparing the $MSIS$ (related to $\alpha = 0.2$) which has the enormous advantage of considering all the forecast horizons within a single metric. If for *Boot* and *Gauss*, $MSIS$ are respectively 1.05 and 1.03, for *Quant* and *Compl*, $MSIS$ are lower and so better (0.89 and 0.95).

V. CONCLUSIONS

The objective of this paper is to present a new method for predicting GHI that is able to take into account fast fluctuations. Often the literature boasts some sophisticated approaches, but when focusing on the existing installations, one remarks that the highly-developed models yield way to simpler methods. Although less effective, they are more robust and easier to use. From a practical point of view, a "good" method concerns a tool that would be easily usable in a stand-alone application (problems of some toolboxes), and which doesn't involve a lot of different concepts or data. The procedures used for smart management shall be self-sufficient and consistent with continuous learning and with some eventual detectors failure. It is in this perspective we tested a new univariate methodology based on the complex-valued time series generated from GHI measurements. With only a few parameters (6 complex numbers in the studied case) and some basic mathematical operations, this approach makes it possible to predict GHI with accuracy compared with classical probabilistic and deterministic predictions. This method proposes the lowest MIL considering a fixed nominal coverage rate (80%). Once the parameters have been estimated and provided that real-time GHI measurements are available, a simple spreadsheet can become a tool of choice in the management of PV installations. The validation of this approach will require many more tests by varying time steps, horizons, and forecastability[48] or predictability[49]. However, this new forecast methodology is simple to implement and may facilitate the integration of renewable energies and improve the management of installations using solar radiation as energy sources (smart grid, building, district, etc.). Interesting perspectives will be to apply it to other kinds of time series (not necessarily in connection with renewable energies), to construct the imaginary part concerning other variables than volatility (residuals, exogenous or ordinal data, etc.), and perhaps adapt the method to others predictors kinds (artificial neural network, support vector regression, etc.).

Conflict of Interest. The authors have no conflicts to disclose.

Data Availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Appendix: Data Driven Correction

The data-driven method proposed here allows for improving probabilistic forecasting. From Eq.9, it would be useful to determine experimentally (and not theoretically with Eq:16) $\mu_{t+\ell}$ such as $\mathbb{P}(GHI \in \Lambda) \to (1-\alpha)$ when the number of observations is large enough. All along the training step, curves fitting to inverse cumulative distribution functions are fixed ($\mu_{t+\ell}$ as a function of α). Its use requires a few assumptions (less than in the theoretical case presented above in this paper). The fine advantage lies in the fact that the Gaussian hypothesis no longer has any reason to exist (non-parametric method). Nonetheless, two new much less restrictive hypotheses must be formulated. The first one is that there is the same probability of the risk below and above the median (common postulation[40]) and the second one is that the σ_τ distribution is symmetric (the mean and the median are identical). The sample skewness is worth 0.1, hence it is regular to consider the second assumption as verified (since comprised between -1 and 1[50]). In Table III are shown the f_1 and f_2 parameters values concerning the fit $\mu_{t+\ell} = f_1 e^{f_2 \alpha}$. This data-driven method may be used to estimate the quantiles ($\widehat{Q}$) and so the cumulative

TABLE III. $\mu_{t+\ell}$ adjustment for each horizon ℓ (f_1 and f_2 the constants of the exponential decay fit)

ℓ	$f_1(CB95\%)$[a]	$f_2(CB95\%)$[a]	R^2[b]
1	1.916(1.745,2.087)	-3.034(-3.322,-2.747)	0.999
2	2.739(2.705,2.773)	-3.163(-3.218,-3.109)	0.994
3	2.828(2.797,2.860)	-2.811(-2.853,-2.769)	0.995
4	2.869(2.834,2.904)	-2.605(-2.647,-2.563)	0.994
5	2.821(2.787,2.855)	-2.393(-2.432,-2.354)	0.993
6	2.707(2.673,2.741)	-2.175(-2.214,-2.136)	0.992

[a] estimates with 95% confidence bounds
[b] coefficient of determination

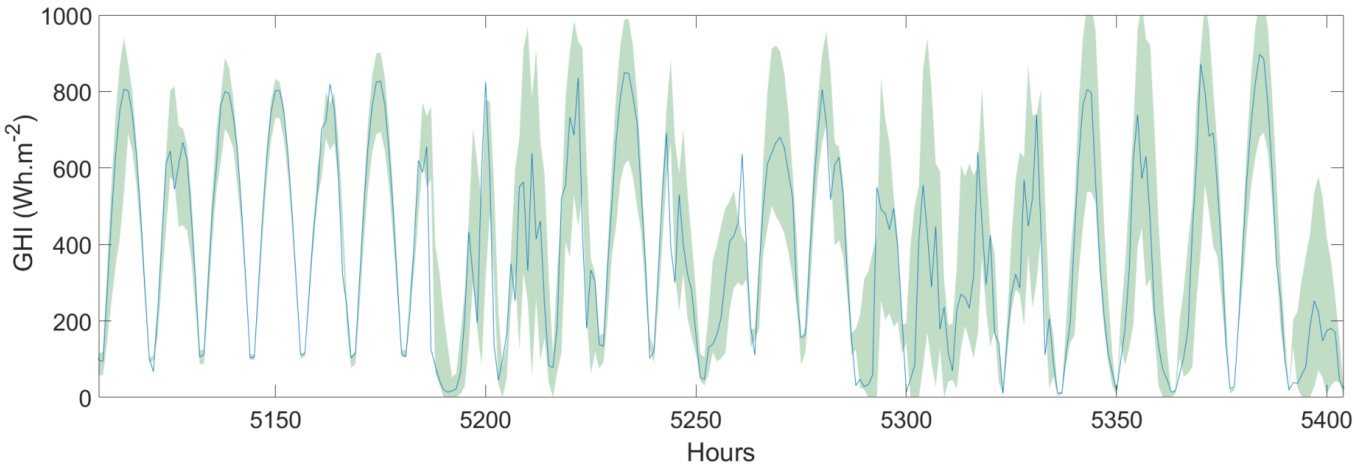

FIG. 3. 80% prediction interval with complex-valued approach versus measures (blue line)

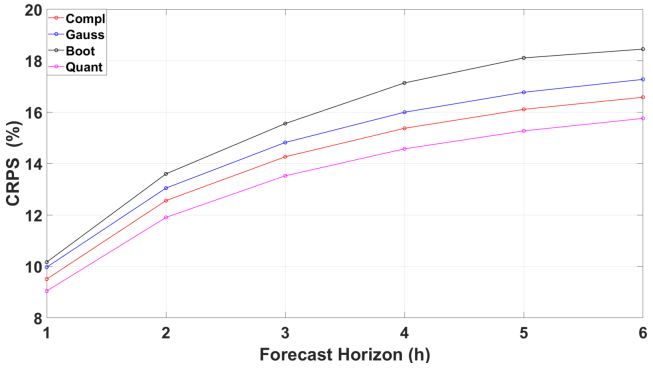

FIG. 4. *CRPS* for the probabilistic comparison

distribution function. Indeed, considering $\Delta q \in [0, 0.5]$, we assume,

$$\begin{cases} \widehat{Q}(0.5+\Delta q) = \widehat{Q}(0.5) + f_1 e^{f_2(1-2\Delta q)}\widehat{\sigma}_\tau \\ \widehat{Q}(0.5-\Delta q) = \widehat{Q}(0.5) - f_1 e^{f_2(1-2\Delta q)}\widehat{\sigma}_\tau \\ \widehat{Q}(0.5) = \widehat{GHI} \end{cases} \quad (A.1)$$

verified if and only if, $\widehat{Q}$ is a continuous function, which implies Eq.A.2 and thereby $f_1 e^{f_2} \to 0$.

$$\lim_{\substack{\Delta q \to 0^+ \\ \Delta q \to 0^-}} \widehat{Q}(0.5 \pm \Delta q) = \widehat{Q}(0.5) \quad (A.2)$$

Taking a concrete example, quantiles $\widehat{Q}(0.1)$ and $\widehat{Q}(0.9)$ could be respectively estimated from a nominal 80% prediction interval ($\alpha = 0.2$ and $\Delta q = 0.4$) with $\widehat{Q}(0.5) - f_1 e^{f_2 0.2}\sigma_\tau$ and $\widehat{Q}(0.5) + f_1 e^{f_2 0.2}\sigma_\tau$. To slightly improve the results, and to position oneself in a totally non-parametric approach, it is doable to use lookup tables rather than curve fitting.

ACKNOWLEDGMENTS

This work was partially supported by ANR grant SAPHIR project ANR-21-CE04-0014-03.

REFERENCES

[1] G. Notton, M.-L. Nivet, C. Voyant, C. Paoli, C. Darras, F. Motte, and A. Fouilloy, "Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting," Renewable and Sustainable Energy Reviews **87**, 96–105 (2018).

[2] Y. Zhou, Y. Liu, D. Wang, X. Liu, and Y. Wang, "A review on global solar radiation prediction with machine learning models in a comprehensive perspective," Energy Conversion and Management **235**, 113960 (2021).

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

[4] D. H. Alamo, R. N. Medina, S. D. Ruano, S. S. García, K. P. Moustris, K. K. Kavadias, D. Zafirakis, G. Tzanes, E. Zafeiraki, G. Spyropoulos, J. K. Kaldellis, G. Notton, J.-L. Duchaud, M.-L. Nivet, A. Fouilloy, and S. Lespinats, "An Advanced Forecasting System for the Optimum Energy Management of Island Microgrids," Energy Procedia Renewable Energy Integration with Mini/Microgrid, **159**, 111–116 (2019).

[5] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," IEEE Transactions on Evolutionary Computation **1**, 67–82 (1997).

[6] V. Cerqueira, L. Torgo, and C. Soares, "Machine learning vs statistical methods for time series forecasting: Size matters," (2019), arXiv:1909.13316 [stat.ML].

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

[8] M. Pagano, "On periodic and multiple autoregressions," Annals of Statistics **6**, 1310–1317 (1978).

[9] E. Gladhyšhev, "Periodically correlated random sequences," Sov. Math. Doklady **2**, 385–388 (1961).

[10] G. E. P. Box and G. M. Jenkins, Time series analysis: forecasting and control (Holden-Day, 1976) google-Books-ID: 1WVHAAAAMAAJ.

[11] Y. Yu and G. Hu, "Short-term solar irradiance prediction based on spatiotemporal graph convolutional recurrent neural net-

work," Journal of Renewable and Sustainable Energy **14**, 053702 (2022), https://doi.org/10.1063/5.0105020.
[12]L. Garcia-Gutierrez, C. Voyant, G. Notton, and J. Almorox, "Evaluation and comparison of spatial clustering for solar irradiance time series," Applied Sciences **12** (2022), 10.3390/app12178529.
[13]P. Ineichen, "A broadband simplified version of the Solis clear sky model," Solar Energy **82**, 758–762 (2008).
[14]E. Dimson and P. Marsh, "Volatility forecasting without data-snooping," Journal of Banking & Finance **14**, 399–421 (1990).
[15]M. Nwogugu, "Further critique of GARCH/ARMA/VAR/EVT Stochastic-Volatility models and related approaches," Applied Mathematics and Computation **182**, 1735–1748 (2006).
[16]M. David, F. Ramahatana, P. J. Trombe, and P. Lauret, "Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models," Solar Energy **133**, 55–72 (2016).
[17]A. Krawiecki, J. A. Hołyst, and D. Helbing, "Volatility Clustering and Scaling for Financial Time Series due to Attractor Bubbling," Physical Review Letters **89**, 158701 (2002).
[18]T. Kaizoji, "Statistical Properties of Absolute Log-Returns and a Stochastic Model of Stock Markets with Heterogeneous Agents," in Nonlinear Dynamics Heterogeneous Interacting Agents, Lecture Notes in Economics and Mathematical Systems, edited by T. Lux, E. Samanidou, and S. Reitz (Springer, Berlin, Heidelberg, 2005) pp. 237–248.
[19]S. Svetunkov, Complex-Valued Modeling in Economics (Springer-Verlag, New York, 2012).
[20]S. Dégerine and S. Lambert-Lacroix, "Characterization of the partial autocorrelation function of nonstationary time series," Journal of Multivariate Analysis **87**, 46 – 59 (2003).
[21]H. Pal, T. H. Seligman, and J. V. Escobar, "Correlation networks from random walk time series," Physical Review E **98**, 032311 (2018).
[22]J. A. Gubner, "Probability and Random Processes for Electrical and Computer Engineers," (2006).
[23]D. W. Adrian, R. Maitra, and D. B. Rowe, "Complex-valued time series modeling for improved activation detection in fmri studies," The Annals of Applied Statistics **12**, 1451–1478 (2018).
[24]M. Schuld, I. Sinayskiy, and F. Petruccione, "Prediction by linear regression on a quantum computer," Phys. Rev. A **94**, 022342 (2016).
[25]Y. Ahmadian, F. Fumarola, and K. D. Miller, "Properties of networks with partially structured and partially random connectivity," Phys. Rev. E **91**, 012820 (2015).
[26]J. Claerbout, Geophysical Image Estimation by Example (Lulu.com, 2014) google-Books-ID: cvszBwAAQBAJ.
[27]A. van den Bos, "Estimation of Complex Parameters," IFAC Proceedings Volumes IFAC Symposium on System Identification (SYSID'94), Copenhagen, Denmark, 4-6 July, **27**, 1429–1433 (1994).
[28]D. W. Marquardt and R. D. Snee, "Ridge regression in practice," The American Statistician **29**, 3–20 (1975).
[29]T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," Journal of the Royal Statistical Society: Series B (Statistical Methodology) **69**, 243–268 (2007).
[30]H. Khajeh and H. Laaksonen, "Applications of probabilistic forecasting in smart grids: A review," Applied Sciences **12** (2022).
[31]C. Lobry and T. Sari, "Non-standard analysis and representation of reality," International Journal of Control **81**, 519–536 (2008).
[32]P. Cartier and Y. Perrin, "Integration over finite sets," in Nonstandard Analysis in Practice, Universitext, edited by F. Diener and M. Diener (Springer, Berlin, Heidelberg, 1995) pp. 185–204.
[33]M. Fliess, C. Join, and F. Hatt, "Volatility made observable at last," arXiv:1102.0683 [cs, q-fin] (2011), arXiv: 1102.0683.
[34]M. Fliess, C. Join, and C. Voyant, "Prediction bands for solar energy: New short-term time series forecasting techniques," Solar Energy **166**, 519–528 (2018).
[35]J. R. Trapero, "Calculation of solar irradiation prediction intervals combining volatility and kernel density estimates," Energy **114**, 266–274 (2016).
[36]C. Voyant, J. G. De Gooijer, and G. Notton, "Periodic autoregressive forecasting of global solar irradiation without knowledge-based model implementation," Solar Energy **174**, 121–129 (2018).
[37]M. P. Deisenroth, A. A. Faisal, and C. S. Ong, Mathematics for Machine Learning (Cambridge University Press, 2020) google-Books-ID: pbONxAEACAAJ.
[38]K. Krishnamoorthy, Handbook of Statistical Distributions (CRC Press, 2006) google-Books-ID: FEE8D1tRl30C.
[39]R. J. Hyndman and Y. Fan, "Sample Quantiles in Statistical Packages," The American Statistician **50**, 361–365 (1996).
[40]P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: required properties and evaluation," Wind Energy **10**, 497–516 (2007), https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.230.
[41]L. Pan and D. N. Politis, "Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions," Journal of Statistical Planning and Inference **177**, 1 – 27 (2016).
[42]P. Lauret, M. David, and H. T. C. Pedro, "Probabilistic Solar Forecasting Using Quantile Regression Models," Energies **10**, 1591 (2017).
[43]D. W. van der Meer, J. Widén, and J. Munkhammar, "Review on probabilistic forecasting of photovoltaic power production and electricity consumption," Renewable and Sustainable Energy Reviews **81**, 1484–1512 (2018).
[44]T. C. Carneiro, P. A. Rocha, P. C. Carvalho, and L. M. Fernández-Ramírez, "Ridge regression ensemble of machine learning models applied to solar and wind forecasting in brazil and spain," Applied Energy **314**, 118936 (2022).
[45]C. Voyant, G. Notton, J.-L. Duchaud, L. A. G. Gutiérrez, J. M. Bright, and D. Yang, "Benchmarks for solar radiation time series forecasting," Renewable Energy **191**, 747–762 (2022).
[46]P. Lauret, M. David, and P. Pinson, "Verification of solar irradiance probabilistic forecasts," Solar Energy **194**, 254–271 (2019).
[47]R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," International Journal of Forecasting **22**, 679–688 (2006).
[48]C. Voyant, P. Lauret, G. Notton, J.-L. Duchaud, A. Fouilloy, M. David, Z. M. Yaseen, and T. Soubdhan, "A monte carlo based solar radiation forecastability estimation," Journal of Renewable and Sustainable Energy **13**, 026501 (2021), https://doi.org/10.1063/5.0042710.
[49]X. Yang, D. Yang, J. M. Bright, G. M. Yagli, and P. Wang, "On predictability of solar irradiance," Journal of Renewable and Sustainable Energy **13**, 056501 (2021), https://doi.org/10.1063/5.0056918.
[50]D. George and P. Mallery, "Spss for windows step-by-step: A simple guide and reference, 14.0 update (7th edition)," http://lst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/[in=epidoc1.in]/?t2000=026564/(100) (2003).