Complex-Valued Time Series Based Prediction: Case of Solar Irradiance Probabilistic Prediction

Cyril Voyant, Gilles Notton, and Jean-Laurent Duchaud
University of Corsica
SPE Laboratory UMR6134, Corte (France)

Philippe Lauret
University of Reunion
PIMENT Laboratory, Saint-Pierre (France)

(Dated: March 15, 2022)

This paper describes a new way to predict real time series using complex-valued elements. An example is given in the case of the short-term probabilistic global solar irradiance forecasts with measurement as real part and an estimate of the volatility as imaginary part. A simple complex autoregressive model is tested with data collected in Corsica island (France). Results show that, even if this approach is simple to set up 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 like Gaussian process, bootstrap methodology or even more sophisticated model like quantile regression. The number of models that it is possible to build by generating complex-valued time series is substantial. Indeed, by using exogenous or ordinal variables and computed quantities coupled with complex (or multi-complex) numbers, many studies and many fields of physics could benefit from this methodology and from the many models that result from it.

I. INTRODUCTION

From an applied study case of time series forecasting, a new predictive methodology based on the generation of complex numbers will be discussed. The domain studied is the energy generation and particularly the renewable energy part and its difficulty of integration.

If conventional generators (using fossil or nuclear fuels) provide all required services to operate the system at given reliability and power quality levels in an electrical grid, sustainable power supply using solar and wind power is more difficult to manage and complicate the grid operator task. These renewable sources are qualified as variable, non-predictable and even non-dispatchable. The main characteristics making these energies difficult to control and limiting their integration level are: their non-synchronous nature (they use power electronics devices and have no inertia), their uncertainty related to their random generation and their variability [1]. 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 [2].

The topic of this paper falls within the development of a forecasting method for PV power generation. Many forecasting methods exist according to the forecasting horizon and the data time-step. Sky imaging for very short-term (1-60 min), time series models for short-term (from 30 min to a few hours) and physical models (or numerical weather prediction models) for medium-term (from a few hours to a few days). [3][4].

Numerous machine learning methods benchmarks have been published in the literature and most of them compare the models in term on accuracy [5][6] with regard to time horizons. Nevertheless, the use of such complex forecasting for real power plants (or energy) management systems, is not always possible (lack of data, acquisition system failures, process execution time, etc.) and leads many grid managers to use the simplest ones sometimes at the expense of their performance [7].

With agreement with the “No Free Lunch theorem” [8], which proposes that no learning algorithm is the most suitable in all scenarios [9], the objective of this paper is to elaborate and test a new probabilistic forecasting method, easy to implement, with a good accuracy and based on a new mathematical method. An 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.

II. DATA

As detailed by Yang in [10], 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 and its two seasonal periods (a yearly cycle and a diurnal cycle). Since 1961 and the first works about stationary processes with a finite second moment [11]
and periodic correlation (or covariance) \[\text{[12]}\], the scientists know that it is important to pay attention to trends when time series is used. Box and Jenkins’ first formalism \[\text{[13]}\] clarified this aspect by proposing a decomposition, especially when seasonality is easily quantifiable. With a knowledge of a multiplicative double-seasonal patterns, it seems natural and legitimate to effectively take it into consideration by computing a classical ratio defined as GHI on the solar irradiance in clear sky condition (denoted $GHI_{CS}$). This one results in a normalized quantity (theoretically comprised between 0 and 1) known as the clear-sky index ($\kappa(t)$) in Eq.\[\text{[1]}\].

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

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

\[\text{翛} \] $GHI$ 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 a daily solar irradiation (annual average) of 4.5 kWh.m$^{-2}$.

\[\text{翛} \] Models are evaluated during only daytime irradiance values, filtering the checked data (according to quality control \[\text{[19]}\]) on solar zenith angle ($GHI = 0$ if $\theta_Z > 85^\circ$),

\[\text{翛} \] $GHI_{CS}$ is computed with the Solis model which proposes an atmospheric scheme based on radiative transfer calculations and the Lambert–Beer relation \[\text{[18]}\].

This paper is dedicated to the volatility prediction which is used to generate $GHI$ prediction intervals with respect to prediction horizon from 1h to 6h with 1h time granularity (training during the years 2008-2017 and testing during the year 2018). Several methods based on the conditional heteroscedasticity and denoted GARCH or 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 methods has never been used, probably due to its complexity \[\text{[17]}\], the restrictive assumptions \[\text{[18]}\] or the quality of its results which seemed even so promising \[\text{[19]}\]. An important conclusion of \[\text{[17]}\] concerning the ARCH family predictors, is an other form 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:

\[\text{翛} \] The prediction of the conditional volatility (i.e. upon an information set) using parameters like the return and its standard deviation (see definition in Eq.\[\text{[2]}\])

\[\text{翛} \] The generation of $\kappa$ and then $GHI$ prediction intervals.

From the computed $\kappa$ time series (Eq.\[\text{[1]}\]), 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 (or volatility $\sigma_{\tau}$) 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 $\kappa$ and $\sigma_{\tau}$, we propose 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($p$) of random variables defined on the same probability space\[\text{[11]}\]). This method is to be compared to that exposed by Ivan Svetunkov concerning the logic of Brown’s exponential smoothing

\[\text{cesso} \] $\sigma_{\tau}(t) = \sqrt{\frac{1}{\tau} \sum_{i=0}^{\tau - 1} (r(t - i) - \frac{1}{\tau} \sum_{j=0}^{\tau - 1} r(t - j))^2} \] \[2\]

Figure \[\text{A}\] shows that the trend of the centered $\sigma_{\tau}$ (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 $\kappa$ and $\sigma_{\tau}$, we propose 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($p$) of random variables defined on the same probability space\[\text{[11]}\]). This method is to be compared to that exposed by Ivan Svetunkov concerning the logic of Brown’s exponential smoothing

\[\text{cesso} \] In the literature, other definitions can be found for the volatility \[\text{[20]}\] 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.

\[\text{cesso} \] With three components: a nonempty sample space, its $\sigma$-algebra denoted event space together with a probability measure $\mathbb{P}$ that make it possible to define distribution function $F$.
methods and the complex-valued time series used to forecast two time series simultaneously [22], 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 \( \kappa \) and its volatility \( \sigma_\tau \) (Eq.3a) with (\( \hat{\kappa} \)) for predicted values) by a single regression equation (Eq.3b, the proof is obvious setting \( \omega = \xi + j\zeta, \omega \in \mathbb{C} \) and \( \xi, \zeta \in \mathbb{R} \)).

\[
\begin{align}
\hat{\kappa}(t+1) &= \sum_{i=0}^{p-1} \kappa(t-i)\xi_i - \sum_{i=0}^{p-1} \sigma_\tau(t-i)\xi_i \\
\hat{\sigma}_\tau(t+1) &= \sum_{i=0}^{p-1} \kappa(t-i)\zeta_i + \sum_{i=0}^{p-1} \sigma_\tau(t-i)\zeta_i \\
\hat{\xi}(t+1) &= \sum_{i=0}^{p-1} z(t-i)\omega_i
\end{align}
\]

Before using an auto-regressive (AR) model, it is important to deal with the model identification (the choice of the parameter or order \( p \) in Eq.3a), a classical tool widely studied in regression analysis is employed. It handles with the interpretation of the partial auto-correlation factor (\( \beta, [13] \)) according to real and imaginary parts of \( z \) (respectively \( \Re(z) \) and \( \Im(z) \)) as described in Eq.4

\[
\forall t > p \in \mathbb{Z}, [23, 24].
\]

Inasmuch as \( p_R, p_\Im \) and \( p \) denote AR orders having connections with \( \Re(z) \), \( \Im(z) \) and \( z \), we are setting \( p = \max(p_R, p_\Im) \) 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 (\( p = 6 \) for the studied site). There are several identification methods, as for example the complex auto-correlation factor [25], however, it seems that this method has the best ratio complexity-efficiency.

The next step is the model estimation (\( \hat{\omega} \)) by transposing what has been done for many years in the real-valued case (least square optimisation) to the complex-valued case. Considering an input matrix \( I \in \mathbb{C}^{D \times p} \) (Eq.5) and an output column vector \( o \in \mathbb{C}^{D \times 1} \) (Eq.6), the solution of the AR(p) least squares problem consists in determining unknown parameters (\( \hat{\omega} \in \mathbb{C}^{p \times 1} \), Eq.7). By the way, the problem, already raised and well detailed in [26] for spatial data-based model is resumed as in the classical real-valued case by \( \hat{\omega} = \hat{\omega} \) and can be solved from the formulation of the mean square error estimation \( \mathbb{E}[e^2] = ||\hat{\omega} - \hat{o}||^2 \) [27]. 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 have to prove that there is equivalence with the maximum likelihood and have to formulate hypotheses on the complex residual distribution.

\[
I = \begin{pmatrix}
(z(t-1)) & (z(t-2)) & \cdots & (z(t-p)) \\
(z(t-2)) & (z(t-3)) & \cdots & (z(t-p-1)) \\
\vdots & \vdots & \ddots & \vdots \\
(z(t-D)) & (z(t-D-1)) & \cdots & (z(t-D-p+1))
\end{pmatrix}_{D \times p}
\]

\[
\hat{\omega} = (\omega_1, \ldots, \omega_p)^{'p \times 1}, \hat{o} = (z(t), \ldots, z(t-D+1))^{'D \times 1}
\]

The complex-valued case differs from the real-valued one, replacing the \( l_2 \)-norm by the Frobenius norm introducing the Frobenius inner product on \( \mathbb{C}^D [25] \) \( \mathbb{E}[e^2] = < (\omega - \hat{\omega}), (\omega - \hat{\omega}) >_F \). Classically, the minimum of the squared expected value (argmin(\( \mathbb{E}[e^2] \)) = \( (\omega \in \mathbb{C}^p)\forall \omega^* \in \mathbb{C}^p : \mathbb{E}[e^2(\omega^*)] \geq \mathbb{E}[e^2(\omega)] \)) is carried out computing its differentiating (Eq.7) and by letting \( \partial\mathbb{E}[e^2]/\partial(\omega^H) = 0 \)

\[
\frac{\partial\mathbb{E}[e^2]}{\partial(\omega^H)} = \frac{\partial(I\omega - \hat{\omega})^H(I\omega - \hat{\omega})}{\partial(\omega^H)} = \frac{\partial(\omega^H I^H - \hat{\omega}^H)(I\omega - \hat{\omega})}{\partial(\omega^H)}
\]

It must be emphasised that \( \partial\mathbb{E}[e^2]/\partial(\omega^H) \) is the complex conjugate transpose of \( \partial\mathbb{E}[e^2]/\partial\omega \), thus, setting one to zero also sets the other to zero. The normal equation, in this complex-valued case becomes Eq.8

\[
I^H(I\omega - \hat{\omega}) = 0
\]

Furthermore, the solution of this matrix equation corresponds to a regression coefficients estimated by \( \hat{\omega} = (I^H I)^{-1} I^H \hat{\omega} \) [29]. 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 [30].

On top of that, to improve the condition number of the problem, one can introduce a constrained minimization with \( ||\omega||^2 < r(\lambda) \) where \( r \) is a bijective function and \( \lambda \) is the Lagrange multiplier of the constraint \( (I^H I + \lambda^H \lambda)^{-1} \) defined positive and so invertible. This approach denoted \textbf{Ridge} approach [31] can also be used in the complex-value case, in the form \( \hat{\omega}_\lambda = (I^H I + \lambda I)^{-1} I^H \hat{\omega} \). A machine learning-like approach consists to perform a cross-validation and select the \( \lambda \) value that minimizes the out-sample sum of squared residuals, in the experimental

\[ (\cdot)^H \text{for conjugate transpose} \]
setup, we chose $\lambda = 3.74$. Now the identification and optimization problems have been analyzed (the outcome of the experiment is detailed in Figure II), it is required to theoretically validate the use of predictions of both the $GHI$ and its volatility in the case of the probabilistic forecasting [32]. The goal is to show that the volatility as detailed previously, could be used to capture the idea of unpredictable and quick fluctuations. Thereby, considering the $\ell$-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 $\Lambda$ satisfying the Eq $\ref{9}$. The following is dedicated to the answer to this question: is there any theoretical meaning of this, and if so, what does $\mu_{t+\ell}$ worth?

$$\Lambda = \left[\hat{G}HI(t + \ell) - \mu_{t+\ell}GHI_{CS}(t + \ell)\hat{\sigma}_\tau(t + \ell), \right.$$  
$$\left. \hat{G}HI(t + \ell) + \mu_{t+\ell}GHI_{CS}(t + \ell)\hat{\sigma}_\tau(t + \ell)\right]$$  

(9)

First of all, it is important to consider this study within non-standard analysis framework with $S$-integrable time series (additive decomposition [33]) and by referring to the Cartier-Perrin theorem [34], as suggested by Fliess [35] [36]. 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 $\ref{2}$, it is conceivable to interpret the volatility thanks to the Backshift operator $\mathbf{B}$, it holds Eq $\ref{10}$ where $\mathbb{E}[r^2] \xrightarrow{\tau \rightarrow +\infty} \mathbb{E}[r^2]$, and $\mathbf{B}^\ell \kappa(t) = \kappa(t - \ell)$.

$$\sigma^2(t) = \mathbb{E}[r^2] - \mathbb{E}[r^2]^2 \simeq \mathbb{E}[r^2] = \mathbb{E}[(\kappa - \mathbf{B}^\ell \kappa)^2]$$  
$$= \frac{1}{\tau} \sum_{i=0}^{\tau-1} \left( \kappa(t - i) - \kappa(t - i - 1) \right)^2$$  

(10)

This equation is not without reminding the classical formulation of the variance in which the mean of $\kappa$ is replaced by the $\kappa(t - i - 1)$. Of course, it sounds appealing to propose probabilistic prediction using variance of residual; but in $GHI$ prediction, the Gaussian hypothesis is never verified [37] [38]. It is also known that the predictions intervals are too wide and become quickly unusable with the hypothesis of the persistence of the variance [39]. 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 $\ref{10}$, the attentive reader will recognize the formulation of the mean square error with respect to a persistence model or the fact that $\sigma^2(t) = \mathbb{E}[(\frac{\partial \kappa}{\partial t})^2]$. With this equation, it is tempting to believe that we have stumbled upon a deep, but certainly not fast fluctuations. From Eq $\ref{2}$, it is conceivable to interpret the volatility thanks to the Backshift operator $\mathbf{B}$, it holds Eq $\ref{10}$ where $\mathbb{E}[r^2] \xrightarrow{\tau \rightarrow +\infty} \mathbb{E}[r^2]$, and $\mathbf{B}^\ell \kappa(t) = \kappa(t - \ell)$.

$$\sigma^2(t) = \frac{1}{\tau} \sum_{i=0}^{\tau-1} \left( (T(t - i) - \mathbb{E}[T]) - (\mathbf{B}^\ell T(t - i) - \mathbb{E}[T]) \right)^2$$  
$$+ (\epsilon(t - i) - \mathbb{E}[\epsilon]) - (\mathbf{B}^\ell \epsilon(t - i) - \mathbb{E}[\epsilon]) \right)^2$$  

(11)

To go further, we shall consider the co-variance $(\sigma_\epsilon(t, t - 1))$ and the partial auto-correlation function $(\beta_\epsilon(t, t - 1))$ for dependency between $\epsilon$ and himself 1 lag delayed. By contrast with the standard analysis, here the high frequency term $(\epsilon)$ has no mean and co-variance functions tending to 0 $(\beta_\epsilon$ is a function close to 1 and $\sigma^2_\epsilon <<< \sigma^2_\tau$) which means that Eq $\ref{11}$ can be replaced by Eq $\ref{12}$.

$$\sigma^2(t) = 2\sigma^2_\tau(t) - 2\sigma_\epsilon(t, t - 1) + 2\sigma^2_\epsilon(t) - 2\sigma_\tau(t, t - 1)$$  
$$= 2\sigma^2_\tau(t) (1 - \beta_\epsilon(t, t - 1) + 2\sigma^2_\epsilon(t) (1 - \beta_\epsilon(t, t - 1))$$  

(12)

Bearing in mind the above, this equation shows that there is a link between the volatility as described in Eq $\ref{2}$ and the variance of the high frequency component $(\epsilon)$. By means of the Konig-Huygens’ theorem and characteristics of linear correlation coefficient of Bravais-Pearson with the fact that the covariances between $T$ and $\epsilon$ are close

---

\(^{iv}\) Over a sufficiently large interval, $\kappa(t)$ oscillates around a constant mean value, making the average of the return close to 0

\(^{v}\) identical to the auto-correlation function for the lag 1
The most common way is to center the prediction interval on the point mean values: \( T(t) = \mathbb{E}[\kappa(t-n : t+n)] \) and \( \epsilon \) 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 \( \Gamma: \sigma_{\epsilon} \in [0,1] \rightarrow \Gamma(\sigma_{\epsilon}) \in \mathbb{R}^+ \) in order to handle with \( \sigma_{\epsilon}^2 \) (Eq.13 with \( \beta < 1 \), \( \sqrt{2}\sigma_{\epsilon} = \Gamma(\sigma_{\epsilon}) \)).

Once the prediction \( \hat{z}(t+1) \) is obtained, it is easy to compute next value of the \( \hat{\kappa}(t+1) = \mathbb{R}[\hat{z}(t+1)] \) and the associated volatility \( \hat{\sigma}_{\epsilon}(t+1) = 3[\hat{z}(t+1)] \).

Whence, we propose to build an estimate of the probabilistic GHI prediction based on the point prediction \( \hat{GHI}(t+1) = \hat{\kappa}(t+1)GHI_{CS}(t+1) \) and the cumulative distribution function \( \hat{F}_\kappa(x) = \mathbb{P}(\kappa < x) \). This last term corresponds to the probability that the random \( \kappa \) variable takes on a value less than or equal to \( x \). The probability that \( \kappa \) lies in the semi-closed interval \( (a,b) \), is therefore \( \mathbb{P}(a < \kappa \leq b) = \hat{F}_\kappa(b) - \hat{F}_\kappa(a) \). In the Gaussian case (assumed here), \( \hat{F}_\kappa \) and its inverse \( \hat{F}_\kappa^{-1} \) are defined with the error function \( \text{erf} \) as described respectively in Eqs.14 \( (x \in \mathbb{R}) \) and 15 \( (q \in \mathbb{R} \mid 0 < q < 1) \).

The 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 \( q \in [0,1] \) \( \hat{Q}(q) = \inf \{ x \in \mathbb{R} : \hat{F}_\kappa(x) \geq q \} \). Consider that, if the the function \( \hat{F}_\kappa \) is continuous and strictly monotonically increasing, we have the quantile function defined by \( \hat{Q}(q) = \hat{F}_\kappa^{-1}(q) \) (denoted \text{probit} function in the Gaussian case)\(^vi\). In that instance of a central prediction interval with a nominal coverage rate of \( (1 - \alpha)100\% \), the lower bound \( (GHI) \) is estimated by using the \( \alpha/2 \) quantile and the upper bound \( (GHT) \) 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).

\[
\begin{align*}
\hat{GHI} = \hat{Q}(\alpha/2) &= \hat{GHI} - \text{erf}^{-1}(1 - \alpha)GHI_{CS}\Gamma(\hat{\sigma}_{\epsilon}) \\
\hat{GHI} = \hat{Q}(1 - \alpha/2) &= \hat{GHI} + \text{erf}^{-1}(1 - \alpha)GHI_{CS}\Gamma(\hat{\sigma}_{\epsilon})
\end{align*}
\]

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)\(^vii\).

We previously treated the \( t+1 \) case, nevertheless the reasoning for \( \mu_{t+t} \) is rather similar replacing in Eq.9 \( \hat{z}(t+1) \) by \( \hat{z}(t+\ell) \). Assuming all the approximations made so far (normal assumption of \( \sigma_{\epsilon} \) in Eqs.14 and 15) the arbitrary choice of \( \tau \) in Eq.2 and the hypothesis on \( \mathbb{E}[\epsilon]^2 \) in Eq.10, it is doable and advisable to calibrate the \( \mu_{t+t} \) 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 \( \mu_{t+t} \) 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).

\[
\begin{array}{cccccc}
\alpha & 0.2 & 0.4 & 0.6 & 0.8 \\
\alpha_{t+1} (\text{Eq.16}) & 1.15 & 0.76 & 0.47 & 0.23 \\
\text{Data driven} & 1.04 & 0.57 & 0.31 & 0.17 \\
\end{array}
\]

\( ^vi \) In this case, the quantities mean, expectation, median and mode of the distribution are identical.

\( ^vii \) 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 (sic).
interval methodologies and is proved that the complex approach is equivalent in term of deterministic prediction ($nRMSE$ nearly identical for all five methods) but grants, considering a nominal coverage rate of 80%, a significant decrease of interval length ($MIL$) 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 the both $Quant$ and $Compl$ methods, $\alpha = 0.2$ effectively corresponds to a $PICP$ close to 80% unlike the two other cases. This conclusion can be confirmed in Figure III, the $RD$ diagram allows for visualize the conditional distribution of measurements, given the forecast probability.

It groups the forecasts into bins according to the issued probability (horizontal axis). The frequency with which the event was observed to occur for this sub-group of forecasts is then plotted against the vertical axis. For perfect reliability, the forecast probability and the frequency of occurrence should be equal, and the plotted points should lie on the diagonal (as it is the case for $Quant$). $Compl$ departs from the diagonal but for quantiles $< 0.2$ and $> 0.8$ the proportions are respected. It is this result which allowed good results ($PICP$) in Table II. One can estimate that this model is well calibrated for forecast intervals with nominal coverage rate greater than 60% ($\alpha < 0.4$) which corresponds precisely to what is the most used in operational mode.

In Figure IV, one can observe how this forecast interval is distributed with the complex-valued method considering 1h horizon predictions. The main attraction of the method lies in the fact that the days close to the 5150th and 5350th hours are (from a measurement point of view) akin and yet, seeing that the model detects (or not) periods of high variability (related to the previous days with the $\tau$ sliding window), the forecast band will be more (or less) 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 Figure V. 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 to the absolute error if the forecast is deterministic, and allows us for comparison between probabilistic and point forecasts [14]. We may note that even if the quantile regression is the best tool considering this metric, the errors observed by the complex-valued methodology is 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).

![Reliability diagram for the probabilistic comparison](image)

**FIG. 3.** Reliability diagram for the probabilistic comparison

### V. CONCLUSIONS

The purpose of this paper is not only to present that $GHI$ should be predicted using a complex-valued AR but mainly to raise awareness on the fact that simple methods based on complex number generation and least square optimization could in some areas be used and involve good results.

Often the literature (focused on $GHI$ prediction) boast some sophisticated approaches but when focusing on the existing installations, one remarks that the highly-developed models yield way to simple methods. Although less effective, they are more robust and easier to use. From a practical point of view, a “good” method concerns tool which would be easily usable in a stand alone application (problems of some toolboxes), and which doesn’t involves a great deal of different concepts or data. The procedures used for the smart management shall be self-sufficient and consistent with a continuous learning and with some eventual detectors failure. It is in this per-
FIG. 4. 80% prediction interval with complex-valued approach versus measures (blue line)

FIG. 5. CRPS for the probabilistic comparison

spective 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 basics mathematical operations, this approach makes it possible to predict GHI with accuracy comparing with classical probabilistic predictions and lets the generation of efficient prediction intervals (with the lowest MIL considering a nominal coverage rate of 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.

Obviously, it will be ineluctable to test this approach on other time steps, horizons and sites (with higher variabilities), and especially to estimate the tilted GHI (the solar panels are rarely horizontal). 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 kind 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.).

Appendix: Data driven correction

The data driven method proposed here allows to improve probabilistic forecasting. From Eq.7, 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 \( (\mu_{t+\ell} \text{ as a function of } \alpha) \). 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 \[42\]) and the second one that the \( \sigma_{\tau} \) 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 \[48\]). 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 \( \hat{Q} \) and so the cumulative distribution function. Indeed, considering \( \Delta q \in [0, 0.5], \) we assume Eq.\[A.1\]

\[
\begin{align*}
\hat{Q}(0.5 + \Delta q) &= \hat{Q}(0.5) + f_1 e^{f_2 (1 - 2\Delta q)} \hat{\sigma}_{\tau} \\
\hat{Q}(0.5 - \Delta q) &= \hat{Q}(0.5) - f_1 e^{f_2 (1 - 2\Delta q)} \hat{\sigma}_{\tau} \\
\hat{Q}(0.5) &= GHI
\end{align*}
\]  

(A.1)
Verified if and only if, $\hat{Q}$ is a continuous function, which implies Eq. (A.2) and thereby $f_1 e^{f_2} \to 0$.

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

Taking a concrete example, quantiles $\hat{Q}(0.1)$ and $\hat{Q}(0.9)$ could be respectively estimated from a nominal 80% prediction interval ($\alpha = 0.2$ and $\Delta q = 0.4$) with $\hat{Q}(0.5) - f_1 e^{f_2} 0.2 \sigma_1$ and $\hat{Q}(0.5) + f_1 e^{f_2} 0.2 \sigma_1$. To slightly improve the results, and to position oneself in a totally non-parametric approach, it is doable to use lookup tables rather than curves fitting.

| $\ell$ | $f_1$ (CB95%) | $f_2$ (CB95%) | $R^2$ |
|-------|---------------|---------------|-------|
| 1     | 1.916(1.745,2.087) | -3.034(-3.322,-2.747) | 0.995 |
| 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 |

[1] IRENA, Transforming small-island power systems, report of international renewable energy agency (2018).
[2] G. Notton, M.-L. Nivet, C. Voiyant, C. Paoli, C. Darra, 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 (2018).
[3] S. Pelland, J. Remund, J. Kleissl, T. Oozeki, and K. C. De Brabandere, Photovoltaic and Solar Forecasting: State of the Art (IEA, 2013).
[4] B. Elliston and I. Macgill, The potential role of forecasting for integrating solar generation into the Australian electricity market, Solar 2010, the 48th AuSES Annual Conference, Canberra (2010).
[5] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, Review of solar irradiance forecasting methods and a proposition for small-scale insular grids, Renewable and Sustainable Energy Reviews 27, 65 (2013).
[6] C. Voiyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, Renewable Energy, 569 (2017).
[7] 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. Kaldelis, 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 (2019).
[8] D. H. Wolpert and W. G. Macready, No free lunch theorems for optimization, IEEE Transactions on Evolutionary Computation 1, 67 (1997).
[9] V. Cerqueira, L. Torgo, and C. Soares, Machine learning vs statistical methods for time series forecasting: Size matters (2019), arXiv:1909.13316 [stat.ML].
[10] D. Yang, Choice of clear-sky model in solar forecasting, Journal of Renewable and Sustainable Energy 12, 026101 (2020).
[11] M. Pagano, On periodic and multiple autoregressions, Annals of Statistics 6, 1310 (1978).
[12] E. Gladyszhev, Periodically correlated random sequences, Sov. Math. Doklady 2, 385 (1961).
[13] G. E. P. Box and G. M. Jenkins, Time series analysis: forecasting and control (Holden-Day, 1976) google-Books-ID: 1WVHAAAAMAAJ.
[14] 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 (2018).
[15] B. Espinar, P. Blanc, L. Wältl, P. C. Hoyer-Klick, M. Schroedter-Homscheidt, and T. Wanderer, On quality control procedures for solar radiation and meteorological measures, from subhourly to monthly average time periods, in European Geosciences Union 2012 (Vienna, Austria, 2012) p. 9134.
[16] P. Ineichen, A broadband simplified version of the Solis clear sky model, Solar Energy 82, 758 (2008).
[17] E. Dimson and P. Marsh, Volatility forecasting without data-snooping, Journal of Banking & Finance 14, 399 (1990).
[18] M. Nwogu, Further critique of GARCH/ARMA/VAR/EVT Stochastic-Volatility models and related approaches, Applied Mathematics and Computation 182, 1735 (2006).
[19] 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 (2016).
[20] A. Krawiec, J. A. Holyst, and D. Helbing, Volatility Clustering and Scaling for Financial Time Series due to Attractor Bubbling, Physical Review Letters 89, 158701 (2002).
[21] 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.
[22] S. Svetunkov, Complex-Valued Modeling in Economics (Springer-Verlag, New York, 2012).
[23] S. Dégérine and S. Lambert-Lacroix, Characterization of the partial autocorrelation function of non-stationary time series, Journal of Multivariate Analysis 14, 399 (1984).
[24] H. Pal, T. H. Seligman, and J. V. Escobar, Correlation networks from random walk time series, Physical Review E 98, 032311 (2018).
[25] J. A. Gubner, Probability and Random Processes for Electrical and Computer Engineers (2006).
[26] 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 (2018).
[27] M. Schuld, I. Sinayskiy, and F. Petruccione, Prediction by linear regression on a quantum computer, Phys. Rev. A 94, 022342 (2016).
[28] Y. Ahmadian, F. Fumarola, and K. D. Miller, Properties of networks with partially structured and partially random connectivity, Phys. Rev. E 91, 012820 (2015).
[29] J. Claerbout, Geophysical Image Estimation by Example (Lulu.com, 2014) google-Books-ID: cvszBwAAQBAJ.
[30] 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 (1994).
[31] D. W. Marquardt and R. D. Snee, Ridge regression in practice, The American Statistician 29, 3 (1975).
[32] 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 (2007).
[33] C. Lobry and T. Sari, Non-standard analysis and representation of reality, International Journal of Control 81, 519 (2008).
[34] 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.
[35] M. Fliess, C. Join, and F. Hatt, Volatility made observable at last. arXiv:1102.0683 [cs, q-fin] (2011) arXiv: 1102.0683.
[36] M. Fliess, C. Join, and C. Voyant, Prediction bands for solar energy: New short-term time series forecasting techniques, Solar Energy 166, 519 (2018).
[37] J. R. Trapero, Calculation of solar irradiation prediction intervals combining volatility and kernel density estimates, Energy 114, 266 (2016).
[38] 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 (2018).
[39] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, Mathematics for Machine Learning (Cambridge University Press, 2020) google-Books-ID: pbONxAEACAAJ.
[40] K. Krishnamoorthy, Handbook of Statistical Distributions (CRC Press, 2006) google-Books-ID: FEESD1tRI30C.
[41] R. J. Hyndman and Y. Fan, Sample Quantiles in Statistical Packages, The American Statistician 50, 361 (1996).
[42] P. Pinson, H. A. Nielsen, J. K. Möller, H. Madsen, and G. N. Kariniotakis, Non-parametric probabilistic forecasts of wind power: required properties and evaluation, Wind Energy 10, 497 (2007).
[43] L. Pan and D. N. Politis, Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions, Journal of Statistical Planning and Inference 177, 1 (2016).
[44] P. Lauret, M. David, and H. T. C. Pedro, Probabilistic Solar Forecasting Using Quantile Regression Models, Energies 10, 1591 (2017).
[45] P. Lauret, M. David, and P. Pinson, Verification of solar irradiance probabilistic forecasts, Solar Energy 194, 254 (2019).
[46] R. J. Hyndman and A. B. Koehler, Another look at measures of forecast accuracy, International Journal of Forecasting 22, 679 (2006).
[47] E. Spiliotis, V. Assimakopoulos, and S. Makridakis, Generalizing the theta method for automatic forecasting, European Journal of Operational Research 284, 550 (2020).
[48] D. George and P. Mallery, Spss for windows step-by-step: A simple guide and reference, 14th edition, http://gst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/[in=epidoc1.in]/?t2000=026564/(100) (2003).