Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models

Mathieu David, Faly H. Ramahatana Andriamasomanana, P.J. Trombe, Philippe Lauret

Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models

M. David[a*], F. Ramahatana[a], P.J. Trombe[b], P. Lauret[a]

[a]PIMENT, University of La Reunion, 97430 Tampon, Reunion

[b]DTU Compute, Technical University of Denmark, 2800 Kgs Lyngby, Denmark

Abstract

Forecasting of the solar irradiance is a key feature in order to increase the penetration rate of solar energy into the energy grids. Indeed, the anticipation of the fluctuations of the solar renewables allows a better management of the production means of electricity and a better operation of the grid-connected storage systems. If numerous methods for forecasting the mean of the solar irradiance were recently developed, there are only few works dedicated to the evaluation of prediction intervals associated to these point forecasts.

Time series of solar irradiance and more specifically of clear sky index show some similarities with that of financial time series. The aim of this paper is to assess the performances of a commonly used combination of two linear models (ARMA and GARCH) in econometrics in order to provide probabilistic forecasts of solar irradiance. In addition, a recursive estimation of the parameters of the models has been set up in order to provide a framework that can be applied easily in an operational context.

A comprehensive testing procedure has been used to assess both point forecasts and probabilistic forecasts. Using only the past records of the solar irradiance, the proposed model is able to perform point forecasts as accurately as other methods based on machine learning techniques. Moreover, the recursive ARMA-GARCH model is easier to set-up and it gives additional information about the uncertainty of the forecasts. Even if some strong assumption has been made regarding the statistical distribution of the error, the reliability of the probabilistic forecasts stands in the same order of magnitude as other works done in the field of solar forecasting.

Keywords: probabilistic solar forecasts; clear sky index; ARMA; GARCH; operational framework; recursive least square

1. Introduction

Forecasts of the power output of solar renewables are required to improve their penetration rate into electricity grids and also to ensure the security of the supply-demand balance. Indeed, accurate forecasts allow a better scheduling of the energy resources and a better operation of the units commitment. The recent development of grid-connected storages associated with intermittent renewables (solar, wind and wave) also requires power forecasts in order to optimize their operational management (Hernández-Torres et al., 2015)(Haessig et al., 2015)(Hanna et al., 2014). In the case of solar renewables, the power output is directly related to the level of received solar irradiance. Thus, forecasting the solar irradiance is the key feature of the PV power prediction.

* Corresponding author

Many tools and methods for forecasting solar irradiance were developed for various time horizons and spatial resolutions (Diagne et al., 2013). For day ahead forecasting, the most accurate methods are based on Numerical Weather Predictions (NWP) (Perez et al., 2013). For horizons of several hours, satellite based models are the most suitable (Kühnert et al., 2013)(Dambreville et al., 2014). And finally, for very short term horizons, from few minutes up to 2 hours, the literature is dominated by statistical approaches based on time series modeling (Reikard, 2009)(Lauret et al., 2015). Most of these works are dedicated to the development of point forecasts that is the mean solar irradiance over the considered time step. No forecast is perfect. Prediction intervals are necessary to assess the uncertainty associated to the point forecasts. In the field of wind power prediction, an extensive literature exists about probabilistic forecasting (Zhang et al., 2014)(Jung and Broadwater, 2014). But, in the solar energy domain, only few works were published recently in relation to probabilistic forecasts.

A first approach used numerical weather predictions (NWPs) models in order to derive the uncertainty of the solar forecasts. These models provide forecasts for a large range of time horizons, from 1 hour up to 72 hours. Iversen et al. (2014) proposed a stochastic differential equation framework with NWP for modeling the uncertainty associated with the point irradiance forecasts. Alessandrini et al. (2015) assessed uncertainty of power forecasts of solar plants with an analog ensemble method. The uncertainty of the forecasts is obtained by gathering past conditions corresponding to analog sets of a reduced number of parameters provided by NWPs. Mathiesen et al. (2013) proposed a postprocessing of different NWPs to obtain an ensemble probabilistic forecasting of the solar irradiance. The forecast system is based on studying the correlation of uncertainty to local meteorological conditions describing synoptic-scale atmospheric flow. Another simple approach has been proposed by Lorenz et al. (2009). A normal distribution with zero mean and a standard deviation dependent on the solar zenith angle and the cloud situation is assumed to describe forecast errors of solar irradiance. A last approach is based on the probabilistic behavior of time series. Bacher et al. (2009) used a weighted quantile regression conditioned to a normalized power to produce the uncertainty associated to the point forecasts obtained with an autoregressive model (ARX). They applied this model to forecast the power output of a small solar plant for horizon ranging from 1 hour to 1 day. Bracale et al. (2013) propose 1 hour ahead probabilistic forecasts of the clearness index. The method uses only ground measurements and it is based on a Bayesian autoregressive model associated with a Monte Carlo simulation to generate the density probability function of the error. Bessa et al. (2015) proposed an original method to provide 6 hours ahead probabilistic forecasts of the output power of a small fleet of PV systems. This last method is based on a vector autoregressive framework (VAR). It enables to assess the uncertainty of the forecast from a spatially distributed network of measurements.

This paper focuses on the very short term solar forecasting, from 10 minutes to 6 hours, applied to a single location. For these short horizons, the operational models should have low computational requirements in order to operate fast and the inputs must be available in real-time. The same issues are usually met in the financial domain. Furthermore, time series of solar irradiance recorded at high frequency (1-10 minutes) are composed of periods of low variability (clear sky or overcast sky) and periods with high variability (passing clouds). This behavior is also observed in financial time series that exhibit stable periods and clusters of volatility. This characteristic of times series is called heteroskedasticity (Engle, 1982).

The first aim of this work is to test models commonly used in econometrics. The point forecasts are generated with the well-known AutoRegressive Moving Average model (ARMA) (Tsay, 2005). The associated prediction interval is produced with a Genearalized AutoRegressive Conditional Heteroskedasticity model (GARCH)(Bollerslev, 1986). The combination of these two models has been

extensively tested in econometrics but to our best knowledge it is the first time in the field of solar forecasting. The second objective is to propose an operational framework of the forecasting methodology. More precisely only past data of the measured solar irradiance are used to run the models and the parameters of the two models are estimated with a recursive method (Ljung and Söderström, 1983).

The remainder of this paper is organized as follows. Section 2 describes the data used to build and to test the different models. Section 3 discusses the errors metrics used to assess the accuracy of the point and probabilistic forecasts. Section 4 depicts the structure of the different forecasting models with a special emphasis on the recursive estimation of the parameters. Section 5 details the results of point forecasts while section 6 gives the results of the probabilistic forecasts. Finally, Section 7 gives some concluding remarks.

2. Data

	Desert Rock (USA)	Fort Peck (USA)	Oahu (Hawaii)	Fouillole (Guadeloupe)	Saint-Pierre (Reunion)	Tampon (Reunion)
Provider	SURFRAD		NREL	LARGE	PIMENT	
Position	36.6N 119.0W	48.3N 105.1W	21.3N 158.1W	16.6N 61.5W	21.3S 55.5E	21.3S 55.5E
Elevation	1007m	634m	11m	6m	75m	550m
Climate type	Desert	Continental	Insular tropic	Insular tropic	Insular tropic	Insular heights
Years of records	2012 2013	2012 2013	2010 2011	2010 2011	2012 2013	2012 2013
Annual solar irradiance (MWh/m^2)	2.105	1.447	1.969	1.893	2.053	1.712
Ratio of missing or removed days	0.68%	1.91%	16.44%	0.82%	0%	15.21%
Ratio of interpolated data	0.28%	0.53%	0.00%	0.00%	0.05%	0.00%
Solar variability 10-min ($\sigma\Delta kt^*_{10min}$)	0.107	0.137	0.193	0.194	0.151	0.183
Solar variability 1-hour ($\sigma\Delta kt^*_{1hour}$)	0.146	0.181	0.209	0.213	0.191	0.241

Table 1: Main characteristics of the solar measurements used in this work

In this work, we used data recorded at 6 different locations in the world (see Table 1). These sites were selected because they offered high frequency measurements of global horizontal solar irradiance (GHI) for long periods. They also represent a large variety of climates. The time step of record is different between the considered stations but never exceed 1 minute. For this study we computed 10-min. and 1-hour averages of GHI directly from the raw data for 2 consecutive years. To get continuous and workable time series of GHI, we applied the following rules to treat the data:

- To avoid the side effects induced by the low accuracy of the solar measurements at sunrise and sunset, we remove all the data for a solar zenith angle superior to 80°. Thus, the obtained time

series do not contain the null night values. This filtering method removes less than 1% of the annual total solar energy.

- If a gap smaller than 20 minutes is detected, a linear interpolation is done in order to fill the missing data with the initial time step.
- If a gap longer than 20 minutes is detected, the whole day is removed from the time series.

The resulting data samples are continuous because the first data of the current day follow the last data of the previous day. Hence, the first forecasts of a day are done using the sky conditions observed before the sunset of the previous day. The measurements of the six selected stations are consistent and we detected only few missing data (table 1). The sites of Oahu and Tampon present a significant amount of missing days, around 15%, because their periods of record are shorter than two years.

A GHI time series can be divided into a deterministic part and a stochastic part. A clear sky model can accurately estimate the annual and diurnal cycles that represent the deterministic part (Iqbal, 2012). The aim of this work is to forecast the stochastic part induced by the cloud cover. Therefore, in order to remove the deterministic part from the radiation time series, we compute the clear sky index (kt^*) by dividing the measured GHI with the GHI observed under a clear sky (eq. 1). The resulting time series is almost stationary and a linear process, like an ARMA, can be used to model it.

(1) $$kt^* = \frac{GHI}{GHI_{Clear}}$$

In this work, we use the clear sky data provided by the McClear model (Lefèvre et al., 2013) available for free on the SoDa website ("SoDa," 2015). This model use the AODs, water vapor and ozone date from the MACC project.

Figure 1 plots the clear sky index distribution and table 1 (last rows) gives the solar variability. These two indicators allow a better understanding of the characteristics of the solar irradiance for each site. In figure 1, one can notice a significant number of occurrences of overirradiance, that is to say a clear sky index above 1. These cloud enhancement events can be observed everywhere in the world (Almeida et al., 2014). This phenomenon is more visible when using short time scales as in our study with 10-min mean of solar irradiance. The solar variability metric was defined by Perez et al. (2012) as the standard deviation of the changes in the clear sky index ($\sigma\Delta kt^*_{\Delta t}$). In our case we choose to compute the solar variability for a sampling intervals of 10 minutes and 1 hour that coincide with the time steps of our time series.

The continental stations of Desert Rock and Fort Peck experience a weather dominated by clear skies with a relatively low variability. For the 4 insular locations, the occurrences of clear skies are lower and the variability is higher than the continental sites. However some differences exist between them. For example, Oahu presents more occurrences of clear sky than Fouillole but these two locations have the same variability. Lauret et al. (Lauret et al., 2015) showed that the ability to generate an accurate point forecast with a linear process depends on the level of variability of the solar irradiance. Thus, the typology of the solar irradiance (i.e. radiation level, distribution and variability) of a site is a key factor to understand the level of error of a solar forecasting method.

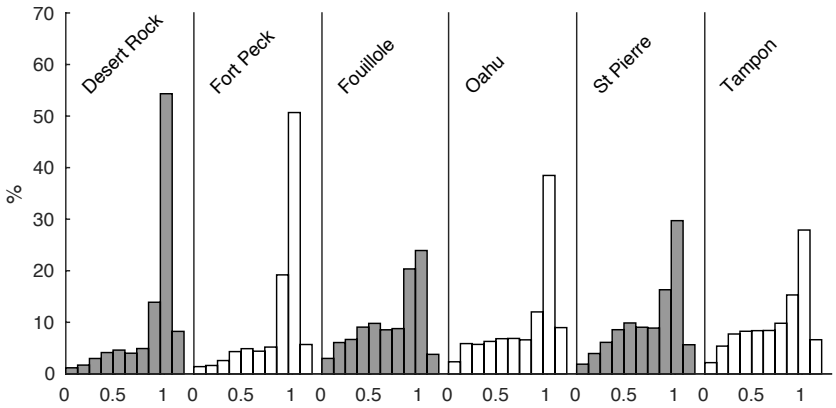


Figure 1: Statistical distribution of the 10-min. clear sky index for the 6 considered locations

3. Error metrics

Several attempts have been made to define the state-of-the art validation metrics for point forecasts (Hoff and Perez, 2012)(David et al., 2012)(Coimbra et al., 2013). In the realm of the solar forecasting community, the commonly used error metrics for point forecasts are: the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE) and the Forecast Skill (FS). The formulas used to evaluate these metrics are given in appendix C. In their work, David et al. (David et al., 2012) show that the RMSE and MAE metrics remain the best indicators. These absolute indices are usually normalized by the mean GHI of the test period as the utility industry wants to understand error in relative terms rather than in absolute terms (Hoff and Perez, 2012). In this work, in order to characterize the quality of the point forecasts, we provide the standard set of relative error metrics.

To evaluate and to rank the probabilistic forecasts, and following (Iversen et al., 2014), (Alessandrini et al., 2015), (Sperati et al., 2015), (Bessa et al., 2015) and (Pinson et al., 2010), we propose to use the reliability diagram, the rank histogram (or Talagrand histogram), the continuous ranked probability score (CRPS) and its associated skill score the CRPSS.

The reliability diagram is a well-known representation to verify a probabilistic forecast. We used the methodology defined by (Pinson et al., 2010) to create it. The latter is specially designed for the case of continuous variables commonly met in the field of energy meteorology. On the horizontal axis, the reliability diagram gathers the forecasts into bins according to their cumulative probability. The frequency with which the meteorological parameter was observed to occur for each bin of the forecasts probability is then plotted against the vertical axis (Hamill, 1997). This representation is appealing because the deviations from perfect reliability (the diagonal) can be visually assessed (Pinson et al., 2010).

The rank histogram is a tool for evaluating ensemble forecasts. They are useful for determining the reliability of probabilistic forecasts and for diagnosing errors in its mean and spread (Hamill, 2001). A rank histogram is generated by repeating the following two steps for every the time step of the verification sample: first sorting the members of the ensemble forecast and then tallying the rank of the corresponding observation. For a number of members N, the number of ranks of the histogram of an ensemble is $N+1$. A flat rank histogram shows that the members of an ensemble system are statistically indistinguishable from the observations (Alessandrini et al., 2015). A ∪-shape indicates a lack of spread while a ∩-shape indicates that the probabilistic forecasts are over-dispersive. In this work, we propose a

parametric model to generate the probabilistic forecasts. To plot the rank histograms, we created ensemble forecasts based on this model. The members of the ensemble are generated with a random variable that follow the distribution law assumed in the parametric model.

The CRPS is also a commonly used verification index in order to compare a full probabilistic distribution with the observations, when both are represented as cumulative distribution functions (CDF) (Sperati et al., 2015). This index contains a sense of distance of how far the forecast was found from reality (Hersbach, 2000). In addition, for a deterministic forecast, the CRPS is the mean absolute error (MAE) and has a clear interpretation. The formulation of the CRPS is:

(2) $\mathrm{CRPS} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \int_{-\infty}^{+\infty} \left[P_{fct}^{i}(x) - P_{obs}^{i}(x)\right]^2 dx$

where $P_{fct}^{i}(x)$ is the cumulative distribution function (CDF) of the forecast and $P_{obs}^{i}(x)$ is the CDF of the observation for the ith ensemble forecast/observation pair. *N* is the number of pairs. The CRPS has the same dimension as the forecasted variable and a lower value (0 is the perfect score) of the CRPS corresponds to a better performance. The interested reader can refer to Hersbach (2000) that gives a complete and detailed presentation of how to asses the CRPS. As the metrics used to assess the accuracy of the deterministic forecasts, the relative counterpart of the CRPS is obtained by dividing the absolute value of this metric by the average GHI of the test period.

As proposed by Bessa et al. (2015), the forecast skill of the proposed model is evaluated by computing the improvement over a reference model in terms of CRPS. This comparative metric will be called further in this work the Continuous Rank Probability Skill Score (CRPSS). The reference model used to derive the CRPSS is a persistence ensemble (see sub-section 4.1). The formulation of the CRPSS is given in appendix C.

4. Forecasting methods

In this section the forecasting models and the recursive method used to estimate their parameters are presented. To clarify the equations we used the following notation. A variable with a hat (^) is a forecast. *h* is the horizon of forecast also called the lead time. *t* denotes the time when the forecasts are generated. We choose to create one model per horizon of forecast in order to avoid the error propagations usually observed while running the same model iteratively. The models and the method of regression presented in the following sub-section are expressed taking into account this approach.

4.1 Reference models

Following Perez et al. (2010), we propose to test our forecasting method against reference models like persistence, smart persistence and climatology.

The persistence model is expressed as follows:

(3) $\widehat{kt^*}_{(t+h)} = kt^*_{(t)}$

This model assumes that the clear sky index for each time horizon 'h' only depends on the previous clear sky index values, which means that the sky conditions remain invariant between time '*t*' and time '*t*+*h*'. The next model represents an easy way to improve persistence model and it is called Smart Persistence ($\mathrm{Pers_{smart}}$). It consists in forecasting clear sky index for each time horizon '*h*' as the mean of the '*h*' previous clear sky values (Perez et al., 2014). Smart-persistence model is defined by the following equation:

(4) $\widehat{kt^*}_{(t+h)} = mean\left[kt^*_{(t)}, \dots, kt^*_{(t-h)}\right]$

We also propose the climatological mean model, which is independent of the forecast time horizon. More precisely, this model performs a constant forecast of the clear sky index that corresponds to its mean historical value:

(5) $\widehat{kt^*}_{(t+h)} = mean\ (kt^*)$

During this work, we used the mean clear sky index of the first year of records in order to forecast the second year.

Finally, we propose to use a persistence ensemble as reference model (Alessandrini et al., 2015) to make a comparison with our probabilistic forecasts. The members of the persistence ensemble are derived from the 10 last records of the clear sky index. Thus, the probability density function (PDF) of a forecast is equal the PDF obtained from the 10 most recent observations:

(6) $pdf\left[\widehat{kt^*}_{(t+h)}\right] = pdf(kt^*_t, kt^*_{t-1}, \cdots, kt^*_{t-9})$

4.2 ARMA model

The AutoRegressive Moving Average model (ARMA) is one of the most used time series models in forecasting. In particular, it has been extensively studied in renewable energy forecasting and, owing to its parsimony, it has turned out to be a very tough competitor. Applications already include among others forecasting of solar irradiance (Bacher et al., 2009). A general formulation of an ARMA(*p*,*q*) model with *p* autoregressive (AR) terms and *q* moving average (MA) terms is given by Tsay et al. (Tsay, 2005). Its application to the h-ahead forecast of the clear sky index is given by the following equation:

(7) $\widehat{kt^*}_{(t+h)} = \alpha_0 + \sum_{i=1}^{p} \alpha_i \cdot kt^*_{(t-i+1)} + \sum_{j=1}^{q} \beta_j \cdot \varepsilon_{(t-j+1)}$

with *h=1,2,...* the forecast horizon. The error term ε is the difference between the previous forecasts and observations as defined in the following equation:

(8) $\varepsilon_{(t-j)} = \widehat{kt^*}_{(t-j)} - kt^*_{(t-j)}$

More broadly, an ARMA model can be formulated in a canonical way as follows:

(9) $\widehat{kt^*}_{(t+h)} = \varphi_{(t)}{}^T \theta_{(t)}$

with $\theta_{(t)} = \left[\alpha_0, \alpha_0, \cdots, \alpha_p, \beta_1, \cdots \beta_q\right]$ the vector of parameters to be estimated and $\varphi_{(t)}{}^T = \left[1, kt^*_{(t)}, \cdots, kt^*_{(t-p+1)}, \varepsilon_{(t)}, \cdots, \varepsilon_{(t-q+1)}\right]$ the vector of inputs.

There are several methods to estimate the vector of parameters θ, with the two most widely implemented being the least squares (LS) and the Maximum Likelihood Estimation (MLE) methods. Here, we chose a variation of the LS method, namely the Recursive Least Squares (RLS) method. This method offers the advantage of reducing the computational cost for estimating the model with its parameters being updated in real-time as new data become available. This contrasts with more intensive estimation methods operating on a sliding window with the estimation process being carried out at each time step. The RLS method is particularly useful in an operational context where forecast have to be timely delivered. In the RLS method, θ is time varying and, by convention, it is subsequently denoted $\theta_{(t)}$. The RLS method is detailed in the following sub-section 4.4.

4.3 GARCH Model

The family of the ARCH (AutoRegressive Conditional Heteroskedasticity) models, introduced by Engle in the early 80's (Engle, 1982), is used to model the volatility (i.e. the variance) of time series in the financial domain. These models are particularly suitable to describe changes in variance over the time. This family of models is commonly used in econometrics in order to forecast the realized volatility of high frequency data (Andersen, 2000)(McAleer and Medeiros, 2008) or to assess the volatility of the error of point forecast models like linear regressions (AR, ARMA, etc.) (Bollerslev, 1986). In this work, we applied this second approach to compute the prediction intervals associated with the point forecasts generated by a recursive ARMA model.

Bollerslev (1986) proposed a Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model, which provides a more parsimonious representation than the simple ARCH model in many applications. GARCH models express the conditional variance as a linear function of lagged squared error terms and also lagged conditional variance terms (Taylor, 2004). The general formulation of a GARCH(p,q) model, with p error terms, q conditional variance terms and an horizon of forecast h, is given by the equations 10 and 11 (Tsay, 2005):

(10) $\hat{y}_{(t+h)} = \varepsilon_{(t)}\hat{\sigma}_{(t+h)}$

where ε is a random variable uniformly distributed with a zero mean and a unitary variance.

(11) $\hat{\sigma}^2_{(t+h)} = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon^2_{(t-i+1)} + \sum_{j=1}^{q} \beta_j \sigma^2_{(t-j+1)}$

with the error term $\epsilon_{(t-i)} = \hat{y}_{(t-i)} - y_{(t-i)}$.

The basic idea is that the series is either serially uncorrelated or with minor lower order serial correlations, but it is a dependent series. GARCH models attempt to capture such dependence in the return series (Tsay, 2005). Figure 2 shows a representation of the serial correlations of the time series of forecasting errors. These errors are obtained for the site of Saint-Pierre with a recursive ARMA(p,q) and a lead time of 10 minutes. Figure 2(a) plots the autocorrelation function (ACF) of the errors, which suggests no serial significant correlations. Figure 2(b) shows the ACF of the squared errors of forecast. This second plot suggests that the time series contains autocorrelation. Similar ACF plots and characteristics can be observed for other lead times and for the other sites. Bollerslev (1986) shows that such time series can be modeled with a GARCH(1,1). Other models of heteroskedasticity exist, like the ARCH model. With a deeper analysis of the correlation in the data, they could also be set up with the same objectives (Boland, 2015).

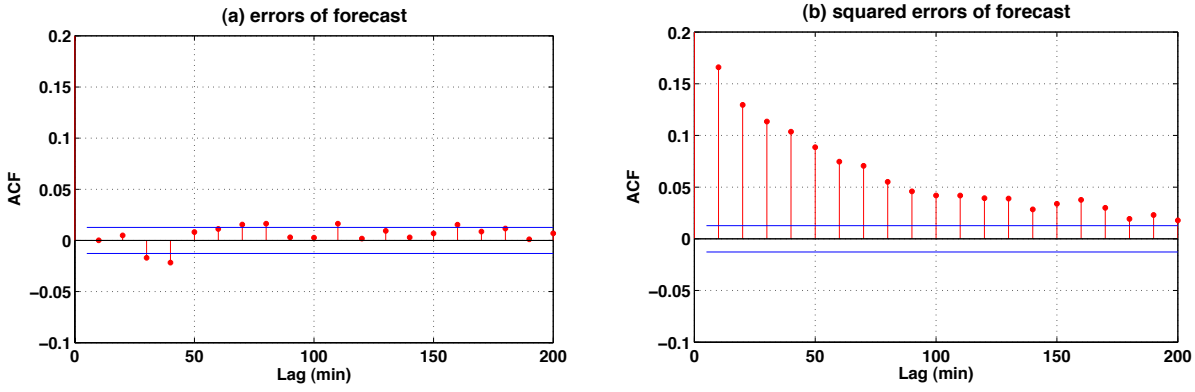


Figure 2: Sample ACF of the 10-min. horizon errors of forecast derived from a recursive ARMA for the site of Saint-Pierre ; (a) ACF of the errors ; (b) ACF of the squared errors

In order to define the confidence intervals of the point forecasts from the variance derived with the GARCH model, we assume a Gaussian distribution of the error. Thus, for each point forecast, the density of probability of the error corresponds to a normal law $N\left(\widehat{kt^*}_{(t+h)}, \hat{\sigma}_{t+h}\right)$. This assumption is not rigorously supported by the data. Figure 3 shows the statistical distribution of the error of forecast for two sites that present significant differences in sky conditions. The distribution of the errors for Desert Rock presents a strong peak around the mean (kurtosis) and a lack of symmetry (skewness). It cannot be considered as gaussian. The distribution of the errors for Fouillole is closer to a normal law but there are also some important differences.

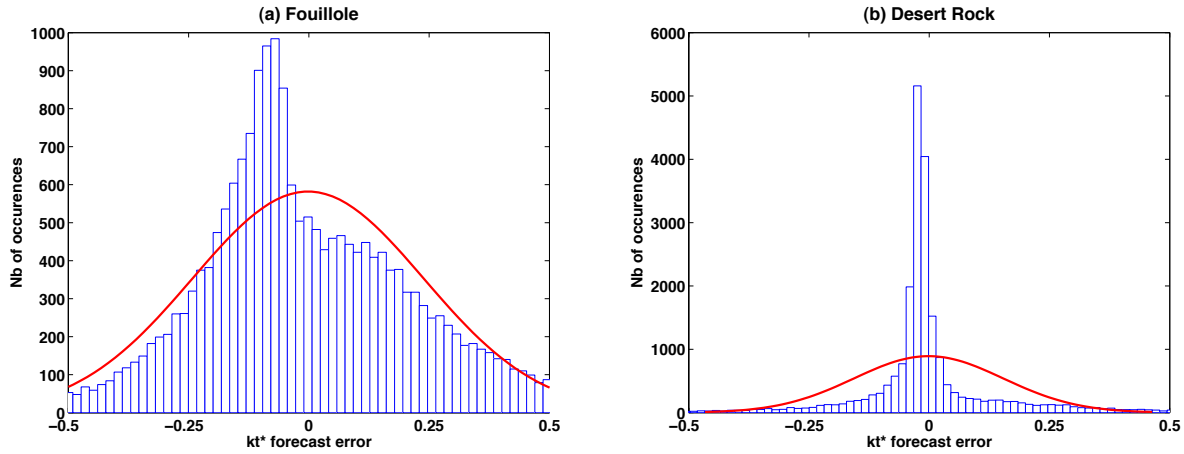


Figure 3: Statistical distribution of the errors of forecast (bar chart) derived from a recursive ARMA model with a horizon of 30 minutes for the sites of Fouillole (a), Desert Rock (b) and the best fit curve corresponding to a normal law (line).

Applied to the modeling of the forecasting error, equation 11 can be expressed with a canonical formulation as follows:

(12) $\hat{\sigma}^2_{(t+h)} = \varphi_{(t)}{}^T \theta_{(t)}$

with $\theta_{(t)} = \left[\alpha_0, \alpha_0, \cdots, \alpha_p, \beta_1, \cdots \beta_q\right]$ the vector of parameters to be estimated and $\varphi_{(t)}{}^T = \left[1, \epsilon^2_{(t)}, \cdots, \epsilon^2_{(t-p+1)}, \sigma^2_{(t)}, \cdots, \sigma^2_{(t-q+1)}\right]$ the vector of inputs. The error term ϵ is the difference between the point forecast and the observation of the clear sky index at time *t* (eq. 13).

(13) $\epsilon_{(t)} = \widehat{kt^*}_{(t)} - kt^*_{(t)}$

A special issue concerning the GARCH model is the estimation of the past values of the variance ($\sigma^2_{(t)}$, …, $\sigma^2_{(t-q+1)}$). The ARMA process generates only one point forecast per time step and it is not possible to assess the variance of past errors. Engle define ARCH processes as time series with non-constant variances conditional on the past, but constant unconditional variances (Engle, 1982). This unconditional variance σ^2 is obtained by applying unconditional expectation to $\hat{\sigma}^2_{(t+h)}$ in equations 11 or 12 (González-Rivera, 1998). The unconditional variance is a function of the parameters $\alpha_i$and β_i of a GARCH model and it can be used to initialize the GARCH process (i.e. $\sigma^2_{(t)} = \sigma^2$ in eq. 11) (Nelson and Cao, 1992). Equation 14 gives the unconditional variance for a GARCH(1,1) process:

(14) $\sigma^2 = \frac{\alpha_0}{1-\alpha_1-\beta_1}$

In this work, we build a GARCH(1,1) model as proposed by Bollerslev (1986). The estimation of the vector of parameters θ of the GARCH model is done with the same recursive method used for the ARMA model.

4.4 Recursive estimation of the parameters

We consider a time series $y_{(t)}$ that can be modeled by a linear regression process:

(15) $y_{(t)} = {\varphi_{(t)}}^T \theta_{(t)} + \eta_{(t)}$

where $\theta_{(t)}$ is the vector of parameters and $\varphi_{(t)}$ the vector of inputs at time *t*. $\eta_{(t)}$ is a random variable with a normal law distribution. A forecast of the time series $\hat{y}_{(t+h)}$ can be done with the same linear regression model considering a horizon of forecast *h* corresponding to *h* time steps:

(16) $\hat{y}_{(t+h)} = {\varphi_{(t)}}^T \theta_{(t)} \leftrightarrow \hat{y}_{(t)} = {\varphi_{(t-h)}}^T \theta_{(t-h)}$

We set up a recursive least square (RLS) to estimate the vector of parameters $\theta_{(t)}$ of this time series. For this work, the RLS methodology has been slightly modified to take into account the different time horizons *h*. For instance, the update step of the recursive process must be lagged in order to take into account that the forecasted $\hat{y}_{(t)}$ was derived at time *t-h*. The adaptation of the RLS method integrates a different time lag between the current time and the forecasted time. In the equation 17 and 18, the gain matrix $P_{(t)}$ is assessed using a lagged input vector at time *t-h*. This change from the initial formulation takes into account that the forecast of the time series at time *t* was done at time *t-h*.

(17) $\theta_{(t)} = \theta_{(t-1)} + P_{(t)}\varphi_{(t-h)}\left(y_{(t)} - {\varphi_{(t-h)}}^T \theta_{(t-1)}\right)$

(18) $P_{(t)} = \frac{1}{\lambda}\left(P_{(t-1)} - \frac{P_{(t-1)}\varphi_{(t-h)}{\varphi_{(t-h)}}^T P_{(t-1)}}{\lambda + {\varphi_{(t-h)}}^T P_{(t-1)}\varphi_{(t-h)}}\right)$

where $P_{(t)}$ is the gain matrix and λ is a forgetting factor. For this work, we use a forgetting factor λ = 0.999.

For the ARMA model, the RLS method was used in a classical way (Ljung and Söderström, 1983) like Madsen (2007) for the estimation of the parameters of their ARX model. The methodology directly minimizes the quadratic error between the observations $y_{(t)}$ and the forecasts $\hat{y}_{(t)}$. For the GARCH model, it is not possible to directly apply the RLS method. The prediction error is different from the GARCH residual. Kierkegaard et al. (2000) transform the residual by making the transformation $\sigma_t^2 = \epsilon_t^2$ in equation 17 in order to assess the GARCH parameters with a pseudo-linear recursive method. This transform works for any situations where the conditional variance is independent from the sign of the prediction error ϵ_t. We use this convenient transform to run our RLS method.

Even if the cyclical effects of the solar path are removed by the use of the clear sky index, the time series still present short-term trends (hurricanes, cold fronts, anticyclone, etc.). It seems important that the method used to assess the vector of parameters takes into account these short-term trends. The recursive methodology with forgetting is able to catch these short-term trends.

4.5 Machine learning techniques

Regarding the point forecasts, we also propose to compare the results obtained with the recursive ARMA with those obtained by nonlinear machine learning techniques (see section 5). In a previous work (Lauret et al., 2015), three machine learning techniques namely Artificial Neural Networks (NNs), Gaussian Processes (GPs) and Support Vector Machines (SVMs) were employed for intraday solar forecasting. These machine-learning techniques are supervised learning methods or data-driven approaches. As a consequence, these techniques rely on the information content embedded in the training data in order to produce forecasts on unseen data. More precisely, the models' parameters are determined with the help of *n* pairs of input and output examples contained in the training data. Once the model is fitted, the model can be evaluated on a test dataset. This second phase, called the generalization phase, consists of evaluating on the test dataset, the ability of the technique to generalize, that is to say, to give correct outputs when it is confronted with examples that were not seen during the training phase.

In our context, $\mathcal{D} = \{x_i, y_i\}_{i=1}^{n}$ represents the training dataset. The vector x_i contains the *p* past values of the clear sky index for training and y_i refer to the corresponding value of the clear sky index for the horizon *h* of interest. The column vector inputs for all *n* training cases can be aggregated in the so-called n×p design matrix X and the corresponding model's outputs (or targets) are collected in the vector y so we can write $\mathcal{D} = \{X, y\}$. Similarly, considering n^* test cases, we have $\mathcal{D}^* = \{X^*, y^*\}$ for the test dataset.

These methods are briefly described below and the interested reader can refer to (Lauret et al., 2015) for details regarding the optimization of these types of models.

4.5.1 Neural networks (NNs)

A NN with *d* inputs, m hidden neurons and a single linear output unit defines a non-linear parameterized mapping from an input vector **x** to an output y. For our application, the prediction for an input test vector $\mathbf{x}^*$ is given by the following equation:

(19) $\widehat{kt^*}_{(t+h)} = y(x; w) = \sum_{j=1}^{m} w_j \, f\left(\sum_{i=1}^{d} w_{ji} \, \mathbf{x}^* + b_1\right) + b_2$

Each of the *m* hidden units are related to the tangent hyperbolic function $f(x) = (e^x - e^{-x})/(e^x + e^{-x})$. The parameter vector $\mathbf{w} = \left(\{w_j\}, \{w_{ji}\}, b_1, b_2\right)$, which contains a set of weights $\{w_j\}, \{w_{ji}\}$ and two biases b_1, b_2, governs the non-linear mapping and is estimated during the training phase.

4.5.2 Gaussian Processes (GPs)

GPs are stated as a kernel-based method. Indeed, it can be shown (Rasmussen and Williams, 2006) that, given *n* training samples, the prediction for an input test vector $\mathbf{x}^*$ can be seen in terms of a linear combination of *n* kernel functions; each one centered on a training point. Therefore, the forecasted clear sky index is given by Eq. (20):

(20) $\widehat{kt^*}_{(t+h)} = \sum_{i=1}^{n} \alpha_i \, k_f(\mathbf{x}_i, \mathbf{x}^*)$

where k_f denotes the squared exponential covariance function $k_f(x_p, x_q) = \sigma_f^2 \exp\left[\frac{-(x_p - x_q)^2}{2l^2}\right]$ and $\mathbf{x}_i$ is the ith input training vector. The coefficients α_i are estimated during the training phase. σ_f^2 and *l* are called hyperparameters of the covariance function. These hyperparameters control the model complexity and can be learned (or optimized) from the training data at hand (Rasmussen and Williams, 2006).

4.5.3 Support vector machines (SVMs)

The support vector machine (SVM) is another kernel based machine learning technique used in classification tasks and regression problems (Cortes and Vapnik, 1995). Support vector regression (SVR) is based on the application of support vector machines to regression problems (Smola and Schölkopf, 2004). This method has been successfully applied to time series forecasting tasks (Muller et al., 1997). In a similar manner as for the GPs, the prediction calculated by a SVR machine for an input test case $\mathbf{x}^*$ is given by Eq. (21):

(21) $\widehat{kt^*}_{(t+h)} = \sum_{i=1}^{n} \alpha_i \, k_{rbf}(\mathbf{x}_i, \mathbf{x}^*) + b$

k_{rbf} denotes the radial basis covariance function $k_{rbf}(x_p, x_q) = \exp\left[-\gamma\left|(x_p - x_q)\right|\right]$ with hyperparameter γ and b is a bias parameter.

In the case of SVMs (unlike GPs), it must be noted that not all the training patterns participate to the preceding relationship. Indeed, a convenient setting of the optimization problem enables to obtain a sparse solution. The latter means that only some of the coefficients α_i will be nonzero. The examples that come with non-vanishing coefficients are called Support Vectors.

5. Results for the point forecasts

The 2 years of data of the six considered sites are split into two sets as follow: the first year is used as training sample and the second year is used as test sample. All the metrics provided in this section are derived from the test samples (i.e. the second years of data). In the case of the recursive estimation of the parameters of the ARMA models, a calibration or training period is not required. Thus, to select the best orders of the ARMA(*p*,*q*) models, we directly used the first years of data of every sites. We selected the pairs of AR and MA lags that presented the lowest RMSE for each site and each forecasting horizon. To find this minimum, we varied the lags in the range [1;10]. Table 2 gives the AR and MA orders resulting from the cross validation. For a 10-min. granularity and for the sites with a low variability (i.e. Desert Rock, Fort Peck and Saint-Pierre), the AR and MA orders present the same evolution. As shown by Table 2, the AR order increases and the MA order decreases when the forecasting horizon is increasing. Obviously, for these sites that present a relatively low variability in the 10-min time scale, very short-term forecasts (horizons inferior to 30 minutes) are mainly influenced by the trend of the change of the sky condition. For the sites with a higher variability and also for a granularity of 1 hour, no clear trend was seen in the values of the orders.

Granularity	Horizon	Desert Rock		Fort Peck		Fouillole		Ohau		Saint-Pierre		Tampon	
		AR	MA	AR	MA	AR	MA	AR	MA	AR	MA	AR	MA
10-min.	10 min.	1	3	1	3	3	2	2	2	1	4	1	3
	20 min.	2	2	2	2	2	3	2	2	2	4	2	2
	30 min.	2	1	3	2	2	4	4	4	3	2	3	3
	40 min.	3	2	3	1	4	1	3	1	4	1	2	4
	50 min.	4	1	4	1	5	1	5	1	5	1	1	7
	60 min.	6	1	5	1	5	2	4	1	5	7	2	5

1-hour	1 hour	1	1	1	1	1	1	1	1	1	1	1	1
	2 hours	2	1	1	1	2	2	4	1	2	1	9	1
	3 hours	1	1	1	1	3	2	1	2	7	5	8	2
	4 hours	2	2	3	3	2	1	1	1	7	2	9	1
	5 hours	1	1	1	1	5	1	5	2	6	1	8	1
	6 hours	1	1	1	1	5	1	1	7	6	1	7	1

Table 2: Orders of the recursive ARMA(p,q) models for different lead times, granularities and sites

The detailed results of the simulations can be found in the tables A.1 to A.4 of the appendix A. These tables give the values of the error metrics for every site, horizon and granularity. The proposed method presents a very low bias (MBE), inferior to 2% for all sites. Figures 4 and 5 show the relative RMSE of the point forecasts respectively for the 10-min. granularity and for the 1-hour granularity. In terms of quadratic errors, the accuracy of the forecasts decreases when the lead-time increases. The recursive ARMA models outperform the two reference models (i.e. smart persistence and climatology) for all the horizons and granularities. As mentioned above, a comparison is also done with the machine learning techniques briefly described in section 4.5. The Gaussian Process technique from this previous work was not used for the 10-min. granularity because it led to a very long computation time. In terms of quadratic errors, the recursive ARMA slightly outperforms the machine learning models (Gaussian Process, Support Vector Machine and Neural Network) for a granularity of 1 hour. But for a granularity of 10 minutes and for the other error metrics (see tables 5 and 6 in appendices), the accuracy of the recursive ARMA is practically the same as the machine learning methods. The main advantage of the proposed ARMA model compared to the machine learning methods comes from the simplicity provided by the recursive estimation of the parameters. Indeed, this method does not require training period and the parameters are estimated only with the last available records using a simple algorithm. Moreover, an ARMA model has only few parameters to estimate and it does not require complicated method to manage its complexity. Hence, it runs very fast and it is easier to set up for an operational use.

The accuracy of the forecasts also depends on the sky conditions observed on a site. There is a clear relationship between the indices of variability ($\sigma\Delta kt^*_{10min}$ and $\sigma\Delta kt^*_{1hour}$) and the ability to forecast accurately the solar irradiance. The stations of Oahu, Fouillole and Tampon experience the most variable sky conditions and they exhibit the less accurate forecasts. On the contrary, Desert Rock has the lowest index of variability and the accuracy of the forecasts is the best for this site. The more variable the sky conditions are, the less predictable the solar irradiance is. This relationship has already been highlighted in recent works (Lauret et al., 2015)(Voyant et al., 2015).

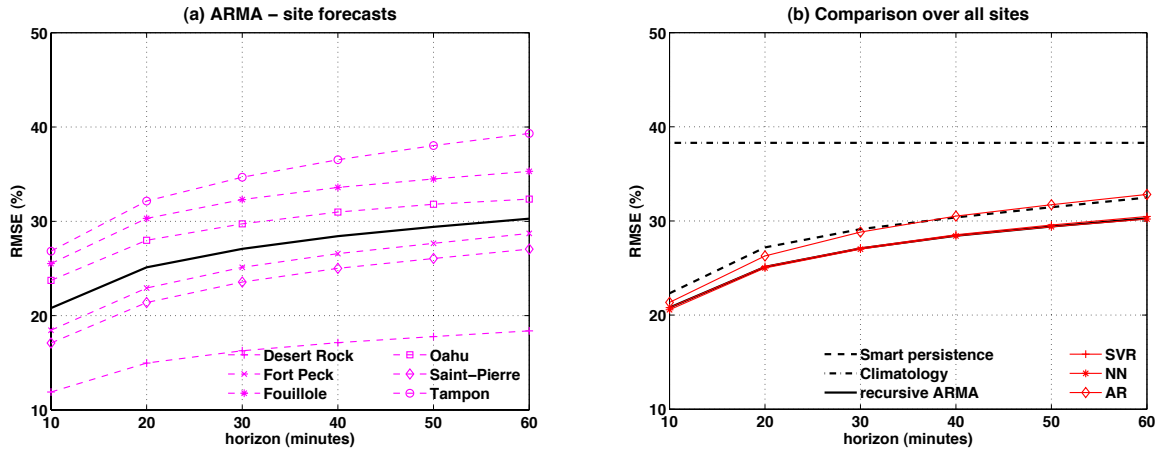


Figure 4: Relative RMSE for horizons of forecast from 10 minutes to 60 minutes and a granularity of 10 minutes. In (a), the dashed lines are the error of the recursive ARMA model for each individual site. The bold line is the average RMSE of the recursive ARMA model considering all the sites. (b) Comparison of the reference models, the recursive ARMA and the models proposed in (Lauret et al., 2015) considering all the sites.

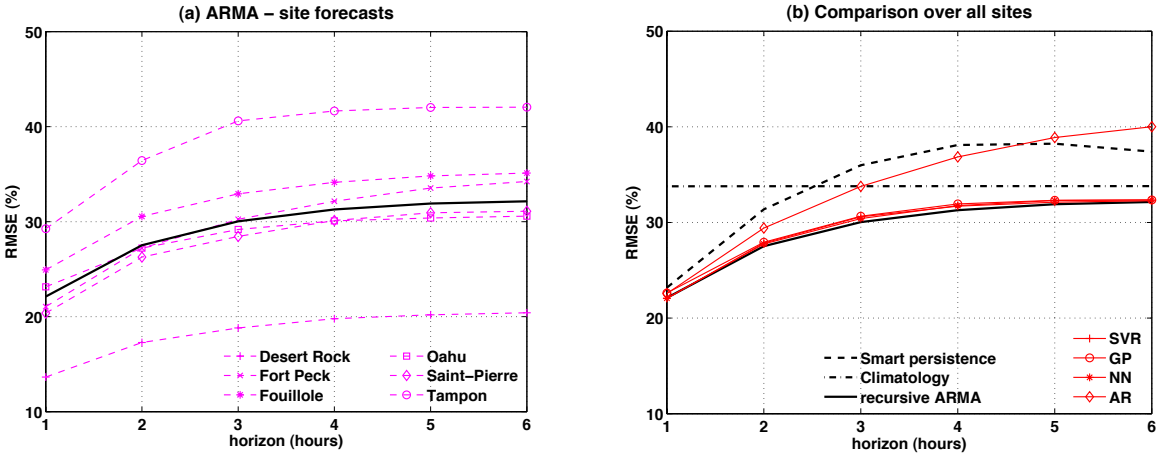


Figure 5: Relative RMSE for horizons of forecast from 1 hour to 6 hours and a granularity of 1 hour. In (a) the dashed lines are the error of the recursive ARMA model for each individual site. The bold line is the average RMSE of the recursive ARMA model considering all the sites. (b) Comparison of the reference models, the recursive ARMA and the models proposed in (Lauret et al., 2015) considering all the sites.

6. Results for the probabilistic forecasts

An uncertainty envelope is associated to the point forecasts thanks to the GARCH(1,1) model (Figure 6). This probabilistic information is derived from the forecasts of the standard deviation of the error between the point forecasts and the observations. To generate the uncertainty, we assume that the distribution of the error of forecast follows a normal law. To assess the ability of the model to reproduce the distribution of his own error, we first use the well-known reliability diagram. This representation allows to graphically assess the reliability for a set of probabilistic forecasts. An example of a reliability diagram is given in

figure 7 for probabilistic forecasts done for the site of Tampon. The dashed line represents a perfect assessment of the errors by the model itself. As shown in figure 7, the probabilistic forecasts do not follow exactly the ideal line. For example, 54% of the observed errors occurred when forecast probability was 60%. In this case, the model is over confident. The probability of the forecasts is higher than the probability of the observations.

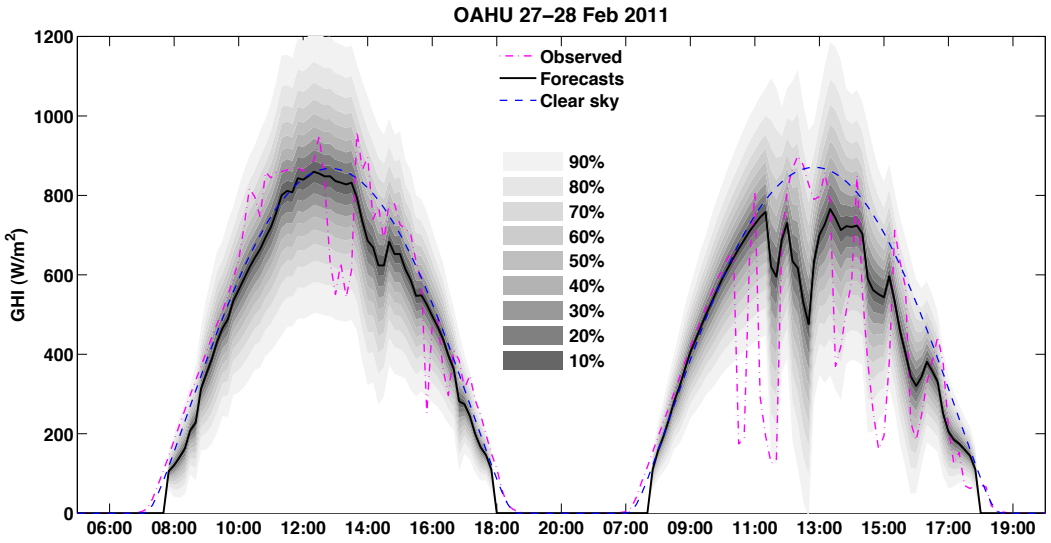


Figure 6: Example of point forecasts with confidence intervals of the GHI for 2 days in Oahu generated with the recursive ARMA-GARCH model. The forecasting time horizon is 1 hour and the granularity is 10 minutes.

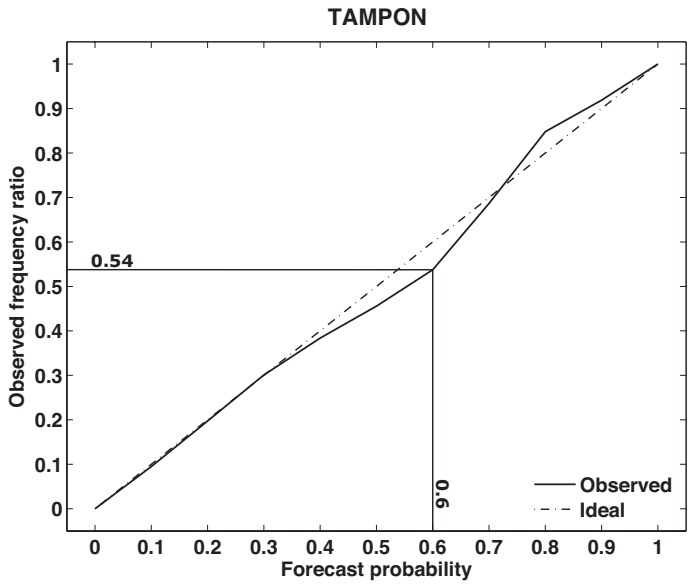


Figure 7: Reliability diagram of the probabilistic forecasts for Tampon with a forecasting time horizon of 1 hour and a granularity of 1 hour

Figure 8 provides respectively the reliability diagrams for the 10-min. (a) and the 1-hour (b) granularities. For all the horizons, the reliability of the ARMA-GARCH models is almost the same. Around a cumulative probability of forecasts of 70%-80% the reliability of the model is changing. Below this threshold, the model is over confident. Inversely, above this threshold, the model is under confident. Strangely, the model based on stochastic differential equations developed by Iversen et al. (2014) shows a similar behavior. In our case, this deviation of the probabilistic forecasts to the ideal line is mainly due to the assumption done about the Gaussian distribution of the error. Figure 9 gives the effect of this assumption on the reliability of the probabilistic forecast. The distance between a Gaussian distribution and the distribution of the error of the forecasting method is measured with the help of the Kolmogorov-Smirnov index (KSI). The KSI for each site is given in the table at the right part of figure 9. The smaller the KSI is, the closer the distribution to a normal law is. The distribution of the forecasting error of the site of Desert Rock is the most distant from a normal law. The worst reliability is also observed for this site. Inversely, the site of Tampon presents the closest distribution to a normal law and the best reliability of our test samples. Thus, the assumption done about the statistical distribution of the error of the point forecasts affects directly the reliability of the probabilistic forecast done with our GARCH model. In our case, the sites that exhibit the lowest error for the point forecasts have the worst reliabilities. Thus, the higher the solar variability is, the better the reliability of the probabilistic forecasts is.

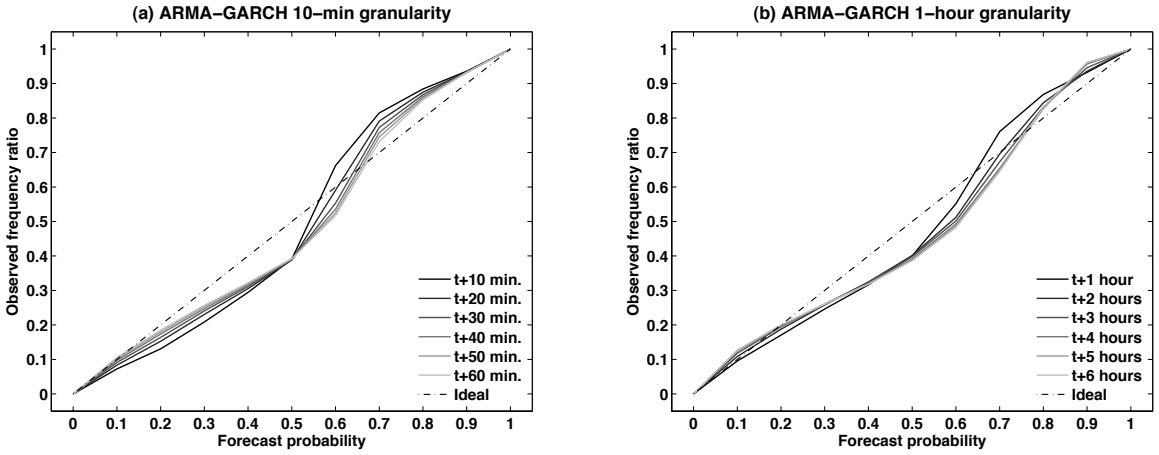
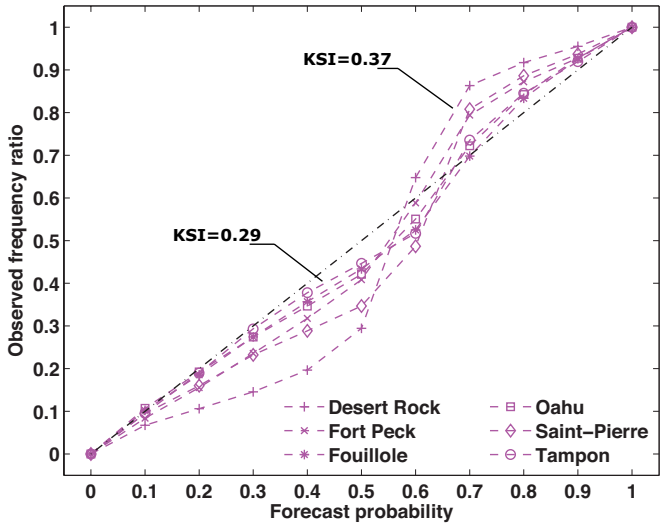


Figure 8: Reliability diagrams of the recursive ARMA-GARCH model considering all the sites. (a) Horizons of forecast from 10 minutes to 60 minutes and a granularity of 10 minutes. (b) Horizons of forecast from 1 hour to 6 hours and a granularity of 1 hour.



	KSI
Desert Rock	0.37
Fort Peck	0.33
Fouillole	0.31
Oahu	0.31
Saint-Pierre	0.33
Tampon	0.29

Figure 9: Reliability diagrams of the recursive ARMA-GARCH for a horizon of 30 minutes and a granularity of 10 minutes. The one-sample Kolmogorov-Smirnov index (KSI) for a correspondence with a normal law is given in the table for the error of forecasts on the kt.*

To complete the analysis done above about the reliability, we plot rank histograms to assess the consistency of our method. Figures 10 and 11 show the rank histograms of our ARMA-GARCH models and of the persistence ensemble forecasts. The rank histogram is initially designed to assess the spread of an ensemble forecast system. In the case of a parametric method, like our approach, a rank histogram can be done by generating an ensemble of forecasts from a random variable that follow the selected distribution. As for the persistence ensemble, we choose to generate 10 members that follow our assumption of normal distribution of the error. The mean comes from the ARMA model and the variance comes from the GARCH model. With a U-shape, the rank histograms of the persistence ensemble show clearly a lack of spread in the forecasts. With a peak of population on the right side, the shape of the rank histograms of our ARMA-GARCH model shows an excess of variability with a negative bias. We can suppose that these bias results from the symmetry of the PDF of the normal law. Indeed, if the sky is clear at time t, at time t+h the probability that the sky will be “less clear” is higher than a probability to be “more clear”. This reality is not supported by our assumption of a Gaussian distribution of the error of forecast.

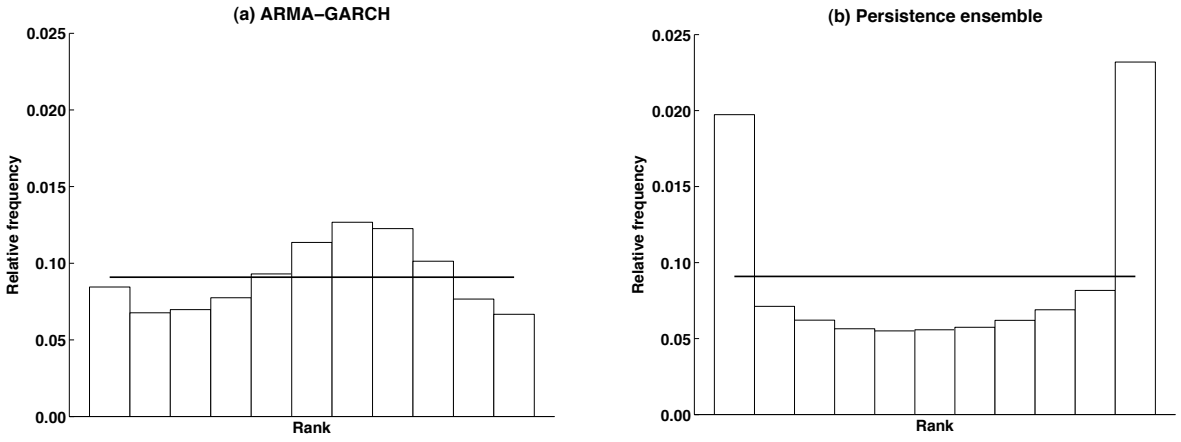


Figure 10: Rank histograms computed over all the lead times of the ARMA-GARCH model (a) and of persistence ensemble (b) for a 10-min granularity.

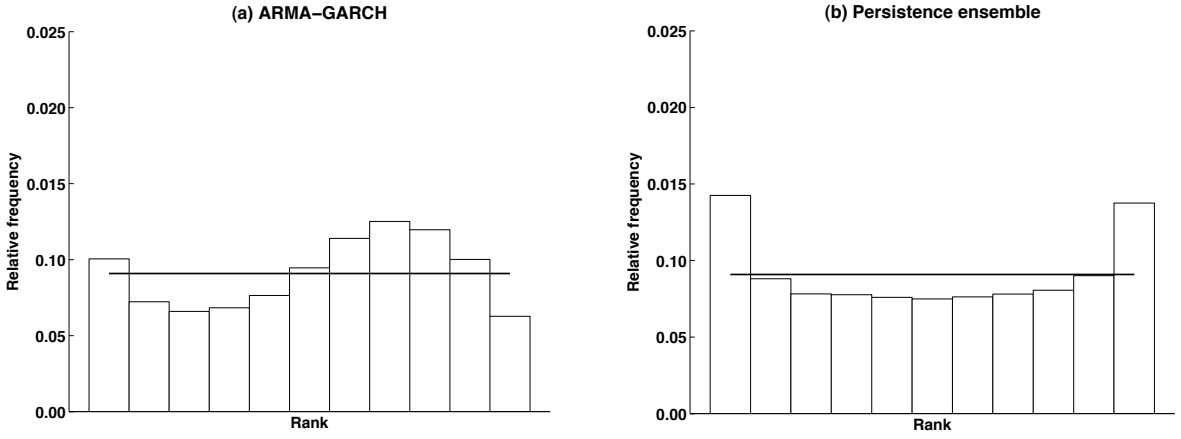


Figure 11: Rank histograms computed over all the lead times of the ARMA-GARCH model (a) and of persistence ensemble (b) for a 1-hour granularity.

The last investigated error metrics are the CRPS and the associated skill score the CRPSS. The CRPS - combines both the error of the point forecasts and the reliability of the probabilistic quantiles. Figure 12 (Tables B.1 and B.2 in appendix B) gives the CRPS of each site for all the granularities and all the horizons. As the CRPS is equivalent to the MAE for a deterministic forecast, its trend is strongly influenced by accuracy of the point forecasts. Again, we can observe that the accuracy of the forecasts is directly dependent to the solar variability of the site. Even if the reliability of the probabilistic forecasts differs between the different sites, it affects very weakly the values of the CRPS. In figure 13, the CRPSS gives the improvement of the ARMA-GARCH model compared to the persistence ensemble. Overall, the ARMA-GARCH model outperforms the persistence ensemble. This improvement is higher for the shortest horizons of forecast. There is also a difference of improvement between the different sites. Again, the variability of the observed sky conditions is the key factor of these differences. By definition, the

persistence ensemble has a good accuracy when the sky conditions remain stable. Thus, the improvement of our model is better for the sites that experience a higher variability.

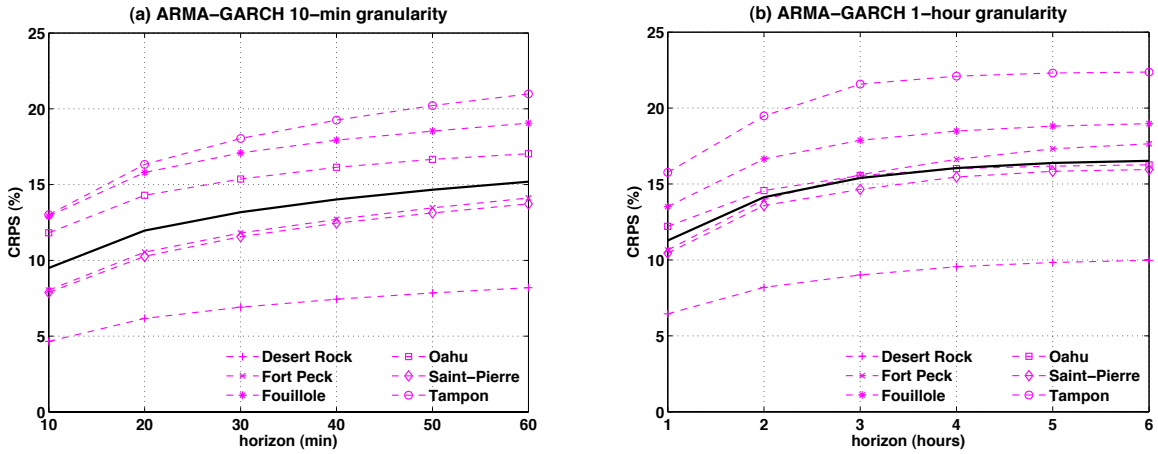


Figure 12: CRPS of the recursive ARMA-GARCH. The dashed lines are the CRPS for each individual site. The bold line is the average CRPS considering all the sites. (a) Horizons of forecast from 10 minutes to 60 minutes and a granularity of 10 minutes. (b) Horizons of forecast from 1 hour to 6 hours and a granularity of 1 hour.

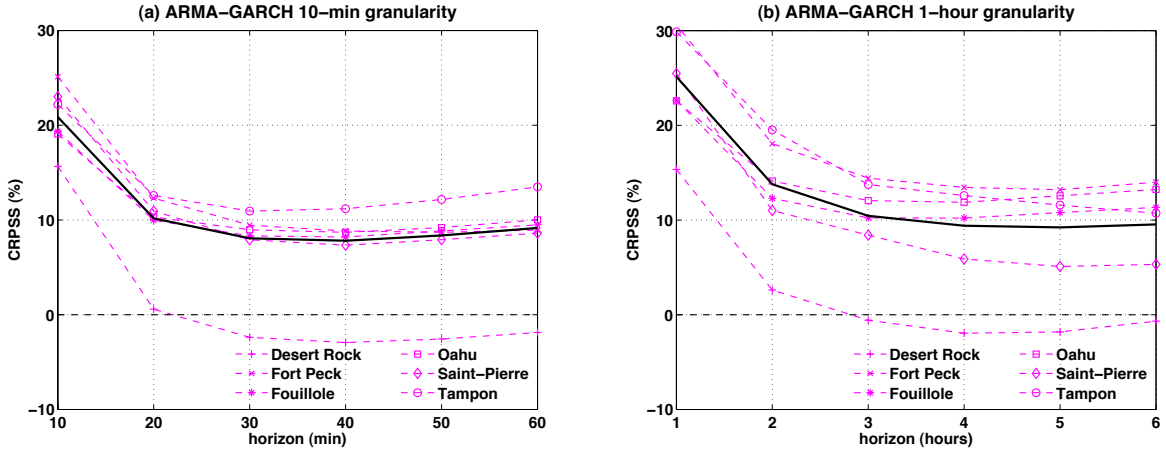


Figure 13: CRPSS of the recursive ARMA-GARCH. The dashed lines are the CRPSS for each individual site. The bold line is the average CRPSS considering all the sites. (a) Horizons of forecast from 10 minutes to 60 minutes and a granularity of 10 minutes. (b) Horizons of forecast from 1 hour to 6 hours and a granularity of 1 hour.

The results obtained for the analysis of the probabilistic forecasts are coherent with the formulation and the behavior of a GARCH model. For instance, a GARCH process is efficient to model time series that tend to "return to the mean" after an event of strong volatility. This mean volatility is also the unconditional volatility of the time series. For our time series of clear sky indices, the mean volatility (or variance) is site dependent and it has a value around 0.2 – 0.3. These mean values do not correspond to a

stable sky condition (i.e. clear sky or overcast) that exhibits an almost null variance. Thus, for stable sky conditions, the GARCH process tends to overestimate the volatility of the error.

7. Conclusions and outlook

In order to provide very short-term forecasts of the solar irradiance, operational tools with low computational requirements must be developed. The approach proposed in this work is built on models used for the high frequency trading in the financial domain. The ARMA-GARCH methodology is an efficient combination of models used to generate very short-term point forecasts of solar irradiance with confidence intervals. Furthermore the recursive estimation of the parameters of these two models facilitates the implementation of these methods in an operational framework. In other words, this approach can be set up using only the real time monitoring of the solar power.

The recursive ARMA model offers a simple and operational tool to generate point forecasts. Using only the actual value of the solar irradiance, this technique outperforms other statistical models as applied in this work. The association of an ARMA model with a GARCH model allows assessing the prediction intervals associated with the point forecasts. This combination is an alternative to the ensemble approach commonly used for generating probabilistic forecasts. Furthermore, a recursive ARMA-GARCH needs lower computational requirement than an ensemble method based on NWP and it is more suitable for very short term forecasting.

In this work, a normal distribution of the error of the point forecasts has been assumed. In reality, the error of the point forecasts does not fit with a Gaussian law. Following this assumption, the model is over confident when the forecast probability is below 60%-70% and under confident above this threshold. An improvement of the method could be achieved by assuming a distribution law that fits better to the data (e.g. a Student law).

New tools are currently developed in order to provide real time monitoring of the sky conditions like fish-eye cameras. The proposed framework can easily integrate exogenous inputs (e.g. through the use of an ARMAX model) in order to improve the quality of the point forecasts. Thus, the recursive ARMA-GARCH model is not only an operational model but it is also an upgradeable tool.

Acknowledgement

The authors would like to thank the laboratory LARGE from the University of Guadeloupe, the National Renewable Energy Laboratory (NREL) and the meteorological network SURFRAD for providing their ground measurements.

This work was supported by the French Environment and Energy Management Agency (ADEME) under the SOLFIN project.

Appendix A: MBE, RMSE and MAE of the point forecasts

	Horizon	Desert Rock (mean = 549W/m²)	Fort Peck (mean = 395W/m²)	Fouillole (mean = 515W/m²)	Ohau (mean = 492W/m²)	Saint-Pierre (mean = 538W/m²)	Tampon (mean = 458W/m²)
MBE	10 min.	-0.1%	-0.2%	-0.3%	0.5%	-0.2%	0.4%
	20 min.	-0.2%	-0.3%	-0.6%	0.6%	-0.3%	0.7%
	30 min.	-0.3%	-0.5%	-0.8%	1.0%	-0.4%	1.0%
	40 min.	-0.3%	-0.6%	-0.9%	1.3%	-0.5%	1.1%
	50 min.	-0.4%	-0.7%	-1.1%	1.3%	-0.6%	1.2%
	60 min.	-0.4%	-0.8%	-1.3%	1.4%	-0.6%	1.2%
RMSE	10 min.	12.0%	18.5%	25.5%	23.7%	17.1%	26.8%
	20 min.	15.0%	22.9%	30.3%	28.0%	21.4%	32.2%
	30 min.	16.3%	25.1%	32.3%	29.7%	23.6%	34.7%
	40 min.	17.1%	26.6%	33.6%	31.0%	25.0%	36.5%
	50 min.	17.8%	27.7%	34.5%	31.8%	26.1%	38.0%
	60 min.	18.4%	28.7%	35.3%	32.4%	27.1%	39.3%
MAE	10 min.	5.4%	10.3%	17.6%	16.1%	10.1%	17.6%
	20 min.	7.4%	13.9%	22.1%	19.8%	13.6%	22.8%
	30 min.	8.6%	15.8%	24.2%	21.5%	15.6%	25.7%
	40 min.	9.3%	17.2%	25.6%	22.8%	17.1%	27.7%
	50 min.	9.9%	18.4%	26.7%	23.6%	18.2%	29.3%
	60 min.	10.4%	19.4%	27.5%	24.3%	19.2%	30.7%

Table A.1: Relative error metrics of the point forecasts done for each site by the recursive ARMA models using a granularity of 10 minutes

	Horizon	Desert Rock (mean = 548W/m²)	Fort Peck (mean = 391W/m²)	Fouillole (mean = 510W/m²)	Ohau (mean = 499W/m²)	Saint-Pierre (mean = 538W/m²)	Tampon (mean = 456W/m²)
MBE	1 hour	-0.2%	-0.7%	-0.9%	1.5%	-0.4%	1.5%
	2 hours	-0.5%	-1.2%	-1.8%	2.1%	-1.0%	0.5%
	3 hours	-1.0%	-1.5%	-2.1%	2.2%	-1.3%	0.3%
	4 hours	-1.3%	-1.5%	-2.2%	2.0%	-1.8%	0.1%
	5 hours	-1.4%	-1.6%	-2.1%	1.8%	-1.9%	-0.1%
	6 hours	-1.4%	-1.6%	-1.9%	2.0%	-1.9%	-0.4%
RMSE	1 hour	13.6%	21.0%	24.9%	23.1%	20.4%	29.3%
	2 hours	17.3%	27.2%	30.6%	27.2%	26.3%	36.4%
	3 hours	18.8%	30.2%	32.9%	29.2%	28.4%	40.6%
	4 hours	19.8%	32.2%	34.1%	30.1%	30.1%	41.7%
	5 hours	20.2%	33.5%	34.8%	30.4%	30.9%	42.0%
	6 hours	20.4%	34.2%	35.1%	30.6%	31.1%	42.0%
MAE	1 hour	7.7%	14.2%	19.1%	17.0%	14.5%	22.7%
	2 hours	10.5%	19.4%	24.1%	20.7%	19.5%	28.4%
	3 hours	11.9%	22.1%	26.0%	22.3%	20.9%	31.8%

4 hours	12.7%	23.8%	27.0%	23.1%	22.3%	32.7%
5 hours	13.1%	25.1%	27.4%	23.4%	22.8%	33.1%
6 hours	13.4%	25.7%	27.6%	23.5%	23.0%	33.2%

Table A.2: Relative error metrics of the point forecasts done for each site by the recursive ARMA models using a granularity of 1 hour

	Horizon	Smart persistence	Climatological mean	Recursive ARMA	SVR	GP	NN	AR
MBE	10 min.	0.1%	-0.3%	0.0%	0.1%	-	0.2%	-1.3%
	20 min.	0.3%	-0.3%	0.0%	0.0%	-	0.3%	-2.3%
	30 min.	0.4%	-0.3%	0.0%	-0.1%	-	0.3%	-3.1%
	40 min.	0.5%	-0.3%	-0.1%	-0.1%	-	0.4%	-4.0%
	50 min.	0.5%	-0.3%	-0.1%	-0.1%	-	0.5%	-5.0%
	60 min.	0.6%	-0.3%	-0.1%	0.1%	-	0.5%	-5.9%
RMSE	10 min.	22.3%	38.3%	20.8%	20.8%	-	20.6%	21.3%
	20 min.	27.2%	38.3%	25.1%	25.1%	-	25.0%	26.3%
	30 min.	29.1%	38.3%	27.1%	27.1%	-	27.0%	28.8%
	40 min.	30.4%	38.3%	28.4%	28.5%	-	28.4%	30.5%
	50 min.	31.5%	38.3%	29.4%	29.6%	-	29.4%	31.7%
	60 min.	32.5%	38.3%	30.3%	30.5%	-	30.2%	32.8%
MAE	10 min.	12.0%	29.5%	12.5%	12.0%	-	12.3%	12.1%
	20 min.	15.7%	29.5%	16.2%	15.8%	-	16.1%	15.8%
	30 min.	17.5%	29.5%	18.1%	17.9%	-	18.0%	17.9%
	40 min.	18.8%	29.5%	19.5%	19.4%	-	19.3%	19.4%
	50 min.	19.9%	29.5%	20.5%	20.5%	-	20.4%	20.8%
	60 min.	20.9%	29.5%	21.4%	21.4%	-	21.2%	21.9%
FS	10 min.	0.0%	-71.7%	6.7%	6.8%	-	7.8%	4.3%
	20 min.	3.0%	-36.7%	10.4%	10.2%	-	10.8%	6.2%
	30 min.	5.0%	-24.9%	11.7%	11.5%	-	12.0%	6.0%
	40 min.	6.5%	-17.8%	12.5%	12.2%	-	12.6%	6.1%
	50 min.	6.8%	-13.4%	12.9%	12.4%	-	12.9%	6.0%
	60 min.	6.7%	-10.0%	13.0%	12.8%	-	13.2%	5.7%

Table A.3: Relative error metrics of the point forecasts done for all the sites (average GHI = 493W/m^2) by the reference models, the recursive ARMA models and the models proposed in (Lauret et al., 2015) using a granularity of 10 minutes

	Horizon	Smart persistence	Climatological mean	Recursive ARMA	SVR	GP	NN	AR
MBE	1 hour	0.5%	-0.4%	0.1%	0.2%	0.4%	0.3%	-1.4%
	2 hours	0.5%	-0.4%	-0.4%	0.7%	0.2%	0.1%	-2.7%

	3 hours	-0.3%	-0.4%	-0.6%	1.2%	-0.2%	-0.1%	-4.2%
	4 hours	-1.2%	-0.4%	-0.8%	1.1%	-0.4%	-0.4%	-5.6%
	5 hours	-1.6%	-0.4%	-1.0%	1.2%	-0.5%	-0.5%	-6.9%
	6 hours	-1.4%	-0.4%	-0.9%	1.2%	-0.5%	-0.5%	-8.1%
RMSE	1 hour	23.2%	33.8%	22.1%	22.1%	22.6%	22.1%	22.6%
	2 hours	31.4%	33.8%	27.5%	27.8%	27.9%	27.8%	29.4%
	3 hours	36.0%	33.8%	30.0%	30.6%	30.7%	30.4%	33.8%
	4 hours	38.1%	33.8%	31.3%	31.7%	31.9%	31.8%	36.9%
	5 hours	38.2%	33.8%	31.9%	32.1%	32.3%	32.3%	38.9%
	6 hours	37.4%	33.8%	32.1%	32.3%	32.4%	32.4%	40.0%
MAE	1 hour	14.7%	25.4%	15.5%	15.3%	15.8%	15.4%	15.2%
	2 hours	21.0%	25.4%	20.0%	20.0%	20.2%	20.1%	20.5%
	3 hours	24.6%	25.4%	22.0%	22.1%	22.5%	22.3%	23.7%
	4 hours	26.2%	25.4%	23.1%	23.0%	23.5%	23.3%	25.8%
	5 hours	26.2%	25.4%	23.6%	23.3%	23.9%	23.9%	27.1%
	6 hours	25.5%	25.4%	23.9%	23.4%	23.9%	23.9%	27.9%
FS	1 hour	0.0%	-45.8%	4.7%	4.8%	2.4%	4.8%	2.6%
	2 hours	-0.9%	-8.6%	11.5%	10.5%	10.2%	10.8%	5.4%
	3 hours	1.4%	7.4%	17.7%	16.2%	16.0%	16.7%	7.5%
	4 hours	6.1%	16.7%	22.9%	21.8%	21.3%	21.7%	9.1%
	5 hours	11.6%	21.9%	26.2%	25.7%	25.3%	25.3%	10.1%
	6 hours	16.4%	24.4%	28.2%	27.9%	27.7%	27.7%	10.6%

Table A.4: Relative error metrics of the point forecasts done for all the sites (average GHI =491W/m^2) by the reference models, the recursive ARMA models and the models proposed in (Lauret et al., 2015) using a granularity of 1 hour

Appendix B: CRPS and CRPSS of the probabilistic forecasts

	Horizon	Desert Rock (mean = 549W/m^2)	Fort Peck (mean = 395W/m^2)	Fouillole (mean = 515W/m^2)	Ohau (mean = 492W/m^2)	Saint-Pierre (mean = 538W/m^2)	Tampon (mean = 458W/$^{-2}$)	Average (mean = 493W/$^{-2}$)
CRPS	10 min.	4.7%	8.0%	12.9%	11.8%	8.0%	13.0%	9.5%
	20 min.	6.2%	10.6%	15.8%	14.3%	10.3%	16.3%	12.0%
	30 min.	6.9%	11.8%	17.1%	15.4%	11.6%	18.0%	13.2%
	40 min.	7.4%	12.7%	17.9%	16.1%	12.5%	19.3%	14.0%
	50 min.	7.9%	13.5%	18.5%	16.7%	13.1%	20.2%	14.7%
	60 min.	8.2%	14.1%	19.1%	17.0%	13.7%	21.0%	15.2%
CRPSS	10 min.	15.7%	25.1%	19.4%	19.1%	23.0%	22.2%	20.7%
	20 min.	0.6%	12.3%	10.0%	10.3%	10.9%	12.6%	10.2%
	30 min.	-2.4%	9.5%	8.3%	9.0%	7.9%	10.9%	8.1%
	40 min.	-2.9%	8.8%	8.2%	8.7%	7.3%	11.2%	7.8%
	50 min.	-2.6%	8.7%	8.8%	9.2%	7.9%	12.3%	8.4%
	60 min.	-1.9%	8.9%	9.5%	10.0%	8.6%	13.5%	9.2%

Table B.1: CRPS and CRPSS of the probabilistic forecasts done for each site by the recursive ARMA models using a granularity of 10 minutes

	Horizon	Desert Rock (mean = 548W/m²)	Fort Peck (mean = 391W/m²)	Fouillole (mean = 510W/m²)	Ohau (mean = 499W/m²)	Saint-Pierre (mean = 538W/m²)	Tampon (mean = 456W/m²)	Average (mean = 491W/m²)
CRPS	1 hour	6.5%	10.7%	13.5%	12.2%	10.5%	15.8%	11.3%
	2 hours	8.2%	14.0%	16.7%	14.6%	13.6%	19.5%	14.1%
	3 hours	9.0%	15.6%	17.9%	15.6%	14.7%	21.6%	14.4%
	4 hours	9.6%	16.6%	18.5%	16.0%	15.5%	22.1%	16.1%
	5 hours	9.8%	17.3%	18.8%	16.2%	15.8%	22.3%	16.4%
	6 hours	10.0%	17.6%	20.0%	16.3%	16.0%	22.4%	16.5%
CRPSS	1 hour	15.3%	30.6%	22.6%	22.6%	25.5%	29.9%	25.1%
	2 hours	2.6%	18.1%	12.3%	14.1%	11.0%	19.5%	13.8%
	3 hours	-0.6%	14.4%	10.2%	12.1%	8.4%	13.8%	10.4%
	4 hours	-2.0%	13.5%	10.2%	11.9%	5.9%	12.6%	9.4%
	5 hours	-1.8%	13.2%	10.8%	12.6%	5.1%	11.6%	9.2%
	6 hours	-0.7%	14.0%	11.3%	13.2%	5.3%	10.7%	9.5%

Table B.2: CRPS and CRPSS of the probabilistic forecasts done for each site by the recursive ARMA models using a granularity of 1 hour

Appendix C: Error metrics formula

(C.1) $\mathrm{MBE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{GHI}_{\mathrm{forecast.i}} - \mathrm{GHI}_{\mathrm{measured.i}})$

(C.2) $\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{GHI}_{\mathrm{forecast.i}} - \mathrm{GHI}_{\mathrm{measured.i}})^2}$

(C.3) $\mathrm{MAE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}|\mathrm{GHI}_{\mathrm{forecast.i}} - \mathrm{GHI}_{\mathrm{measured.i}}|$

(C.4) $FS = \left[1 - \frac{RMSE(forecasting\ model)}{RMSE(Persitance\ with\ kt^*)}\right] * 100$

(C.5) $CRPSS = \left[1 - \frac{CRPS(forecasting\ model)}{CRPS(Persitance\ ensemble\ with\ kt^*)}\right] * 100$

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. doi:10.1016/j.apenergy.2015.08.011

Almeida, M.P., Zilles, R., Lorenzo, E., 2014. Extreme overirradiance events in São Paulo, Brazil. Sol. Energy 110, 168–173. doi:10.1016/j.solener.2014.09.012

Andersen, T.G., 2000. Some Reflections on Analysis of High-Frequency Data. J. Bus. Econ. Stat. 18, 146. doi:10.2307/1392552

Bacher, P., Madsen, H., Nielsen, H.A., 2009. Online short-term solar power forecasting. Sol. Energy 83, 1772–1783. doi:10.1016/j.solener.2009.05.016

Bessa, R.J., Trindade, A., Silva, C.S.P., Miranda, V., 2015. Probabilistic solar power forecasting in smart grids using distributed information. Int. J. Electr. Power Energy Syst. 72, 16–23. doi:10.1016/j.ijepes.2015.02.006

Boland, J., 2015. Spatial-temporal forecasting of solar radiation. Renew. Energy 75, 607–616. doi:10.1016/j.renene.2014.10.035

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31, 307–327. doi:10.1016/0304-4076(86)90063-1

Bracale, A., Caramia, P., Carpinelli, G., Di Fazio, A., Ferruzzi, G., 2013. A Bayesian Method for Short-Term Probabilistic Forecasting of Photovoltaic Generation in Smart Grid Operation and Control. Energies 6, 733–747. doi:10.3390/en6020733

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.

Cortes, C., Vapnik, V., 1995. Support-Vector Networks. Mach. Learn. 20, 273–297. doi:10.1023/A:1022627411411

Dambreville, R., Blanc, P., Chanussot, J., Boldo, D., 2014. Very short term forecasting of the Global Horizontal Irradiance using a spatio-temporal autoregressive model. Renew. Energy 72, 291–300. doi:10.1016/j.renene.2014.07.012

David, M., Diagne, M., Lauret, P., 2012. Output and errors indicators for solar forecasting models, in: Proceedings of World Renewable Energy Forum (WREF) 2012. Presented at the World Renewable Energy Forum (WREF) 2012, American Solar Energy Society, Denver, Colorado, USA, pp. 831–836.

Diagne, M., David, M., Lauret, P., Boland, J., Schmutz, N., 2013. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renew. Sustain. Energy Rev. 27, 65–76. doi:10.1016/j.rser.2013.06.042

Engle, R.F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50, 987–1007.

González-Rivera, G., 1998. Smooth-Transition GARCH Models. Stud. Nonlinear Dyn. Econom. 3. doi:10.2202/1558-3708.1041

Haessig, P., Multon, B., Ben Ahmed, H., Lascaud, S., Bondon, P., 2015. Energy storage sizing for wind power: impact of the autocorrelation of day-ahead forecast errors: Energy storage sizing for wind power. Wind Energy 18, 43–57. doi:10.1002/we.1680

Hamill, T.M., 2001. Interpretation of Rank Histograms for Verifying Ensemble Forecasts. Mon.

Weather Rev. 129, 550–560. doi:10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2

Hamill, T.M., 1997. Reliability Diagrams for Multicategory Probabilistic Forecasts. Weather Forecast. 12, 736–741. doi:10.1175/1520-0434(1997)012<0736:RDFMPF>2.0.CO;2

Hanna, R., Kleissl, J., Nottrott, A., Ferry, M., 2014. Energy dispatch schedule optimization for demand charge reduction using a photovoltaic-battery storage system with solar forecasting. Sol. Energy 103, 269–287. doi:10.1016/j.solener.2014.02.020

Hernández-Torres, D., Bridier, L., David, M., Lauret, P., Ardiale, T., 2015. Technico-economical analysis of a hybrid wave power-air compression storage system. Renew. Energy 74, 708–717. doi:10.1016/j.renene.2014.08.070

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

Hoff, T.E., Perez, R., 2012. Modeling PV fleet output variability. Sol. Energy 86, 2177–2189. doi:10.1016/j.solener.2011.11.005

Iqbal, M., 2012. An Introduction To Solar Radiation. Elsevier.

Iversen, E.B., Morales, J.M., Møller, J.K., Madsen, H., 2014. Probabilistic forecasts of solar irradiance using stochastic differential equations: PROBABILISTIC SOLAR FORECASTS BY SDES. Environmetrics 25, 152–164. doi:10.1002/env.2267

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

Kierkegaard, J.L., Nielsen, J.N., Jensen, L., Madsen, H., 2000. Estimating GARCH models using Recursive Methods [WWW Document]. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.40.703&rep=rep1&type=pdf

Kühnert, J., Lorenz, E., Heinemann, D., 2013. Satellite-Based Irradiance and Power Forecasting for the German Energy Market, in: Solar Energy Forecasting and Resource Assessment. Elsevier, pp. 267–297.

Lauret, P., Voyant, C., Soubdhan, T., David, M., Poggi, P., 2015. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol. Energy 112, 446–457. doi:10.1016/j.solener.2014.12.014

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter-Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J.W., Morcrette, J.-J., 2013. McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. Atmospheric Meas. Tech. 6, 2403–2418. doi:10.5194/amt-6-2403-2013

Ljung, L., Söderström, T., 1983. Theory and Practice of Recursive Estimation. MIT Press, Cambridge, Massachusetts.

Lorenz, E., Hurka, J., Heinemann, D., Beyer, H.G., 2009. Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems. IEEE J. Sel. Top. Appl. Earth Obs. Remote

Sens. 2, 2–10. doi:10.1109/JSTARS.2009.2020300

Madsen, H., 2007. Time Series Analysis. Chapman & Hall, London.

Mathiesen, P., Brown, J.M., Kleissl, J., 2013. Geostrophic Wind Dependent Probabilistic Irradiance Forecasts for Coastal California. IEEE Trans. Sustain. Energy 4, 510–518. doi:10.1109/TSTE.2012.2200704

McAleer, M., Medeiros, M.C., 2008. Realized Volatility: A Review. Econom. Rev. 27, 10–45. doi:10.1080/07474930701853509

McClear service for estimating irradiation under clear sky [WWW Document], 2015. . Sol. Irradiat. Data. URL http://www.soda-pro.com/web-services/radiation/mcclear

Muller, K.-R., Smola, A.J., Ratsch, G., Scholkopf, B., Kohlmorgen, J., Vapnik, V., 1997. Predicting Time Series with Support Vector Machines.

Nelson, D.B., Cao, C.Q., 1992. Inequality Constraints in the Univariate GARCH Model. J. Bus. Econ. Stat. 10, 229. doi:10.2307/1391681

Perez, R., Kankiewicz, A., Schlemmer, J., Hemker, K., Kivalov, S., 2014. A new operational solar resource forecast model service for PV fleet simulation, in: Photovoltaic Specialist Conference (PVSC), 2014 IEEE 40th. Presented at the Photovoltaic Specialist Conference (PVSC), 2014 IEEE 40th, pp. 0069–0074. doi:10.1109/PVSC.2014.6925204

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., Hoff, T.E., 2010. Validation of short and medium term operational solar radiation forecasts in the US. Sol. Energy 84, 2161–2172. doi:10.1016/j.solener.2010.08.014

Perez, R., Lorenz, E., Pelland, S., Beauharnois, M., Van Knowe, G., Hemker, K., Heinemann, D., Remund, J., Müller, S.C., Traunmüller, W., Steinmauer, G., Pozo, D., Ruiz-Arias, J.A., Lara-Fanego, V., Ramirez-Santigosa, L., Gaston-Romero, M., Pomares, L.M., 2013. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. Sol. Energy 94, 305–326. doi:10.1016/j.solener.2013.05.005

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. doi:10.1002/qj.559

Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian processes for machine learning, Adaptive computation and machine learning. MIT Press, Cambridge, Mass.

Reikard, G., 2009. Predicting solar radiation at high resolutions: A comparison of time series forecasts. Sol. Energy 83, 342–349. doi:10.1016/j.solener.2008.08.007

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222. doi:10.1023/B:STCO.0000035301.49549.88

Sperati, S., Alessandrini, S., Pinson, P., Kariniotakis, G., 2015. The “Weather Intelligence for Renewable Energies” Benchmarking Exercise on Short-Term Forecasting of Wind and Solar Power Generation. Energies 8, 9594–9619. doi:10.3390/en8099594

Taylor, J., 2004. Volatility forecasting with smooth transition exponential smoothing. Int. J. Forecast. doi:10.1016/j.ijforecast.2004.09.010

Tsay, R.S., 2005. Analysis of financial time series. Wiley, Hoboken, N.J.

Voyant, C., Soubdhan, T., Lauret, P., David, M., Muselli, M., 2015. Mean absolute log-return and solar radiation forecastability. Presented at the ISES Solar World Congress, Daegu, Korea.

Zhang, Y., Wang, J., Wang, X., 2014. Review on probabilistic forecasting of wind power generation. Renew. Sustain. Energy Rev. 32, 255–270. doi:10.1016/j.rser.2014.01.033