Calculation of solar irradiation prediction intervals combining volatility and kernel density estimates.

Juan R. Trapero[a,*]

[a] *University of Castilla-La Mancha*
Department of Business Administration, Ciudad Real 13071, Spain

Abstract

In order to integrate solar energy into the grid it is important to predict the solar radiation accurately, where forecast errors can lead to significant costs. Recently, the increasing statistical approaches that cope with this problem is yielding a prolific literature. In general terms, the main research discussion is centered on selecting the "best" forecasting technique in accuracy terms. However, the need of the users of such forecasts require, apart from point forecasts, information about the variability of such forecast to compute prediction intervals. In this work, we will analyze kernel density estimation approaches, volatility forecasting models and combination of both of them in order to improve the prediction intervals performance. The results show that an optimal combination in terms of prediction interval statistical tests can achieve the desired confidence level with a lower average interval width. Data from a facility located in Spain are used to illustrate our methodology.

Keywords: forecasting, solar irradiation, prediction intervals, volatility, kernel density estimation.

1. Introduction

Solar power generation has been steadily increasing worldwide as a response to environmental concerns. Unfortunately, the integration of solar energy into the energy mix of a country brings new challenges. The main problem is due to the variability of the solar energy, which is not available

*Corresponding author.
Email address: juanramon.trapero@uclm.es (Juan R. Trapero)

"on demand". Therefore, reasonably accurate forecasts are required to make this kind of energy economically viable. Depending on the objective, different forecasting horizons are needed, for instance, long-term forecasts are useful for locating potential solar power plants, energy resource planning and scheduling employs mid-term (up to 48 hours) solar forecasts, whereas intra-day forecasts are required for load following and predispatch (Kleissl, 2013). Kraas et al. (2013) economically quantified the impact of forecasting errors in the Spanish electricity system for a concentrating solar power plant. In Spain, forecasts of production have to be provided to the transmission system operator (TSO). In case of deviations from the scheduled production, the TSO applies a cost penalty. If the forecast is higher than the real production, the TSO charges falling penalties, and conversely, charges rising penalties for a production dispatch above the forecasted value. In that reference, forecasting improvement may significantly reduce penalty charges by 47.6% compared to the simple persistence forecasts.

In general terms, most of the published literature on solar energy forecasting is based on the application of different techniques in order to provide point forecasts. Among the diverse techniques provided, the selection of the best technique is usually based on choosing the one with the lowest forecast error. Nevertheless, the need of the users of such forecast (or stakeholders) require more information apart from the point forecast. In particular, they require both uncertainty and variability forecasts (chapter 14, Kleissl (2013)) that they can be even more useful than typical point forecasts (Chu et al., 2015). In this context, point forecasts and its associated uncertainty are usually given as hourly-average time series, whereas irradiance variability or ramp events are measured in an intra-hour time scale, for example, minutes. Note that this difference between variability and uncertainty might not be unanimous in other research disciplines out of the solar energy literature. For instance, in supply chain applications, variability is commonly employed as a measure of uncertainty. Nonetheless, in this work we follow the distinction made in (chapter 14, Kleissl (2013)).

This work focuses on the uncertainty and their use to compute prediction intervals. Essentially, prediction intervals require a forecast of the uncertainty, usually measured as the standard deviation of the forecast error or volatility in finance terms. Traditionally, forecast error is assumed to be constant, independent and normally distributed. However, since the solar irradiance time series involves very complex processes, it is expected that those assumptions may be violated and thus, further improvements on the

standard deviation forecasts should be proposed. Depending on which assumptions we want to be relaxed, different approaches can be used. Although a detailed literature review will be shown in the next section we can summarize the following ideas. On the one hand, if we assume normality but independence and constant standard deviation assumptions are relaxed, then volatility models as generalised autoregressive conditional heteroskedastic (GARCH) models coming from finance literature can be employed. On the other hand, if we relax the normality hypothesis but we assume that error variability is constant and independent, then non-parametric approaches as either histograms or kernel density estimation can be utilized. Finally, if any of the aforementioned assumptions are expected to be fulfilled, then a combination of volatility models and non-parametric approaches should be employed.

The results show that such a combination provides a robust performance by achieving a compromise between prediction interval coverage and average interval width with respect to other techniques. Additionally, the combination can be optimized by maximizing the conditional coverage test developed by Christoffersen (1998).

In order to illustrate the methodology employed, we are going to focus on one-step-ahead uncertainty forecasts obtained from Global Horizontal Irradiation (GHI) data that is crucial for photovoltaic generators. Note that this study can be extended in a straightforward manner to analyse the Direct Normal Irradiation (DNI) that is more relevant for Concentrated Solar Power applications.

The rest of the paper is organised as follows: Section 2 reviews the literature on prediction intervals and it describes the models employed in this article. Section 3 describes the case study dataset. Section 4 lays out the experimental setup and the discussion of initial results. Section 5 defines a new approach to compute prediction intervals based on combining previous models, and finally, section 6 presents the concluding remarks.

2. Literature review

2.1. Estimation of prediction intervals

There are two principal streams to compute prediction intervals. In the first place, a theoretical prediction interval can be calculated based on a forecasting model that assumes to reproduce the data correctly and that the forecast error follows a determined distribution (Chatfield, 2000). Since the

model is assumed to be specified correctly, the forecasts are unbiased and the forecast errors have zero mean and constant variance, which is function of a certain forecasting horizon and the model parameters (see for example, ARMA models (Box et al., 1994), State Space models (Harvey, 1989), Exponential Smoothing models (Hyndman et al., 2008)). In this sense, Gaussian distribution is usually employed to facilitate the derivation of the theoretical formulae. Therefore, a prediction interval can be defined by the lower ($\hat{L}_{t+h}$) and upper limit ($\hat{U}_{t+h}$) for a certain forecasting horizon h and a confidence level $100\alpha\%$, such as:

$$\left[\hat{L}_{t+h}, \hat{U}_{t+h}\right] = \left[\hat{F}_{t+h} \pm z_{(1-\alpha)/2} \cdot \hat{\sigma}_h\right] \tag{1}$$

where $\hat{F}_{t+h}$ is the point forecast at $t+h$. Note that we have information available up to time t. $z_{(1-\alpha)/2}$ can be calculated from a standard normal distribution table for a certain confidence level α and $\hat{\sigma}_h$ is the theoretical standard deviation determined by the forecasting horizon and the model forecasting parameters as previously discussed.

Nonetheless, if either the model is misspecified or aforementioned assumptions about forecast errors are not valid, the theoretical prediction intervals derived tend to be too narrow reducing its credibility. In other words, if the confidence level is 90%, there are more than 10% of forecasts located out of the prediction interval. In the particular case of solar irradiation forecasting, it is reasonable to doubt about obtaining a forecasting model capable of capturing the complex solar irradiation dynamics.

If theoretical formulae are not available, or there are doubts about the validity of the "true" model, empirical approaches to compute prediction intervals can be used as an useful alternative (Chatfield, 2000). Note that, depending on the forecasting horizon deterministic approaches denoted by Numerical Weather Prediction models are widely used to provide solar irradiation forecasts. As they do not rely on a statistical model, the only way to compute prediction intervals for them is to use empirical approaches.

Empirical approaches can also be divided in two parts (Lee and Scholtes, 2014). The first option is to use empirical residual errors, also known as in-sample errors. In this sense, the forecasting model is applied to the same data employed for estimating its parameters to provide a sample of forecast errors at different lead times. Then, on this in-sample errors data, various non-parametric methods such as Chebyshev's inequality (Gardner, 1988) or kernel density estimators (Wu, 2012), and semi-parametric techniques such

as quantile regression (Taylor and Bunn, 1999) can be employed to build the prediction intervals. However, prediction intervals based on residuals errors may also be narrow given that residuals errors usually are lower than out-of-sample errors (Tashman, 2000). In order to overcome such a limitation, empirical approaches are also developed on the basis of out-of-sample errors, i.e., forecast errors obtained from the forecasting model applied on data that have not been used in the fitting process. This kind of empirical approach was initially suggested by Williams and Goodman (1971) and it has gained popularity in diverse applications, particularly in those areas where data requirements are not a limitation (Isengildina-Massa et al., 2011). In addition, Lee and Scholtes (2014) have explored the robustness of the empirical approach from an asymptotic perspective.

Focusing on constructing empirical prediction intervals based on out-of-sample forecast errors, we can employ parametric and non-parametric approaches. For instance, the $p\%$ forecast error quantile denoted by $\hat{Q}_h(p)$ can be estimated non-parametrically by means of the empirical distribution of the generated forecast errors. Alternatively, $\hat{Q}_h(p)$ can be parametrically determined by assuming a forecast error distribution. For the sake of simplicity if the distribution is assumed to be normal with finite mean μ_h and variance σ_h, then $\hat{Q}_h(p) = \hat{\mu}_{t+h} + z_p\hat{\sigma}_{t+h}$. Note that, unlike to the theoretical way to compute prediction intervals where the forecast errors had zero mean and the variance depended on the forecasting model parameters, in the empirical approach the sample mean and the sample variance are calculated as:

$$\hat{\mu}_{t+h} = \frac{\sum_{i=t-n}^{t-1} \hat{\epsilon}_{i+h}}{n} \tag{2}$$

$$\hat{\sigma}^2_{t+h} = \frac{\sum_{i=t-n}^{t-1} (\hat{\epsilon}_{i+h} - \hat{\mu}_{i+h})^2}{n} \tag{3}$$

Then, the parametric empirical prediction interval with the nominal $100\alpha\%$ coverage is

$$\left[\hat{L}_{t+h}, \hat{U}_{t+h}\right] = \hat{F}_{t+h} + \hat{\mu}_{t+h} \pm z_{(1-\alpha)/2} \cdot \hat{\sigma}_{t+h} \tag{4}$$

It should be noted that equation (4) assumes that the forecast errors are normally distributed but it also assumes that forecast errors mean and variance are time constants. In this way, equation (2) and (3) can be seen as forecasts of the mean and variance by means of moving averages of past

forecast errors for a data window of size n. Although moving averages can be used to forecast variables that fluctuates around a level, this option has two important disadvantages, firstly, we have to determine the data window size and, secondly, this technique gives the same weights to both the old and recent past data. In order to overcome such limitations, this work suggests to use volatility forecasts that, although they have been widely used in finance applications, to the best of our knowledge, this is the first time they are used to compute prediction intervals of solar irradiation forecasts. Note that, forecast errors mean is expected to fluctuate around zero in a slow manner, nonetheless, the forecast errors variance is expected to change abruptly as a consequence of clouds passing and thus, volatility forecasts capable of handling time-varying variance may be a well-suited option.

In the remainder of the section, we will review the main terms in (4) required to provide a prediction interval. In the first place, we will investigate the point forecasting models found in the solar irradiation forecasting literature to estimate $\hat{F}_{t+h}$. In the second place, we will explore the different volatility forecasting models that can be employed to forecast $\hat{\sigma}_{t+h}$. Since point forecasts are nearly unbiased, for the sake of simplicity $\hat{\mu}_{t+h}$ will be assumed to be zero.

2.2. Point forecast

A literature review about forecasting models can be found in (Diagne et al., 2013). If irradiance measurements are available and the forecasting horizon is defined at short/mid-term, statistical forecasting models are typically applied (Reikard, 2009).

Given that the solar irradiation is non-stationary, exhibiting a seasonal pattern due to the planetary motions. The different statistical approaches may be categorized depending on how the seasonality is dealt with (Diagne et al., 2013). In this sense, three main approaches can be found in the literature: i) remove the seasonality using a deterministic approach that makes use of our understanding of solar geometry and modelling the remainder; ii) remove the seasonality using statistical means and then modelling the resulting stationary series; and iii) model the raw time series data with models capable of dealing with the seasonality. Under the first category, models can make use of the *clearness index* or the *clear-sky* index. The *clearness index* is defined as the ratio of irradiance at ground level with respect to extraterrestial irradiance. On the other hand, the *clear-sky* index is the ratio

of irradiance at ground level to clear sky irradiance, which it is the global irradiance in clear sky conditions computed by means of *clear sky* models that use atmospheric conditions information. However, some authors criticize the use of such indices given that they are mostly random and recommend the use of irradiance data (Sfetsos and Coonick, 2000). Under the second category various techniques such as Fourier series, high order polynomial models, trigonometric function models, Gaussian models and time series decomposition have been employed to make the time series stationary, which is then model as an autoregressive process or by using exponential smoothing models without seasonal component (Boland, 1995; Dong et al., 2013; Al-Sadah et al., 1990; Kaplanis, 2006; Baig et al., 1991; Yang et al., 2015). Under the third category seasonality capable models are used, such as dynamic harmonic regression and exponential smoothing with seasonal terms, and therefore the need for the two-stage approach described before is no longer needed, thus eliminating potential biases introduced by it (Trapero et al., 2015). In this work we adhere to the third category of time series models, attempting to avoid introducing errors coming from either the calculation of the indices or from the decomposition of the time series, so that any observed errors are solely due to the forecasting model fit and simplifying the interpretation of our results.

2.2.1. ARIMA model

Box et al. (1994) propose a general framework based on a Seasonal AutoRegressive Integrated Moving Average (ARIMA) process of order $(p, d, q) \times (P, D, Q)_s$ to model stationary and nonstationary time series. The process can be expressed by:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D y_t = \theta(B)\Theta(B^s)\epsilon_t, \tag{5}$$

where y_t is an observable time series (solar irradiance). Typically, ϵ_t is a white noise process having mean zero and constant variance[1] σ_ϵ^2. The backward shift operator is denoted by $By_t = y_{t-1}$. The non-seasonal Autoregressive and Moving Average operators are defined by $\phi(B)$ and $\theta(B)$ polynomials of order p and q respectively, and d denotes the order of differencing that is required to make the time series stationary, such as:

[1]Note that, in this work we will relax the constant variance assumption allowing a time-varying variance as it will be described later on.

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p, \quad (6)$$
$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q. \quad (7)$$

The seasonal ARIMA part is represented by $\Phi(B^s)$ and $\Theta(B^s)$ polynomials of order P and Q respectively and D denotes the order of seasonal differencing, where s is the number of periods per season.

ARIMA models can be difficult to identify, in particular, when the process is a mixture of AR and MA structures, (Makridakis et al., 1998). In order to provide the same forecast errors for all the volatility forecasting models, the same point forecast or conditional mean should also be provided. In this work we propose an $ARIMA(1,0,0) \times (1,1,0)_{24}$. The reason to use this model is twofold. Firstly, it has been previously analyzed in (Reikard, 2009). Secondly, it yields reasonable results in our dataset.

2.3. Volatility models

2.3.1. Moving average

According to Boudoukh et al. (1997), a widely used approach to forecast volatility over some prior window is to use a moving average, as it is expressed in (3). Moving average has been typically used via the bollinger bands in finance applications, within the technical analysis. In principle, the bollinger bands are defined as:

$$[\hat{L}^B_{t+1}, \hat{U}^B_{t+1}] = [\hat{F}_{t+1} - k \cdot \hat{\sigma}_{t+1}, \hat{F}_{t+1} + k \cdot \hat{\sigma}_{t+1}] \quad (8)$$

where $[\hat{L}^B_{t+1}, \hat{U}^B_{t+1}]$ are the lower and upper limit at time t, $\hat{F}_{t+1}$ is the forecast at time $t+1$ and $\hat{\sigma}_{t+1}$ is the volatility forecast at time $t+1$. It should be noted that the bollinger bands follow an expression very similar to the prediction intervals defined in (1), where the volatility forecast is given by a moving average of the squared forecast error (9), selecting the order n and the value of k in a quite *ad-hoc* manner.

$$\hat{\sigma}^2_{t+1} = \frac{\sum_{i=t-n}^{t-1} \hat{\epsilon}^2_{i+1}}{n} \quad (9)$$

However, this estimator has two disadvantages. On the one hand, if volatility clusters, it makes sense to give more weights to recent data. On the other hand, the choice of the length of the observation period (n) is somewhat arbitrary.

2.3.2. Single exponential smoothing

In order to overcome such limitations many practitioners prefer using exponentially weighted moving averages, where the one-step-ahead variance estimator can be expressed as the simple exponential smoothing recursive form (Taylor, 2004), with smoothing parameter a. Typically, the smoothing parameter is referred to α. Nonetheless, to avoid any confusion with the confidence level (α) defined in (1), we have denoted it as a, whose values vary between 0 and 1 (Ord and Fildes, 2012). Thus, the exponentially weighted moving average expression is:

$$\hat{\sigma}^2_{t+1} = a\hat{\epsilon}^2_t + (1-a)\hat{\sigma}^2_t \tag{10}$$

Equation (10) substantially reduces any storage problem, because it only needs the most recent squared forecast error ($\hat{\epsilon}^2_t$) and the most recent volatility forecast ($\hat{\sigma}^2_t$). Unlike expression (9), the weights of old data decay exponentially over time. In order to observe the exponential decay, it is convenient to develop expression (10). If we substitute $\hat{\sigma}^2_t$ with a comparable recursive equation, we obtain:

$$\begin{aligned} \hat{\sigma}^2_{t+1} &= a\hat{\epsilon}^2_t + (1-a)(a\hat{\epsilon}^2_{t-1} + (1-a)\hat{\sigma}^2_{t-1}) \\ \hat{\sigma}^2_{t+1} &= a\hat{\epsilon}^2_t + a(1-a)\hat{\epsilon}^2_{t-1} + (1-a)^2\hat{\sigma}^2_{t-1} \end{aligned} \tag{11}$$

If this substitution process is repeated by replacing $\hat{\sigma}^2_{t-1}$ as a function of $\hat{\sigma}^2_{t-2}$, and so on, the result is:

$$\begin{aligned} \hat{\sigma}^2_{t+1} = \; & a\hat{\epsilon}^2_t + a(1-a)\hat{\epsilon}^2_{t-1} + a(1-a)^2\hat{\epsilon}^2_{t-2} + \ldots + \\ & a(1-a)^{t-1}\hat{\epsilon}^2_1 + (1-a)^t\hat{\sigma}^2_1 \end{aligned} \tag{12}$$

Given that $0 < a < 1$, the weights for all past squared errors decrease exponentially, hence the name exponential smoothing. Additionally, the weights sum is approximately one.

2.3.3. Generalized AutoRegressive Conditional Heteroskedasticity models

Looking back at expression (5), although the term ϵ_t in most of the solar irradiance forecasting literature is assumed to be homoscedastic, in this work we intend to include a changing variance. In this sense, the error term can

be expressed as $\epsilon_t = \sigma_t \cdot v_t$, where

$$\sigma_{t+1}^2 = \omega + \sum_{i=1}^{p} a_i \epsilon_{t-i+1}^2 \tag{13}$$

is the AutoRegressive Conditional Heteroscedasticity (ARCH) model introduced by (Engle, 1982) for estimating and forecasting volatility and v_t is an independent and identically distributed (iid) process. ARCH models express the conditional variance as a linear function of p lagged squared error terms. Bollerslev (1986) proposed the Generalized AutoRegressive Conditional Heteroskedasticity GARCH models that represent a more parsimonious and less restrictive version of the ARCH(p) models. Generally speaking, GARCH models with respect to ARCH models bring similar advantages as exponential smoothing with regards to moving average (Boudoukh et al., 1997). GARCH models express the conditional variance of the forecast error (or return) (ϵ_t) at time t, as a linear function of q lagged squared error terms and p lagged conditional variance terms. For example, GARCH(1,1) model is given by:

$$\sigma_{t+1}^2 = \omega + a_1 \epsilon_t^2 + \beta_1 \sigma_t^2 \tag{14}$$

It should be noted that exponential smoothing has the same formulation as the integrated GARCH model (IGARCH) (Nelson, 1990), in which $\beta_1 = 1 - a_1$, plus an additional restriction in which $\omega = 0$.

3. Case study data

Solar irradiance data have been collected by the Spanish Institute for Concentration Photovoltaics Systems (ISFOC), located in Ciudad Real in the region of Castilla-La Mancha in Spain (at 38.67 ° N, 4.15 ° W, 687 m). Minute-by-minute solar irradiance measurements have been recorded using pyranometers and pyrheliometers, which comply with the international standards of Baseline Surface Radiation Network (BSRN) (McArthur, 2004).

Global Horizontal Irradiance (GHI) data have been provided. GHI is the total solar radiation on a horizontal surface and it is calculated as the sum of Direct Horizontal Irradiance, the Diffuse Horizontal Irradiance and the Albedo. Diffuse irradiance refers to all the solar radiation scattered from the sky. Albedo is the fraction of the incident solar radiation reflected and scattered by the ground.

From the original minute-by-minute data hourly series of solar irradiation data (Wh/m^2) are constructed for this study, aggregating ground-based solar irradiance data (W/m^2). The data span from January 2011 to December 2011, resulting in 8,760 hourly observations. Fig. 1 provides examples of GHI data in winter and summer. For the provided dataset the mean and standard deviation of GHI is approximately 200 Wh/m^2 and 285 Wh/m^2, respectively.

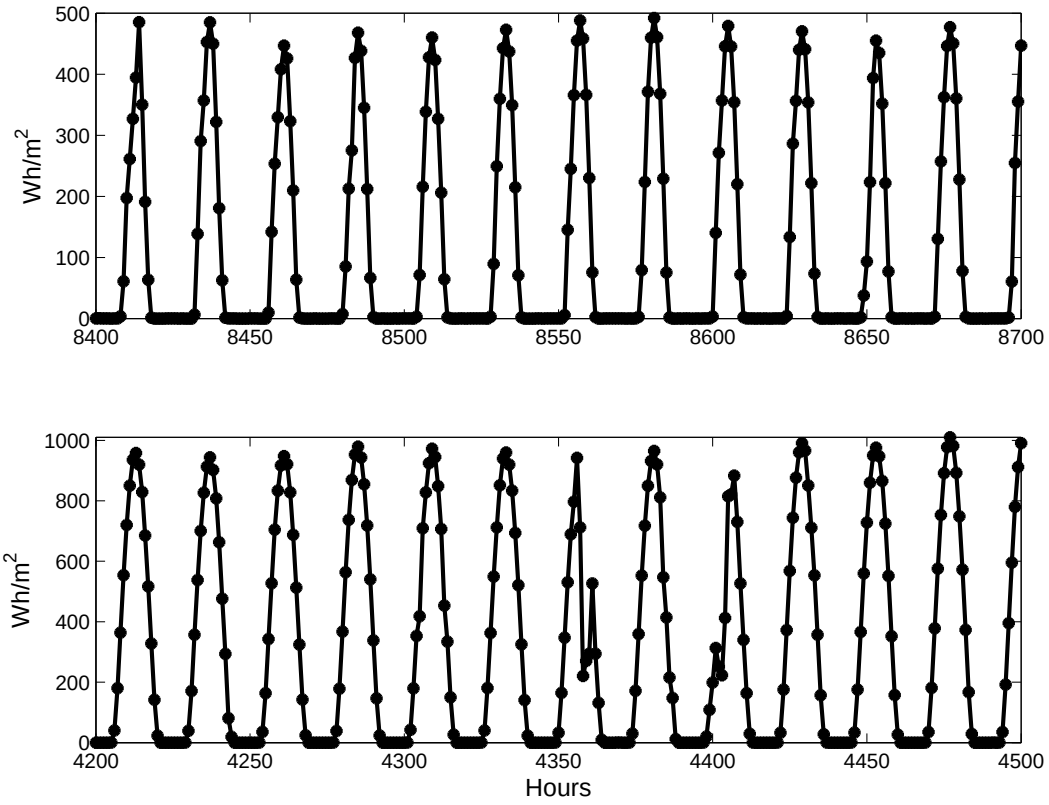

Figure 1: Examples of hourly solar irradiation for GHI in winter (upper plot) and in summer (lower plot).

The time series are characterised by strong seasonality, caused by the day-night cycle. The amplitude of the daily seasonality is changing over the year, following the winter-summer cycle. This results in a complex seasonality that does not fully fall under neither stochastic and deterministic conventional definitions of seasonality. The observed seasonal profile changes in terms

of amplitude and duration (sunrise and sunset) over time, a characteristic that could characterise the seasonality as stochastic. However the way that these changes occur are not random and are controlled by the planetary motions, thus having deterministic elements. On top of this seasonality, atmosphere conditions, such as cloud cover, cause additional variability to the data, increasing the complexity of the data.

4. Experimental setup

The data (8,760 observations) have been split down in two parts of the same size approximately. The first part (4,392 observations) has been used to estimate the parameters of the ARIMA model (in-sample data) and the GARCH parameters. Once the forecasts have been calculated, we have removed the nights by eliminating observations with an elevation angle lower than 15°. Note that we have not removed the nights in the calculation of the ARIMA-GARCH model in order to use the same model than (Reikard, 2009) with a fundamental period of 24 hours. The exponential smoothing parameter has been estimated by considering the ARIMA point forecasts for the in-sample data without nights. An example of the point forecasts vs. the actual values is depicted in the upper plot in Figure 2. The one-step-ahead forecast errors are shown in the lower plot of the same figure. During the first observations, this figure shows that the forecasts are accurate and the out-of-sample forecast error is close to zero, perhaps because the sky is clear. Then, as a result of clouds passing, the forecast error increases. It is interesting to note that the forecast error may also cluster. This fact is illustrated with ellipses in the figure, where cluster 1 indicates that the forecast error is negative for almost the whole day (24 hours), conversely, the cluster 2 provides positive forecast errors. In other words, the presence of clusters implies that the error is not purely random and it may have time correlation that should be explored with volatility models.

The second part of the data (4,368 observations) has been reserved as out-of-sample data. Here, we are only investigating the forecast errors one-step-ahead. In that second part, we have reserved a sample of 100 observations to initialize the kernel density estimate on out-of-sample errors. Then, the rest of the sample is utilized by means of a rolling origin experiment (Tashman, 2000) to test the effectiveness of the different alternatives to compute the prediction intervals.

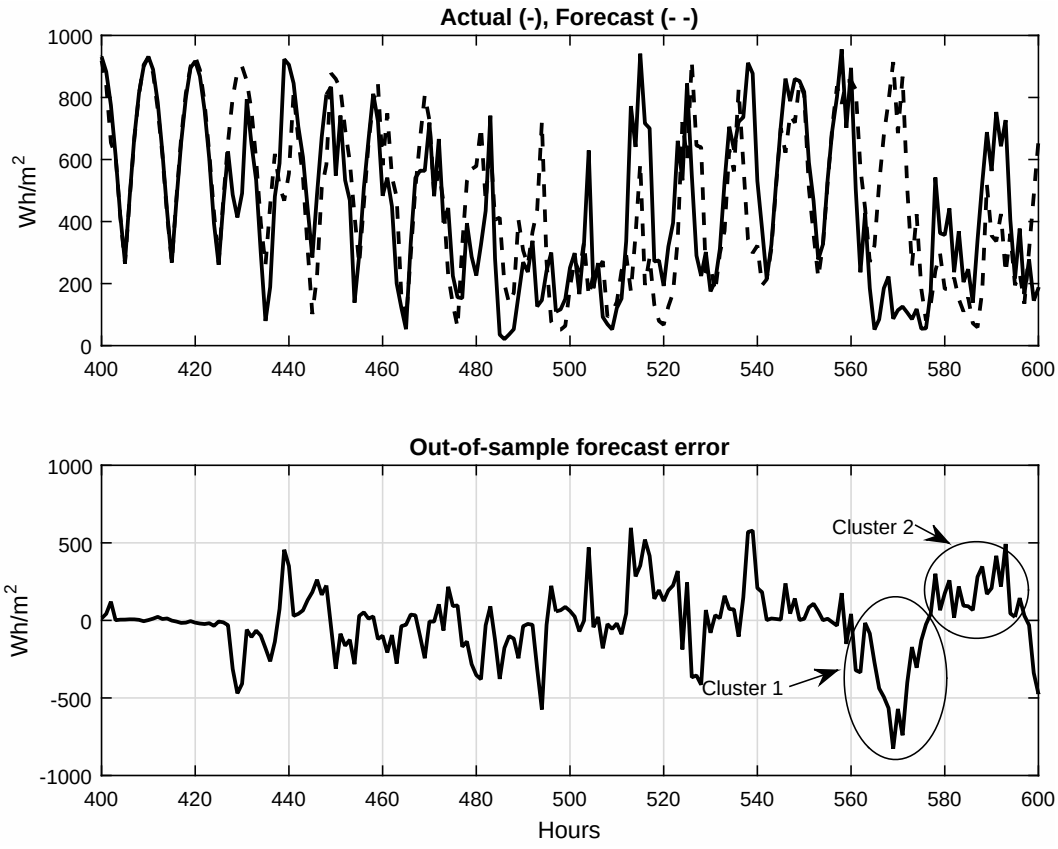

Figure 2: Upper panel, actual values (solid line) vs. forecast values (dashed line). Lower panel, one-step-ahead forecast error

4.1. Prediction interval alternatives

Once we have defined how to forecast the conditional mean and variance of the variable under study, the next step is to use them efficiently to provide a prediction interval. In this sense, we can distinguish two alternatives depending on the error distribution dynamics. Firstly, we assume that the one-step-ahead forecast error distribution properties are approximately constant, although unknown, and are estimated by a non-parametric approach as the kernel density estimation. Secondly, the error distribution is assumed to be normal (parametric hypothesis) and its variance has time-varying properties estimated by GARCH and SES models.

4.1.1. Non-parametric approach. Kernel density estimation

The first approach to compute prediction interval is based on the non-parametric kernel density estimation. This technique represents the probability density function ($f(x)$) of the forecast errors without the need of making assumptions about the distribution of the data. Its formula for a series X at a point x is given by:

$$f(x) = \frac{1}{Nh} \sum_{j=1}^{N} K\left(\frac{x - X_j}{h}\right) \tag{15}$$

where N is the sample size, $K(\cdot)$ is the kernel smoothing function that integrates to one and h is the bandwidth. In the present case study, kernel density estimates were computed using the MATLAB library of statistical functions. We have chosen the MATLAB default values, i.e., a normal kernel smoothing function and an optimal bandwidth for estimating normal densities. More details about density estimation can be found in (Silverman, 1986)

Once the kernel density is estimated the quantiles (Q_k) can be computed as well as the prediction intervals, such as:

$$\left[\hat{L}_1, \hat{U}_1\right] = \left[\hat{F}_{t+1} + Q_k\left(\frac{1-\alpha}{2}\right), \hat{F}_{t+1} + Q_k\left(\frac{1+\alpha}{2}\right)\right] \tag{16}$$

Since we are using a rolling origin experiment with the out-of-sample data to compute the empirical quantiles, the shape of the kernel estimate can be changing as observations are added. Figure 3 shows kernel density estimates and histograms at the beginning of the out-of-sample data with a sample size N=100 (upper panel) and at the end of the out-of-sample with a sample size

N=100 (lower panel), since we keep a constant data window, i.e., each time the newest observation is added, the oldest observation is removed. Note that the kernel density estimates have been scaled to overlay the histograms.

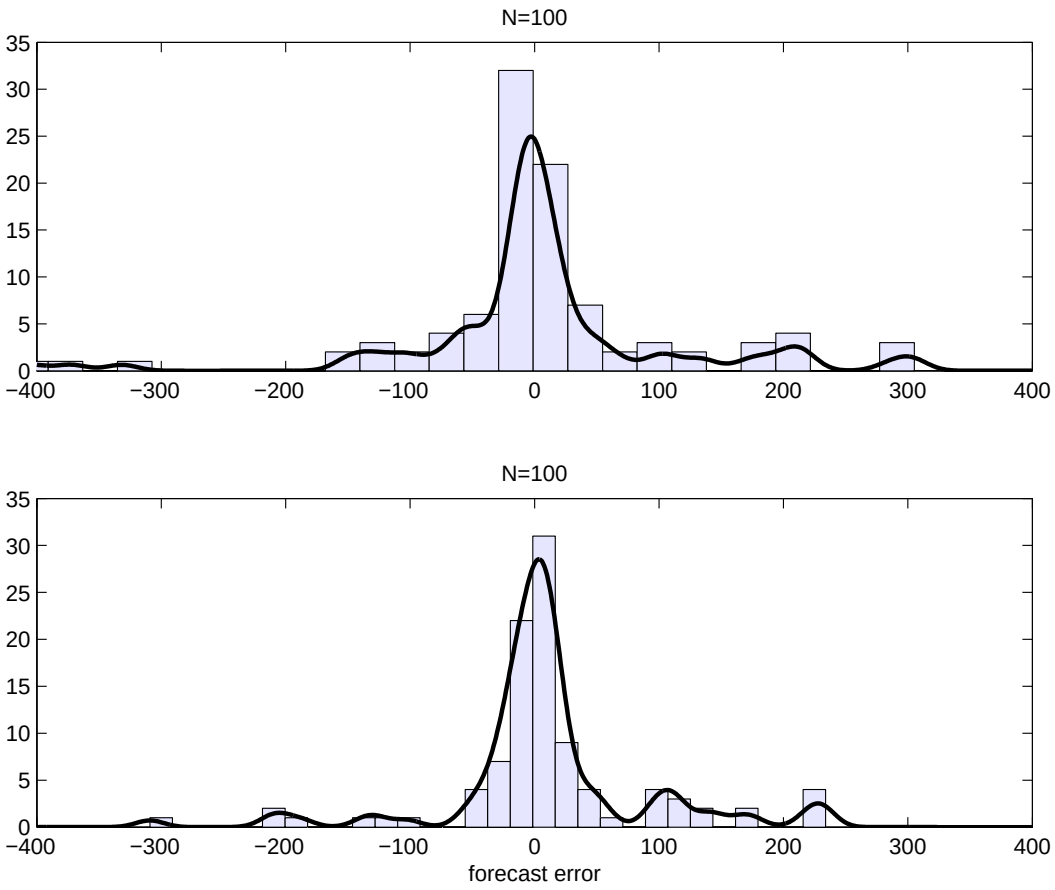

Figure 3: Forecast error kernel density estimates and histograms at the beginning of the out-of-sample (Upper panel) and at the end (Lower panel)

4.1.2. Parametric approaches. GARCH and SES estimation.

In order to incorporate volatility forecasts, we assume that the forecast errors are normally distributed with time-varying variance, such as:

$$\left[\hat{L}_i, \hat{U}_i\right] = \hat{F}_{t+h} + \pm z_{(1-\alpha)/2} \cdot \hat{\sigma}_{i,t+1} \tag{17}$$

where $i = 2$ if the time-varying standard deviation $\hat{\sigma}_{i,t+1}$ is given by the square root of the GARCH(1,1) model in (14) and $i = 3$ if the estimation

	Parameter	Value	Standard Error
ARIMA	c	0.0014	0.003
	ϕ_1	0.0399	0.001
	Φ_{24}	-0.1071	0.0006
GARCH	ω	0.0014	0.003
	a_1	0.99	0.03
	β_1	6.8e-05	2.8e-05
SES	a	0.623	–

Table 1: Parameter estimation ARIMA-GARCH and Single Exponential Smoothing.

of the standard deviation is based on the Single Exponential Smoothing[2] in (10).

4.2. Estimation of parameters

The ARIMA model proposed has been estimated together with the GARCH(1,1) model by using maximum likelihood based on a t-distribution with the Econometrics toolbox from MATLAB. Table 1 shows the estimates of the ARIMA-GARCH model. Since the exponential smoothing does not rely on a statistical model, typical model optimization routines based on maximum-likelihood cannot be used. Instead, literature recommends the minimisation of the sum of in-sample one-step-ahead prediction errors. In fact, RiskMetrics (1996) suggests the following minimisation:

$$min \sum_i (\hat{\epsilon}_i^2 - \hat{\sigma}_i^2)^2 \tag{18}$$

where $\hat{\epsilon}_i^2$ estimates the unobservable actual variance. The last row in Table 1 shows the parameter estimate of the single exponential smoothing model. It is interesting to note that the estimated values of the GARCH(1,1) model suggests the use of a IGARCH model, that is very similar to a single exponential smoothing model with a parameter value close to 1.

4.3. Performance metrics

The error metrics employed to compare the point forecasting models are: i) relative Mean Biased Error (rMBE); and ii) relative Root Mean Squared

[2]The use of Single Exponential Smoothing to volatility forecasting has also been employed by the package RiskMetrics (1996).

	Cond. Mean	Cond. Variance	
	ARIMA	GARCH(1,1)	SES
rMBE(%)	-0.02	2.68	-0.13
rRMSE(%)	31.55	171.0	98.8

Table 2: Conditional mean and variance forecasting accuracy for one-step-ahead horizon

Error (rMSE). The calculations are carried out as follows, let $\hat{\epsilon}_t$ be the forecast error given by:

$$\hat{\epsilon}_t = (y_t - \hat{F}_t), \qquad t = 1, \ldots, T \tag{19}$$

where y_t and $\hat{F}_t$ stand for the actual value and the forecast, respectively, at time t. T is the out-of-sample size. The rMBE and rMSE are defined such as:

$$rMBE = \frac{\sum_{t=1}^{T} \hat{\epsilon}_t / T}{\bar{y}} \cdot 100 \tag{20}$$

$$rRMSE = \frac{\sqrt{\sum_{t=1}^{T} \hat{\epsilon}_t^2 / T}}{\bar{y}} \cdot 100 \tag{21}$$

Here, $\bar{y}$ is the mean of the actual values in the hold-in sample.

Volatility forecasting accuracy can also be measured by utilising equations (19)-(21), where y_t is replaced by the actual variance proxy $\hat{\epsilon}_t^2$ and $\hat{F}_t$ by the squared volatility forecast $\hat{\sigma}_t^2$ provided by either the GARCH(1,1) model or the single exponential smoothing.

4.4. Results

Table 2 shows the conditional mean and variance performance of the models proposed. The first column in the table assesses the forecast error of the point forecasting model (ARIMA). The results show that the ARIMA provides a reasonable rRMSE. The ARIMA model proposed also provides a small bias measured by the rMBE. On the other hand, the conditional variance performance is shown in the second and third columns. Note that the SES provides a more precise volatility forecast according to the rRMSE with a nearly zero bias. The forecasting accuracy improvement of SES over GARCH may be the result of removing the nights in the estimation process.

Intuitively, we assume that a better volatility forecast would provide a better prediction interval. In this part of the text, we are going to validate such an assumption by comparing the prediction intervals performance by means of the prediction interval coverage and the average interval width. Prediction interval coverage is based on the expectation of the hit rate, which indicates the proportion of times that a forecast is included within the prediction interval. Hit rate can be calculated on the basis of the following indicator:

$$I_t = \begin{cases} 1 & \text{if } y_t \in [\hat{L}_t(\alpha), \hat{U}_t(\alpha)] \\ 0 & \text{if } y_t \notin [\hat{L}_t(\alpha), \hat{U}_t(\alpha)] \end{cases} \tag{22}$$

where $[\hat{L}_t(\alpha), \hat{U}_t(\alpha)]$ are the lower and upper limits of the prediction interval for the actual value y_t with a target coverage or confidence level α. Statistical tests to assess the conditional coverage were developed by Christoffersen (1998). In this work we have used the implementation developed by (Hurlin and Perignon, 2013) in MATLAB. The Christoffersen test for correct conditional coverage is the combination of the tests for unconditional coverage and independence. Additionally, it is also important to measure the precision of the prediction interval, i.e., how wide or narrow is the interval. To do this, average interval width is calculated by dividing interval range by its midpoint. In summary, an ideal prediction interval forecasting method should possess a close prediction interval coverage with regards to the desired confidence level and a low average interval width.

Table 3 shows the aforementioned prediction interval performance metrics for the proposed methods. It is interesting to note that the non-parametric kernel density estimation method provides a higher hit rate (third column) with respect to the desired coverage (second column) than the parametric approaches (GARCH and SES). Nonetheless, the non-parametric alternative provides very wide prediction intervals as it can be verified by the average interval width shown in the last column of the table.

It is also interesting to analyze the results regarding the p-values of the Likelihood ratio tests for unconditional coverage (LR_u), independence (LR_i) and conditional coverage (LR_c) that can be found in the fourth, fifth and sixth columns of the table, respectively. P-values lower than 0.05 mean that the null hypothesis of unconditional coverage, independence and conditional coverage, respectively, are rejected at the 5% significance level.

In the same table, we can observe that, although, the kernel density estimation method provides a higher hit rate, it does not pass the unconditional

coverage test. This may happen because it also does not pass the independence test. Consequently, the conditional coverage is not passed neither. However, the parametric GARCH and SES pass the independence test for mostly all the desired coverage α, although given the low hit rate, the unconditional coverage test is not passed, and thus, the conditional coverage is not passed.

5. Combination of prediction intervals

Since the non-parametric approach gives a higher hit rate and the parametric approaches, specially the SES approach, tend to be more independent, another alternative is to combine both methods to compute a prediction interval. Actually, Christoffersen (1998) pointed out that combining non-parametric error distribution with time-varying variance estimators is likely to present a favorable alternative. In this section we will test such an option.

Given that the non-parametric and SES were the methods with better results, we can combine them to obtain the combined prediction interval $[\hat{L}_c, \hat{U}_c]$ in the following form:

$$\begin{aligned} \hat{L}_c &= w \cdot \hat{L}_1 + (1-w) \cdot \hat{L}_3 \\ \hat{U}_c &= w \cdot \hat{U}_1 + (1-w) \cdot \hat{U}_3 \end{aligned} \tag{23}$$

where $0 < w < 1$ is a constant that weights the importance of each technique in the combined prediction interval. In other words, if w is close to 1 the combined prediction interval will be similar to the prediction interval given by the non-parametric kernel density estimate. A possible way to define w is to maximize the conditional coverage Christoffersen test p-value. Following that idea, a rolling origin experiment has been carried out, where the parameter (w) has been recomputed each time a new observation was added. Figure 4 depicts the average optimal weight (w^*) in function of the desired coverage. Interestingly, the typical 50%-50% combination would not be adequate and according to the Christoffersen tests, when the desired coverage is low, we should give more importance to the SES approach, nonetheless, if the desired coverage increases, the weight of the non-parametric approach should be bigger. Actually, such an increase follows an exponential growth.

Last row in Table 3 summarizes the performance of the combined prediction intervals. Essentially, the hit rate is very close to the desired coverage without implying very high interval width. Additionally, most of the

Method	$\alpha(\%)$	Hit Rate (%)	LR_u	LR_i	LR_c	Av. width
KERNEL	**70**	77.14	0.01	0.00	0.00	1.08
	75	80.00	0.04	0.00	0.00	1.20
	80	84.13	0.06	0.00	0.00	1.33
	85	89.21	0.03	0.00	0.00	1.46
	90	94.60	0.00	0.00	0.00	1.60
	95	97.14	0.06	0.24	0.09	1.74
GARCH	**70**	47.94	0.00	0.25	0.00	0.59
	75	51.43	0.00	0.04	0.00	0.64
	80	57.46	0.00	0.39	0.00	0.69
	85	62.54	0.00	0.23	0.00	0.74
	90	67.94	0.00	0.28	0.00	0.80
	95	73.97	0.00	0.22	0.00	0.88
SES	**70**	63.17	0.01	0.21	0.01	0.75
	75	69.84	0.04	0.39	0.07	0.81
	80	73.02	0.00	0.16	0.00	0.86
	85	78.10	0.00	0.55	0.00	0.93
	90	83.81	0.00	0.77	0.00	1.00
	95	87.94	0.00	0.75	0.00	1.09
COMBINED	**70**	69.52	0.82	0.02	0.07	0.78
	75	74.60	0.85	0.03	0.08	0.85
	80	78.41	0.47	0.09	0.18	0.92
	85	83.49	0.44	0.09	0.17	0.99
	90	89.84	0.91	0.66	0.90	1.15
	95	95.24	0.86	0.74	0.93	1.32

Table 3: Efficiency of prediction intervals depending on the method.

Christoffersen tests are successfully passed, making the combined prediction intervals the most reliable approach among the methods investigated.

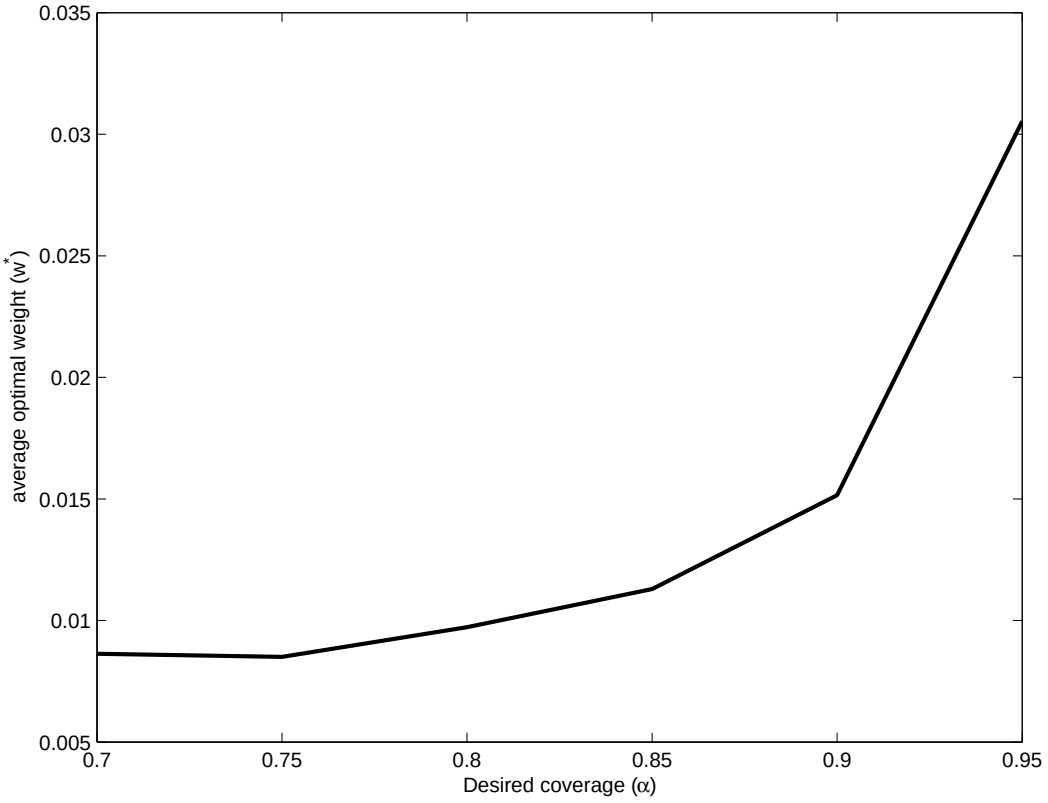

Figure 4: Relationship between optimal weigth (w^*) and desired coverage (α)

6. Conclusions

Forecasts of solar irradiation are required to incorporate the electricity generated into the grid. Recently, a large variety of approaches have proliferated so as to provide point forecasts. Nonetheless, the literature about the uncertainty associated to such forecasts is scarce. This work examines the aforementioned uncertainty through prediction intervals computed by means of non-parametric kernel density estimates and parametric approaches based on volatility forecasting models. The results show that the out-of-sample non-parametric approach is indicated to provide a high prediction interval

coverage but it also yields a high average interval width. Additionally, its main limitation is its lack of independence, i.e., actual observations out of the prediction intervals tend to cluster together. In order to overcome that limitation parametric techniques capable of estimating time-varying variances as GARCH or Single Exponential Smoothing applied to forecast errors are suggested. These techniques pass the independence test but they offer a prediction interval coverage excessively low. This work proposes a novel approach based on the simple combination of the non-parametric and parametric approaches. The results of the combined approach show a good compromise between coverage (unconditional and conditional) and average interval width, where the combination weight is obtained by maximizing the Christoffersen conditional coverage test p-value.

Further research should extend the results provided here in the following directions. Firstly, the forecasting horizon should be expanded. Here, we have just considered one-step-ahead as an starting point. However, further forecasting horizons would require potential seasonal correlations and thus, seasonal GARCH models or exponential smoothing versions able to deal with seasonality should be explored. Secondly, since Numerical Weather Prediction models can predict climate variables like cloud cover, this information may be incorporated to enhance the variability forecasts by means of, for example, adaptive exponential smoothing models, where the parameters are not constant but they depend on a key variable like the cloud cover. An idea in that direction has been investigated by Chu et al. (2015), where they take advantage of sky images to predict cloud cover and use it to enhance intra-hour prediction intervals calculation. Finally, it would be interesting to corroborate the results of this paper by computing solar irradiation prediction intervals from other geographic locations, as well as, for other key variables as the Direct Normal Irradiation, which is a crucial variable for concentrating solar power plants.

Acknowledgment

The authors are grateful to Alberto Martín and ISFOC for kindly providing the data used in this paper. This work was supported by the European Regional Development Fund and Spanish Government (MINECO/FEDER, UE) under the project with reference DPI2015-64133-R

References

Al-Sadah, F. H., Ragab, F. M., Arshad, M. K., 1990. Hourly solar radiation over bahrain. Energy 15 (5), 395 – 402.

Baig, A., Akhter, P., Mufti, A., 1991. A novel approach to estimate the clear day global radiation. Renewable Energy 1 (1), 119 – 123.

Boland, J., 1995. Time-series analysis of climatic variables. Solar Energy 55 (5), 377 – 388.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31 (3), 307 – 327.

Boudoukh, J., Richardson, M., Whitelaw, R. F., 1997. Investigation of a class of volatility estimators. The Journal of Derivatives 4 (3), 63–71.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., 1994. Time series analysis: Forecasting and Control. (3rd. Edition). Upper Saddle River, New Jersey: Prentice Hall.

Chatfield, C., 2000. Time-series forecasting. CRC Press.

Christoffersen, P. F., 1998. Evaluating interval forecasts. International economic review 39 (4), 841–862.

Chu, Y., Li, M., Pedro, H. T., Coimbra, C. F., 2015. Real-time prediction intervals for intra-hour DNI forecasts. Renewable Energy 83, 234 – 244.

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. Renewable and Sustainable Energy Reviews 27 (0), 65 – 76.

Dong, Z., Yang, D., Reindl, T., Walsh, W. M., 2013. Short-term solar irradiance forecasting using exponential smoothing state space model. Energy 55 (0), 1104 – 1113.

Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the Econometric Society 50, 987–1007.

Gardner, Jr., E. S., 1988. A simple method of computing prediction intervals for time series forecasts. Management Science 34 (4), 541–546.

Harvey, A., 1989. Forecasting Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Hurlin, C., Perignon, C., October 2013. CHRISTOF: MATLAB function to perform Christoffersen's (1998) tests of coverage.
URL `http://www.runmycode.org/companion/view/93`

Hyndman, R. J., Koehler, A. B., Ord, J. K., Snyder, R. D., 2008. Forecasting with Exponential Smoothing: The State Space Approach. Springer-Verlag, Berlin.

Isengildina-Massa, O., Irwin, S., Good, D. L., Massa, L., 2011. Empirical confidence intervals for usda commodity price forecasts. Applied Economics 43 (26), 3789–3803.

Kaplanis, S., 2006. New methodologies to estimate the hourly global solar radiation; comparisons with existing models. Renewable Energy 31 (6), 781 – 790.

Kleissl, J., 2013. Solar Energy Forecasting and Resource Assessment. Elsevier Science.

Kraas, B., Schroedter-Homscheidt, M., Madlener, R., 2013. Economic merits of a state-of-the-art concentrating solar power forecasting system for participation in the spanish electricity market. Solar Energy 93 (0), 244 – 255.

Lee, Y. S., Scholtes, S., 2014. Empirical prediction intervals revisited. International Journal of Forecasting 30 (2), 217 – 234.

Makridakis, S., Wheelwright, S. C., Hyndman, R. J., 1998. Forecasting: Methods and applications., 3rd Edition. John Wiley & Sons, New York.

McArthur, L. J. B., 2004. Operations manual. WMO/TD-N^o, 1274, WCRP/WMO. Tech. rep., Baseline Surface Radiation Network (BSRN).

Nelson, D. B., 1990. Stationarity and persistence in the garch(1,1) model. Econometric Theory 6 (3), 318–334.

Ord, J. K., Fildes, R., 2012. Principles of Business Forecasting. South-Western Cengage Learning, Mason, Ohio.

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

RiskMetrics, 1996. Riskmetrics technical document. Tech. rep., J.P. Morgan/Reuters.

Sfetsos, A., Coonick, A., 2000. Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. Solar Energy 68 (2), 169 – 178.

Silverman, B., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, Bristol.

Tashman, L., 2000. Out-of-sample tests of forecasting accuracy: an analysis and review. International Journal of Forecasting 16, 437–450.

Taylor, J. W., 2004. Volatility forecasting with smooth transition exponential smoothing. International Journal of Forecasting 20 (2), 273 – 286, forecasting Economic and Financial Time Series Using Nonlinear Methods.

Taylor, J. W., Bunn, D. W., 1999. A quantile regression approach to generating prediction intervals. Management Science 45 (2), 225–237.

Trapero, J. R., Kourentzes, N., Martin, A., 2015. Short-term solar irradiation forecasting based on dynamic harmonic regression. Energy 84, 289–295.

Williams, W. H., Goodman, M. L., 1971. A simple method for the construction of empirical confidence limits for economic forecasts. Journal of the American Statistical Association 66 (336), 752–754.

Wu, J. J., 2012. Semiparametric forecast intervals. Journal of Forecasting 31 (3), 189–228.

Yang, D., Sharma, V., Ye, Z., Lim, L. I., Zhao, L., Aryaputera, A. W., 2015. Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. Energy, 81, 111–119.