Solar irradiation forecasting: state-of-the-art and proposition for future developments for small-scale insular grids

Hadja Maïmouna Diagne, Philippe Lauret, Mathieu David

To cite this version:

HAL Id: hal-00918150
https://hal.archives-ouvertes.fr/hal-00918150
Submitted on 3 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ABSTRACT

Because of the great integration of solar energy on the electrical network solar irradiance forecasting is becoming essential. In fact, this integration can offer a better quality of service if the solar irradiation variation can be predicted with great accuracy.

This paper presents an in-depth review of the current methods used to forecast solar irradiation in order to facilitate selection of forecast method according to needs. Our work starts with a presentation of physical models and techniques based on cloud imagery. Next statistical approaches are detailed before to discuss about hybrid model. Finally, we give indications for future solar irradiation forecasting approaches dedicated to the management of small-scale insular grids.

Keywords: Solar irradiance, forecast models

1. INTRODUCTION

The integration of a large amount of photovoltaic (PV) plants into the electrical grid poses technical challenges due to the variable nature of the solar resource. Because of the exponential rate of growth registered recently in the renewable energies field, there is an increasing need for more accurate modeling, forecasting and prediction of solar irradiance. The forecasting of the PV production is especially helpful for the grid operators in order to better accommodate the variable generation of electricity in their scheduling, dispatching and regulation of power. The relevant horizons of forecast can range from 5 minutes to several days.

The most common solar measurements recorded by weather stations correspond to the total irradiation on a horizontal surface, so called Global Horizontal Irradiation or GHI. This parameter is required for the design of photovoltaic applications.

Depending on the forecast horizon different input data and forecasting models are appropriate. Time series models with on-site measured irradiance or power data as input are adequate for the very short-term time scale ranging from 5 minutes up to 6 hours (see Reikard, (1)). Forecasts based on cloud motion vectors from satellite images (Lorenz and al, (2)) show a good performance for a temporal range of 30 min to 6 h. For forecast horizons from about 6 h onwards, forecasts based on numerical weather prediction (NWP) models typically out-perform the satellite based forecasts Perez and al (3), Heinemann and al., (4).

Forecasting models for GHI are divided into two main groups. The first group is based upon the analysis of historical time series of global irradiation. The second group uses forecasted values from a numerical weather prediction (NWP) model and cloud imagery. Hybrid methods can improve some aspect of all of these methods. The models of the first group use the statistical approach to forecast mean hourly solar irradiation. The models in the second group use explanatory variables (mainly hourly mean cloud motion and direction derived from atmosphere) to predict global irradiation N-steps ahead. The models of the second group provide good results for the estimation of solar irradiance more than a day ahead. However, for the short-term horizon (ranging from 1h up to 5h forecasts), the influence of cloud motion becomes more important, so that the use of the models of the first group becomes essential (Glassley and al, (5)).

The paper is organized as follow. In section 2, the physical forecasting models proposed in the literature are reviewed. In section 3, statistical approaches are presented. In section 4, hybrid models are evaluated. Finally the section 5 is dedicated to trends for future solar irradiation forecasting on island environment.
2. PHYSICAL MODEL

Currently, physically based forecasting is primarily conducted using numerical weather prediction (NWP) and cloud observations by satellite or Total Sky Imager (TSI). NWP provides information up to several days ahead, however there are significant biases and random errors in the irradiance estimates (Remund and al., (6), Lorenz and al.(7), Perez and al.,(8) and Mathiesen and Kleissl (9)). The spatial resolution of NWP is coarse with a pixel of 100 km² for the most accurate ones. Most of the clouds remains unresolved in NWP. Frozen cloud advection based on GOES satellites images can provide accurate forecasts up to 6 h ahead (Perez and al., (8) and Schroeder-Homscheidt and al., (10)) at a resolution of 1 km².

2.1 Numerical Weather Prediction (NWP)

Forecasts beyond 6 h, up to several days ahead, are generally most accurate if derived from numerical weather prediction (NWP). NWP models predict GHI using columnar (1D) radiative transfer models (RTM, Heinemann and al., (4)). Heinemann and al. (4) showed that the MM5 mesoscale model can predict GHI in clear skies without mean bias error (MBE). However, the bias was highly dependent on cloud conditions and becomes strong in overcast conditions.

Perez and al. (11) examined the accuracy of the National Digital Forecast Database (NDFD), a derivative of the operational NWP models published by the National Center for Environmental Prediction (NCEP). After a local correction function was applied, Perez and al. (11) found that for 8–26 h forecast horizons, the NDFD had an hourly-average GHI rRMSE of 38%.

Remund and al. (6) evaluated different NWP-based GHI forecasts in the USA, reporting relative RMSE values ranging from 20% to 40% for a 24 h forecast horizon. Similar results were reported by Perez and al. (3) , evaluating NWP-based irradiance forecasts in several places in the USA. Remund and al. (6) examined NWP biases compared to a single site and find that ECMWF and Global Forecast System (GFS) next day GHI forecasts have an MBE of 19%. This MBE was found to be approximately constant for intra-day (hour-ahead) to 3 days ahead forecast horizons.

Lorenz and al. (12) evaluated several NWP-based GHI forecasts in Europe. Overall, results showed relative RMSE values of about 40% for Central Europe and about 30% for Spain. Evaluating European Centre for Medium-Range Weather Forecast (ECMWF), a NWP model, accuracy in Germany, Lorenz and al. (7) showed that NWP MBE was largest for cloudy conditions with moderate clear sky indices (0.3 < kₚ < 0.6), while forecasted clear conditions were relatively unbiased. They reported relative RMSE values of about 35% for single stations for a 24 h horizon forecasts.

To achieve high temporal and spatial resolution for intra-hour forecasts, NWP and satellite forecasts are currently inadequate. Ground observations using a sky imager present an opportunity to fill this forecasting gap and deliver a sub-kilometer view of cloud shadows over a large-scale PV power plant or an urban distribution feeder.

2.1.1 MOS correction

Model Output Statistics (MOS) is a post-processing technique used to objectively interpret numerical model output and produce site-specific forecasts. MOS relates observed weather elements to appropriate variables (predictors) via a statistical approach. These predictors may be NWP model forecast, prior observations, or geoclimatic data.

Consistent error patterns allow for MOS to be used to produce a bias reduction function for future forecasts. Bofinger and Heilscher (13) used MOS locally with ECMWF GHI forecasts to create daily solar electricity predictions accurate to 24.5% RMSE for averaged daily forecasts. Lorenz and al. (7) related forecasted solar zenith angle (SZA) and clear sky index to ECMWF MBE for Germany, revealing a consistent over-prediction (up to 100 W·m⁻²) for moderately cloudy conditions. Using a MOS correction function eliminated bias and reduced RMSE of hourly forecasts by 5% for 24 h forecasts.

Lorenz and al. (7) applied a stepwise multivariate fourth-order regression (see Rogers, (14)) to derive the MOS correction function.

\[
GHIc = GHI - \alpha \cos(\text{SZA}) - \beta(kₚ) + \delta
\]

GHIc is the model led Mean Bias Error (MBE) for a given solar zenith angle (SZA) and clear sky index.

(Mathiesen and Kleissl (9)) presented the analysis and MOS correction of GHI forecasts from three operational NWP models within the continental United States (North American Model, NAM, GFS, and ECMWF). They indicated that MOS application to the NWP irradiance output was successful in minimizing bias and reducing RMSE, but did not provide information as to the source of the MBE. MOS corrections in the measured clear sky regime (ktm > 0.9) did not reduce RMSE. This is because the MOS could not distinguish between RTM errors (over-prediction of GHI even for clear skies, especially for NAM) and cloud model errors (incorrect parameterization of RTM inputs). Consequently, many initially accurate forecasts were unnecessarily corrected. Differentiating between the sources of the error is important to selectively correct forecasts.

Lorenz and al. (7) found reduced errors for averaging over a 100 km × 100 km square grid for the ECMWF forecast.

2.1.2 NWP Limitation

A limitation of NWP forecasting is their coarse resolution. Even the 0.1° × 0.1° NAM spatial resolution is insufficient to resolve most clouds. Only an average cloud cover can be forecasted for a given point. For the GFS and ECMWF the resolution is even coarser.
However, even if the spatial resolution is finer, the temporal output intervals would not permit the assessment of time dependent cloud cover variability, important in predicting ramp rates and ranges of variability for solar power plants. Although NWP model time-steps are on the order of minutes, the radiative transfer models are run less frequently, and the output is only hourly (NAM) or every 3 h (GFS and ECMWF). Consequently, any patterns with characteristic time scales less than an hour are unresolved. Linking observed temporal variability in GHI to native NWP forecasts will require further research.

2.2 Cloud imagery

For physically-based forecasting, cloud cover and cloud optical depth are the most important parameters affecting solar irradiance. Through processing of satellite or ground imagery, clouds can be detected, characterized, and advected to predict GHI accurately up to 6 h in advance ahead. Hammer and al. (16) demonstrated achieved 17% rRMSE in satellite imagery for 30 minutes cloud index forecasts and 30% rRMSE at 2 h forecast horizons. For intra-day forecasts, a reduction in rRMSE by 7–10% compared to persistence forecasts was found.

Chow and al (17) presented a technique for intra-hour, sub-kilometer cloud shadow nowcasting and forecasting using a ground-based sky imager for selected days at the UC San Diego. This technique allows to obtain sky cover, cloud motion, cloud shadows, irradiance, and to forecast cloud locations.

Geostationary satellite images, such as those obtained from the METEOSAT satellite, have been used for the determination and forecasting of local solar radiation conditions. The basis of this method relies upon the determination of the cloud structures during the previous recorded time steps. Extrapolation of their motion leads to a forecast of cloud positions and, as a consequence, to the local radiation situation. This method has the advantage of producing a spatial analysis of an area within certain resolution capabilities. The improvement over the persistent method is small, according to the authors.

Only short deterministic forecast horizons are feasible using a single TSI at a site due to low clouds and large cloud variability at the fine spatial scale studied. Capturing these features deterministically is nearly impossible with satellites or NWP approaches Chow and al (17).

3. STATISTICAL MODELS

Several methods for forecasting solar irradiance in different time scales have appeared recently based on time series models, Artificial Neural Networks (ANNs), Fuzzy Logic (FL) and hybrid system such as ANFIS, ANN-wavelet, and ANN-genetic algorithms (see Bacher and al.(18)). These approaches can be classified into three different types (see Mellit and Kalogirou (19)):

The first one can estimate the solar irradiance (in different time scales) based on some meteorological parameters such as air temperature (T), relative humidity (RH), wind speed (Ws), wind direction(Wd), cloud (Cl), sunshine duration (S), clearness index (Kt), pressure (P), etc. and geographical coordinates as latitude. Mathematically, this approach can be generally expressed as:

\[ G_i = f(T_i, S_i, RH_i, Ws_i, Wd_i, Kt_i, Cl_i, P_i, Lat, Lon) \]

The problem consists in writing an approximate function \( f \) which allows to find a relationship between the inputs and the outputs data. Multilayer Perceptron (MLP) network, Radial Basis Function (RBF) network, and fuzzy logic can resolve this problem.

A second approach allows predicting the future solar irradiance (in different time scales) based on the past observed data Kemmoku and al, (20), Sfetsos and Coonick, (21), Cao and Cao, (22), Mellit and al, (23), Cao and Lin, (24), Cao and Cao, (25), Sfetsos and Coonick, (21) and Kaplanis and Kaplanis, (26)). Mathematically, this method can be formulated as:

\[ G_{t+p} = f(G_{t+p-1}, G_{t+p-2}, G_{t+p-2}, ..., G_t) \]

The problem consists in finding a relationship between the inputs and the outputs data which allows to predict the future time step \( t + p \) based on the observed data at the time \( t + p - 1, t + p - 2, ..., t \). Therefore, in this case time series models, recurrent neural networks (RNN), wavelet-networks, and wavelet-networks-fuzzy are very suitable.

The last one combines the two previous approaches. Mathematical expression of the method is:

\[ G_{t+p} = f(G_{t+p-1}, G_{t+p-2}, G_{t+p-2}, ..., G_t, T_t, S_t, RH_t, Ws_t, Wd_t, Kt_t, Cl_t, P_t, Lat, Lon) \]

In this case, the input parameters are the past observed solar irradiance data at the scale times \( t + p - 1, t + p - 2, ..., t \), and other meteorological parameters so that different ANN-architectures and ANFIS are adequate.

Many different models are used to predict the solar radiation time series, like the classic Autoregressive (AR) model, the Autoregressive and Moving Average (ARMA) model and the Markov Chains. Furthermore, adaptive methods such as the Time Delay Neural Network (TDNN), which has been proven to be reliable in predicting the future trend of a time series, can also be used to solve this problem.

3.1 Persistence

It is useful to check, whether the forecast model provides better results than any trivial reference model. It is worthwhile to implement and run a complex forecasting tool only if it is able to clearly outperform trivial models. The probably most common reference model in the solar or wind forecasting community for short term forecasting
is the Persistence model. The persistence model supposed that global irradiation data at $x_t$ is equal to global irradiation data at $x_{t-1}$. Persistence model also known as the naïve predictor, can be used to benchmark other methods. Persistence forecasts accuracy decrease strongly with forecast duration as cloudiness changes from the current state. Generally, persistence is an inaccurate method for more than 1 hour ahead forecasting and should be used only as a baseline forecast for comparison to more advanced techniques.

In Perez and al. (8), the single site performance of the forecast models is evaluated by comparing it to persistence.

3.2 ARMA model

The Autoregressive Moving Average (ARMA) model is usually applied to auto correlated time series data. This model is a great tool for understanding and predicting the future value of a specified time series. ARMA is based on two parts: autoregressive (AR) part and moving average (MA) part. Also, this model is usually referred as ARMA $(p, q)$. In which $p$ and $q$ are the order of AR and MA respectively.

The popularity of the ARMA model is its ability to extract useful statistical properties and the adoption of the well-known Box–Jenkins methodology. ARMA models are very flexible since they can represent several different types of time series by using different order. It has been proved to be competent in prediction when there is an underlying linear correlation structure lying in the time series. One major requirement for ARMA model is that the time series must be stationary. However, from a stationary test using Augmented Dickey–Fuller (ADF) test, the solar radiation series was found to be non-stationary. Thus a phase of detrending is needed to obtain the stationary series.

3.2.1 Process to obtain stationary solar radiation time series

A stationary solar radiation time series can be obtained by removal of seasonality and trend.

Deleting seasonality

Clear sky index is a reference model to obtain the deterministic daily variation of irradiance. A clear sky model estimates the global irradiance in clear sky conditions at any given time. Fourier series allow to delete cycles. Magnano and Boland (27) shown that it captures not only yearly cycles, but also intra-day cycles.

Detrending models

Because of its unpredictable noise, it is not easy to find the trend in a specific day’s series of solar irradiance. Several models exist to detrend the hourly solar radiation (see Baig and al., (28), Kaplanis (29)).

To judge the goodness of different detrending models, Wu and Chan (30) use the Augmented Dickey–Fuller (ADF) test to measure the stationarity of the detrended series. The ADF test is a test for unit root in a time series. If there is a unit root in time series, the time series is not stationary; otherwise, it should be stationary. Wu and Chan (30) finded that Al-Sadah’s model’s performance is the best in both aspects detrending and fitting.

After the detrending phase, Wu and Chan (30) applied a classical ARMA model to the stationary series, and checked its order according to auto correlation and partial correlation. They concluded that the best order of the model is an ARMA $(1, 1)$. The TDNN model was also used to predict the trend series. It was found to be much more sensitive than the ARMA model, but not as stable.

3.3 ARIMA techniques

The ARIMA (Auto-Regressive Integrated Moving Average) techniques (see Hamilton (15)) are reference estimators in the prediction of global radiation field. It is a stochastic process coupling autoregressive component (AR) to a moving average component (MA). ARIMA models allow to treat non-stationary series.

Reikard applies a regression in log to the inputs of the ARIMA models to predict the solar radiation. He compares ARIMA models with other forecast methods such as ANN. At the 24-h horizon, he states that the ARIMA model captures the sharp transitions in irradiance associated with the diurnal cycle more accurately than other methods.

3.4 Artificial Neural Network (ANN)

As an alternative to conventional approaches, artificial neural networks (ANNs) have been successfully applied for solar radiation estimation. Artificial neural networks (ANNs) recognize patterns in data and have been successfully applied to solar forecasting. Using training data, ANNs have been developed to reduce relative RMSE (rRMSE) of daily average GHI by as much as 15% when compared to 12–18 h ahead NWP forecasts (see Guarnieri and al.(31)).

Sfetsos and Coonick (21) use ANN to make one-step predictions of hourly values of global irradiance and to compare them with linear time series models that work by predicting clearness indexes. Heinemann and al. (4) use satellite images for horizons below 6 h; In Lorenz and al., (32) longer horizons of forecast produces by numerical weather predictions (NWPs) are used as input of an ANN to predict global irradiance.

Mellit and Pavan (33) developed a MLP neural network to forecast 24 h ahead the solar irradiance. The proposed model accepts as input parameters the mean daily irradiance and the mean daily air temperature; The output is represented by the 24 h of the ahead of solar irradiance. Performance prediction of a grid-connected PV plants at Trieste, Italy, have a correlation coefficient of more than 98% for sunny days and less than 95% for cloudy days.

Kemmoku and al, (20) used a multistage ANN to predict the insulation of the next day. The input data to the
network are the average atmospheric pressure, predicted by another ANN and various weather data of the previous day. Authors propose to use a multi-stage NN method for forecasting the insulation of the next day. The insulations forecast by the multi-stage and the single-stage neural networks are compared with the measured ones. The results show that the mean error (MBE) reduces from about 30% (by the single-stage) to about 20% (by the multi-stage).

Sfetsos and Coonick (21) introduced a simple approach for the forecasting of hourly solar radiation using various artificial intelligence based techniques (ANNs and ANFIS). They also investigated other meteorological variables such as temperature, wind speed, and pressure. A comparison between the various models in terms of prediction error and training time indicated that the network trained with the Levenberg–Marquardt algorithm (LM) network was as the optimum prediction model.

Mihalakakou and al., (34) developed a total solar radiation time series simulation model based on ANN and applied it in Athens. The Neural Logic Network was identified as the model with the least error. It incorporates Logic Rules that produced an RMSE of 4.9% lower than the persistent approach.

Time Delay Neural Network (TDNN) is developed from general feed forward neural network to obtain the relationship between the input and output position in time series (Fatih and al., (35). The conventional neurons of a neural network provide their response to the weighted output provided by a given layer depends on the output series (Fatih and al., (35). The conventional neurons of a relationship between the input and output position in time general feed forward neural network to obtain the performance.

A review on the application of the artificial intelligence techniques for modeling and forecasting of the solar radiation is presented Mellit (36). In this review, several approaches have been compared and analyzed.

4. HYBRID MODEL

The use of hybrid models has become more popular as it takes advantage of different models. The basic idea of the model combination is to use each model’s unique feature to capture different patterns in the data. Both theoretical and empirical finding suggests that combining different models can be an efficient way to improve the forecast performance.

Artificial intelligence techniques, such as fuzzy logic and neural networks, have been used for estimating hourly global radiation from satellite images. The results seem to point out that fuzzy and neural network models are better than regression models.

Cao and Cao in (22) and (25) developed a hybrid model for forecasting sequences of total daily solar radiation, which combines artificial neural network with wavelet analysis.

Cao and Lin (24) use an ANN (with a special designed training algorithm) combined with wavelets (based on diagonal recurrent wavelet neural network (DRWNN)) to predict next day hourly values of global irradiance. Different types of meteorological observations are used as input to the models; among others the daily mean global irradiance and daily mean cloud cover of the day are forecasted.

Wu and Chan (30) use a hybrid model of ARMA and TDNN to improve the prediction accuracy. They suppose that the daily solar radiation series is composed by linear and nonlinear component and used the ARMA model to fit the linear component and the TDNN model to find the nonlinear pattern lying in the residual. This hybrid model has the potential to harness the unique feature and strength of both models. It is more accurate than using the ARMA or TDNN model separately.

Reikard (1) compares a regression model, the UCM Model, an ARIMA, a transfer function Model, a neural network Model and an hybrid model. For this work, the author uses a logarithmic scaling of the input. Results show that for the resolutions of 60, 30 and 15min, the ARIMA model show better results.

5. FUTURE SOLAR IRRADIATION FORECASTING APPROACHES FOR SMALL-SCALE INSULAR GRIDS

Insular territories experience unstable electricity network and uses expensive means in order to provide the power for the peak demand. Their grids are generally not interconnected with any continent and all the electricity must be produce inside the territory. The power of grid connected PV plants increases fast and can interfere with the network stability. An efficient forecasting method
will help the grid operators to better manage the electric balance between the demand and the power generation. Kostylev and Pavlovski (37) identify three forecasting horizon (intra-hour, intra-day and day ahead) related to the grid operator activities (ramping events, variability related to operations, unit commitment, transmission scheduling, day ahead markets, hedging, planning and asset optimization). Based on our review, Figure 1 shows the relation between the forecasting horizons, the forecasting models and the related activities for grid operators.

Classification of forecasting model based on spatial resolution of input data and temporal resolution of output or foreseen data are illustrated by Figure 2.

Fig. 1: Relation between forecasting horizons, forecasting models and the related activities

Fig. 2: Classification of forecasting model based on spatial resolution of input data and temporal resolution

A summary of the literature on solar irradiance forecasting models illustrated by Figure 1 and Figure 2 gives indications for future work. However, a further consideration in choosing among forecasting models is their efficiency. Reunion Island which is a small territory with a high relief experiences a lot of microclimates and processes of cloud formation. In this case, the selection of forecasting model is based on small spatial resolution models and small temporal resolution of forecasting model. Based on these situation, we find that ECMWF and GFS that present reliable results are limited by their coarse resolution for Reunion Island even if a pre-
choice of the model depends critically on the horizon of forecast. At longer horizons, the data are dominated by the diurnal cycle. In this case the ARIMA models work better. At higher frequency, the data is more dominated by short-term patterns which can be picked up by persistence or ANN.

6. CONCLUSIONS

Solar irradiance forecasting is important for the integration of photovoltaic plants into an electrical grid. Proper solar irradiance forecasting helps the grid operators to optimize their electricity production and/or to reduce additional costs by preparing an appropriate strategy.

A number of physical methods and statistical techniques have been reviewed in this paper. From the description of the various results of solar irradiance forecasting, we retain that the choice of the appropriate forecasting models depends on forecast horizon and the available data. For forecast horizon from 6 h up to 3 days ECMWF associated with a MOS post-process shows the most accurate results. However, in the case of Reunion Island, WRF model seems to be more pertinent. For a smaller forecast horizon, from 5 minutes to 4 hours ARIMA seems to present the best accuracy. Cloud imagery and hybrid model can improve the results of forecasting when solar irradiance experiences a strong variability like in lot of insular territories.

Future work will include several elements to improve forecast accuracy. Sky imagery technique will be used to take in account process of cloud formation. The interesting methods identified here (WRF, ARIMA and AR) will be combined to sky imagery to yield a comprehensive and more accurate forecast product with different horizons of forecast. The goal is to reach the needs of grid operators.

7. ACKNOWLEDGEMENTS

The authors would like to thank the Reuniwatt company for his support to achieve the present work.

8. REFERENCES


(2) Lorenz E, Heinemann D, Hammer A. Short-term forecasting of solar radiation based on satellite data. Proceedings Eurosun (ISES Europe Solar Congress), 2004; Freiburg, Germany.


