Applied Energy 114 (2014) 218–226

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Bayesian rules and stochastic models for high accuracy prediction of solar radiation

Cyril Voyant [a,*], Christophe Darras [b], Marc Muselli [b], Christophe Paoli [b], Marie-Laure Nivet [b], Philippe Poggi [b]

[a] CHD Castelluccio, Radiophysics Unit, B.P85 20177 Ajaccio, France
[b] University of Corsica/CNRS UMR SPE 6134, Campus Grimaldi, 20250 Corte, France

HIGHLIGHTS

- Global radiation prediction and PV energy integration.
- Artificial intelligence and stochastic modeling in order to use the time series formalism.
- Using Bayesian rules to select models.
- MLP and ARMA forecasters are equivalent (nRMSE close to 40.5% for the both).
- The hybridization of the three predictors (ARMA, MLP and persistence) induces very good results (nRMSE = 36.6%).

ARTICLE INFO

Article history:
Received 27 May 2013
Received in revised form 7 August 2013
Accepted 17 September 2013
Available online 18 October 2013

Keywords:
Mutual information
Pressure
Artifical neural network
Autoregressive and moving average model
Hybrid
Bayes

ABSTRACT

It is essential to find solar predictive methods to massively insert renewable energies on the electrical distribution grid. The goal of this study is to find the best methodology allowing predicting with high accuracy the hourly global radiation. The knowledge of this quantity is essential for the grid manager or the private PV producer in order to anticipate fluctuations related to clouds occurrences and to stabilize the injected PV power. In this paper, we test both methodologies: single and hybrid predictors. In the first class, we include the multi-layer perceptron (MLP), auto-regressive and moving average (ARMA), and persistence models. In the second class, we mix these predictors with Bayesian rules to obtain ad hoc models selections, and Bayesian averages of outputs related to single models. If MLP and ARMA are equivalent (nRMSE close to 40.5% for the both), this hybridization allows a nRMSE gain upper than 14% points compared to the persistence estimation (nRMSE = 37% versus 51%).

1. Introduction

For 60 years, energy needs are multiplying exponentially to support economic developments, comfort and electricity consumption per capita. The gross inland energy consumption was multiplied by a factor 2 according to the world energy assessment report [1]. Currently, we are at one critical moment of the energy exploitation [2]: we carry out the brittleness and the inconsistency of our way. Indeed, the resources of planet in fossil sediment become exhausted, in addition to the economic consequences; so, it is necessary to find alternatives to the current energy sources [3]. In addition, the use of fossil fuels poses another problem: environmental impacts are massive. Even if it has long been ignored, the preservation of the environment is a global issue, with still significant economic challenges. Enhancing the use of renewable energy is one of the solutions but poses a number of challenges in terms of integration.

Because of their random and intermittent trend, the renewable energies must be integrated on a restricted basis in the electrical distribution grid. This parsimonious insertion is in order to protect it and to warrant quality of supply. In France, this limit was set to 30% of the instantaneous power by the ministerial order of April 23rd, 2008 [4]. To be able to increase the insertion rate of the renewable energies on the electrical distribution grid, solutions are studied and applied. In France for example, the CRE (French Energy Regulation Commission) studies the means to control the fluctuations in these intermittent energies [3]. The CRE is a French independent authority (created on March 24th, 2000) managing industrials tenders related to grid integration of "fatal" energy sources. The solutions proposed by the authority are stipulated in the call for tender of the PV energy and the WT energy [3]. These

* Corresponding author. Tel.: +33 495293666; fax: +33 495293797.
E-mail address: cyril.voyant@ch-castelluccio.fr (C. Voyant).

0306-2619/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apenergy.2013.09.051

solutions consist globally in coupling the renewable energies with a storage method (hydrogen, batteries, etc.). However, this coupling is not sufficient if the management of the storage is not mastered. So it is essential to be able to anticipate the renewable energies production. The issue of this paper is the global radiation prediction in order to effectively manage the storage. The association of storage and solar predictive methods allow to guarantee an available energy for the electrical distribution grid [5]. The storage absorbs the strong fluctuations and the surpluses of power and it shall fill the defect of power (we have a PV output power fluctuations smoothing). The manager of the electrical grid could thus estimate the available production of the next day of this power plant [6]; he could so, better manage the supply and demand adequacy. The most interesting prediction horizon for the grid manager is noted $h+24$ representing 24 predictions computed before 18:00 and covering hour-by-hour, the global solar radiation profile of the next day [7]. The goals of this work could be to integrate these predictions tools to a project like the MYRTE platform [3]. So we will know if an energy production system coupling a PV array and storage (electrolyzer, H2 and O2 tanks and fuel cell in our case) associated with a prediction tool, allows to be seen by the electrical distribution grid manager as a reliable energy supply. In this case, load profiles would be created from forecasted meteorological data, and so the storage/destocking would be also created one day ahead. More the state of tank is wrong and more the daily power supply failures are important. In this kind of platform, the storage does not support the PV but the grid by controlling the injected energy into the electrical network.

The global radiation forecasting is the name given to the process used to predict the amount of solar energy available. A lot of predictive methods have been developed by experts around the world. One of the most popular is certainly the numerical weather prediction using mathematical model of the atmosphere to predict the weather based on current weather condition (nRMSE close to 35%) [8]. The second family of models, often called stochastic models, is based on the use of the times series (TS) mathematical formalism [7,9,10]. A TS is described by sets of numbers that measures the status of some activity over time. It is a collection of time ordered observations x_t, each one being recorded at a specific time t (period) [11]. A TS model ($\hat{x}_t$) assumes that past patterns will occur in the future. TS prediction or TS forecasting takes an existing series of data $x_{t-k},\ldots,x_{t-2},x_{t-1}$ and forecasts the x_t data values. The goal is to observe or model the existing data series to enable future unknown data values to be forecasted accurately. Thus a prediction $\hat{x}_t$ can be expressed as a function of the recent history of the time series, $\hat{x}_t = f(x_{t-1}, x_{t-2}, \ldots, x_{t-k})$ [12–15]. In preliminary studies [16,17], we have demonstrated that an optimized multi-layer perceptron (MLP) with endogenous inputs made stationary and exogenous inputs (meteorological data) can forecast the global solar radiation time series with acceptable errors (10–20%) [15]. This prediction model has been compared to other prediction methods [16,17] (AutoRegressive and Moving Average called ARMA, k Nearest Neighbor called k-NN, Markov Chains, etc.) and the conclusion was that MLP and ARMA were the best predictors (nRMSE gain close to 2 points) and were similar for the horizon $h+1$ (prediction one hour ahead, nRMSE close to 15% [16]) and $j+1$ (prediction one day ahead, nRMSE close to 20% [17]). Moreover, we have shown in previous study [7] that MLP modeling of the global solar irradiation TS can be applied to the $h+24$ time horizon prediction (nRMSE lower than 30%). The results demonstrate a higher accuracy with MLP models than with the persistence method. Indeed, the use of MLP to predict the $h+24$ global radiation horizon is interesting but the chosen architectures, stationarization modes, and the choice of a multivariate analyze, modify greatly the results. Different levels of complexity can be considered, namely: multi-output MLP (with or without exogenous data), MLP committee (with or without exogenous data) or ARMA model.

Considering these findings we propose in this paper new methodology of TS stochastic modeling and of data preprocessing for the prediction of the PV energy (Section 2). Then we will present the results of the models mentioned above and expose the performance of prediction in order to cross compare models, and explain the use of Bayesian selection rules (Section 3) and then, finally discuss the use of this predictive approach in the case of a real coupling between a PV array and storage.

2. Methodologies

Most available global radiation measurements are global horizontal radiation, but it is very rare/uncommon to develop PV stations with no tilted PV modules. To use these historic measures while modeling the global radiation, it is therefore important to be able to tilt horizontal data. The next section will describe the adopted tilt methodology.

2.1. The horizontal data problem

Whatever the selected predictive tool, it is necessary in order to develop the majority of the stochastic models to use historical data of global radiation. Depending on the inclination of the PV panels, this type of measures does not often exist on important time intervals. However it is possible to determine the global solar irradiation on tilted plane ($H_{g\beta}$) from horizontal ones. To achieve this, we consider the components of the horizontal global radiation (H_{gh}; available on the French Meteorological Organization database), that is to say the direct component called beam (H_{bh}), and the diffuse component (H_{dh}). According to a past study made on the Mediterranean region [7,14] we decided to use the CLIMED 2 methodology in order to calculate the horizontal diffuse radiation, and the Klucher [18] approach to compute $H_{g\beta}$ (Eq. (1)).

$$H_{g\beta} = H_{bh} \cdot R_b + R_d \cdot H_{dh} + R \quad (1)$$

where R_b is related to the geometric projection (Eq. (2), where θ is the incident angle of the surface, φ is the latitude, δ is the declination, ω_h is the hour angle, β is the tilt angle and γ is the solar azimuth angle), R related to the ground scattering (Eq. (3) where ρ is the ground albedo) and R_d is related to the Klucher methodology (Eq. (4), where θ_z is the zenith angle).

$$R_b = \frac{\cos(\theta)}{\sin(h)} = \frac{(\sin\varphi\cos\beta - \cos\varphi\sin\beta\cos\gamma)\sin\delta + (\cos\varphi\cos\beta + \sin\varphi\sin\beta\cos\gamma)\cos\delta\cos\omega_h + \cos\delta\sin\beta\sin\gamma\sin\omega_h}{\sin\varphi\sin\delta + \cos\varphi\cos\delta\cos\omega_h} \quad (2)$$

$$R = \frac{1}{2} \cdot \rho \cdot H_{gh} \cdot (1 - \cos(\beta)) \quad (3)$$

$$R_d = \frac{1}{2} \cdot \left(1 + \cos\left(\frac{\beta}{2}\right)\right) \cdot \left(1 + F \cdot \sin^3\left(\frac{\beta}{2}\right)\right) \cdot (1 + F \cdot \cos^2(\theta) \cdot \sin^3(\theta_z))$$

where

$$F = 1 - \left(\frac{H_{dh}}{H_{gh}}\right)^2 \quad (4)$$

The various parameters involved in these equations are classical and are related to the celestial mechanics. The CLIMED2/Klucher methodology give us (Eq. (5) with clearness index defined by $k_t = \frac{H_{gh}}{H_0}$, and scattered ratio by $f = \frac{H_{dh}}{H_{gh}}$).