Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 102 (2014) 131–142

SOLAR ENERGY

www.elsevier.com/locate/solener

Time series modeling and large scale global solar radiation forecasting from geostationary satellites data

Cyril Voyant [a,b,*], Pierrick Haurant [c], Marc Muselli [a], Christophe Paoli [d], Marie-Laure Nivet [a]

[a] University of Corsica, CNRS UMR SPE 6134, 20250 Corte, France
[b] Castelluccio Hospital, Radiotherapy Unit, BP 85, 20177 Ajaccio, France
[c] Centre Thermique de Lyon (CETHIL), UMR CNRS 5008, INSA Lyon, France
[d] Galatasaray University, Computer Science Department, Çırağan Cad. No: 36, Ortaköy 34357, Istanbul, Turkey

Received 16 August 2013; received in revised form 11 January 2014; accepted 18 January 2014
Available online 11 February 2014

Communicated by: Associate Editor Christian Gueymard

Abstract

When a territory is poorly instrumented, geostationary satellites data can be useful to predict global solar radiation. In this paper, we use geostationary satellites data to generate 2-D time series of solar radiation for the next hour. The results presented in this paper relate to a particular territory, the Corsica Island, but as data used are available for the entire surface of the globe, our method can be easily exploited to another place. Indeed 2-D hourly time series are extracted from the HelioClim-3 surface solar irradiation database treated by the Heliosat-2 model. Each point of the map have been used as training data and inputs of artificial neural networks (ANN) and as inputs for two persistence models (scaled or not). Comparisons between these models and clear sky estimations were proceeded to evaluate the performances. We found a normalized root mean square error (nRMSE) close to 16.5% for the two best predictors (scaled persistence and ANN) equivalent to 35–45% related to ground measurements. Finally in order to validate our 2-D predictions maps, we introduce a new error metric called the gamma index which is a criterion for comparing data from two matrixes in medical physics. As first results, we found that in winter and spring, scaled persistence gives the best results (gamma index test passing rate is respectively 67.7% and 86%), in autumn simple persistence is the best predictor (95.3%) and ANN is the best in summer (99.8%).

Keywords: Time series; Artificial neural networks; Irradiance; Prediction; Gamma index

1. Introduction

The production and use of non-renewable resources based on fossil fuels combustion are responsible of real public health problem and raise environmental concerns. There are lots of alternatives such as photovoltaic and wind energy sources, which one of the main advantages are the renewable and inexhaustible aspects and the main disadvantage is related to their intermittencies (Hocaoglu, 2011; Voyant et al., 2012). These non-continuities can cause a demand/production unbalance involving irrelevant wind or solar systems uses. To overcome this problem, it is necessary to predict the resource and to manage the transition between different energies sources (Bouhouras et al., 2010; Darras et al., 2012; Muselli et al., 1998b). Considering the grid manager's point of view (Köpken et al., 2004), needs in terms of prediction can be distinguished according to the considered horizon: following days, next day by

* Corresponding author at: Castelluccio Hospital, Radiotherapy Unit, BP 85, 20177 Ajaccio, France. Tel.: +33 495293666, fax: +33 495293797.
E-mail address: cyril.voyant@ch-castelluccio.fr (C. Voyant).

http://dx.doi.org/10.1016/j.solener.2014.01.017
0038-092X/

Nomenclature

$X_t\ (x_i,\ y_j)$ X parameter concerning the pixel $(x_i,\ y_i)$ and the time t
$n_t\ (x_i,\ y_j)$ clearness of the atmosphere (unitless)
$\rho_t\ (x_i,\ y_j)$ measured albedo (unitless)
$\rho_t^{\mathrm{cloud}}(x_i, y_j)$ Albedo of the brightest clouds (unitless)
$\rho_t^{\mathrm{CS}}(x_i, y_j)$ Albedo of the ground under clear sky (unit less)
$\mathrm{CSI}_t(x_i,\ y_j)$ clear sky index (unitless)
$I_t\ (x_i,\ y_j)$ global radiation (W h/m^2)
$I_t^{\mathrm{CS}}(x_i, y_j)$ global radiation under clear sky (W h/m^2)
p number of parameters used for create model
$\varepsilon_{t+1}\ (x_i,\ y_j)$ prediction error (measurement-prediction)
f_n linear or non-linear model
$f,\ g,\ \omega_{ij}^1, b_i^1, \omega_i^2 + b^2,\ H$ and In MLP parameters concerning activation functions, weights and bias, number of hidden and input nodes
$X_t/x = \{x_1 \dots x_n\}$ $Y_t/y = \{y_1 \dots y_n\}$ time series X_t and Y_t and possible associated values
L^{τ} lag operator and associated order
$H(X_t), H(X_t|Y_t)$ and $\mathrm{MI}(X_t,\ Y_t)$ marginal/conditional entropies and the mutual information (bit)
$p(x), p(y)$ and $P(x,\ y)$ marginal and joint probabilities distribution function of X_t, and Y_t
nRMSE normalized root mean square error (%)
$r_p,\ r_m$ pixels distance (polar coordinate) concerning predicted and measured map ($= \sqrt{x_i^2 + y_i^2}$, m)
$r(r_p,\ r_m)$ distance between pixels from predicted (r_p) and measured map (r_m) (m)
γ and $\Gamma(r_p, r_m)$ gamma index and gamma score (unitless)
Tol_r and Tol_I distance (also called DTA) and intensity tolerances (m, W h/m^2)

hourly step, next hour and next few minutes. We choose in this paper to focus only on h + 1 horizon prediction of global radiation as a first step (one hour in advance). Of course, we are aware concerning the importance of other horizons (Voyant et al., 2013b). Note that it is appreciable to know the eventual fluctuations at least 30 min ahead (ignition delay of turbines) for an ideal electrical grid management (Troccoli, 2010). With efficient prediction tools dedicated to grid managers, the PV part in the mix energy would be increased; actually in France, the intermittent energy contribution is limited to 30%.

Several prediction methods have been developed by experts and can be divided in three main groups: methods using mathematical formalism of times series (TS) (De Gooijer and Hyndman, 2006; Elminir et al., 2007; Hamilton, 1994), numerical weather predictions (NWP) and models based on clouds detection (Inness and Dorling, 2012; Perez et al., 2013). In this study, we have chosen to study prediction methods of the first group and we will study if this methodology can be an alternative to the NWP models. Not that in the literature, the NWP models are compared against ground measurements and the error established is approximately 30–40% but depends on the orography and micro-climate studied (Paulescu, 2013). The time series formalism (TS) and modeling is often used in 1 dimensional (1-D) global radiation predictions, i.e. related to one measurement system at ground level (Pons and Ninyerola, 2008). Persistence, autoregressive models, multilayer perceptron (MLP) and more widely artificial neural network (ANN) often applied to this aim (Hocaoglu, 2011; Mellit et al., 2009; Voyant et al., 2012).

In this paper, we will complete the first prediction results exposed in (Haurant et al., 2013) and we will show that the TS formalism applied to surface solar irradiances (SSI) estimations reduced to one point (1-D approach) can be generalized in the 2-D case, even if no ground detector is present. Alternative approaches are available in (Loyola R., 2006; Rahimikhoob et al., 2013). From this point of view, satellite derived SSI maps extracted from the HelioClim-3 (HC-3) (Rigollier et al., 2004) database and centered on Corsica are used as hourly 2-D data generator. Each data series is processed with stochastic estimators in order to generate 1158 predictions per hour (1158 pixels per map separately treated). In fact, for an overall year of prediction, it is necessary to generate more than 10 million of hourly predictions (24 × 365 × 1158). The purpose of this paper is to generate one hour in advance predicted global radiation maps of a specific area from HC-3 SSI maps. But as data used are available for the entire surface, the method can be easily generalized. The geographical effects are taken into account using clear sky index in addition to temporal or seasonal phenomena (Allan, 2011). The uncertainty of the used satellite derived SSI maps is about 16–23% (http://www.soda-is.com/eng/helioclim/helioclim3_uncertainty_eng.html). Moreover, in this paper, inputs of stochastic models are previously measured values, however, in a previous paper we have shown that exogenous data (parameters such as temperature and air pressure) improve the prediction efficiency (Voyant et al., 2013b). As it was our first experiment in using geostationary satellites data, we have preferred to start without the multivariate case.

In the next section, the Meteosat images acquisitions, the SSI computation by Heliosat-II model described in (Rigollier et al., 2004; Gueymard, 2012) and clear sky index computing methodologies (Maini and Agrawal, 2006) will be first explained. Then we will detail the methodologies of prediction we have tested, taking care to explain first