

# ANALOG NOWCASTING OF SOLAR IRRADIANCE FROM GEOSTATIONARY SATELLITE IMAGES

Alex Ayet<sup>1,2</sup> and Pierre Tandeo<sup>3</sup>

**Abstract**—Accurate forecasting of Global Horizontal Irradiance (GHI) is essential for the integration of the solar resource in an electrical grid. We implement a novel data-driven model for up to 6h probabilistic forecasting of GHI. Cloud dynamics are emulated using an analog method on a geostationary satellite database (herein 5 years of hourly images). It contains both the images to be compared to the current meteorological conditions and their successors at one or more hours of interval. No approximation is thus made on the physics of the system, unlike numerical weather forecast. The algorithm is computationally efficient and requires no tuning. It is designed to be easily used on different locations, requiring only GHI satellite images.

## I. INTRODUCTION

In the context of a growing need for sustainable energy, the solar resource ranks among the most promising solutions to meet this upcoming demand. However, the intermittent nature of the production makes its integration into an electrical grid challenging. The main input for most solar power generation systems is Global Horizontal Irradiance (GHI), and its accurate probabilistic and deterministic forecasting is thus essential.

Depending on the forecast horizon, different approaches are used for GHI forecasting (see the reviews [1], [2]). Satellite images have proven to be efficient for the intraday horizon (up to six hours). The popular cloud motion vector methods ([3], [4]) estimate a motion field from successive cloud satellite images, to then advect the clouds, producing the forecast. The main drawback of this methods is the need for post-processing to take into account the cloud dissipation and deformation. Analog methods are also used, e.g. [5] that combines outputs from numerical models and in-situ data as features of a  $k$ -nearest neighbors algorithm. However, to satisfy the forecasting demand for a big amount of different

sites, in locations where numerical models and in-situ observations are sparse, a robust and easy to use method is still needed.

For precipitation nowcasting, atmospheric analogs have become an important topic (e.g. [6], [7]) due to the availability of huge radar datasets. Analogs represent two atmospheric states closely resembling each other [8], with the hypothesis that these states evolve similarly. The forecast is thus issued by finding similar states in an historical database (the analogs), and considering how the atmosphere evolved following these states (the successors). The whole physics of the system is thus contained in the analog-successor pair.

The aim of this paper is to present an operational method to forecast GHI over a *precise* solar energy source (e.g. a solar photovoltaic panel). It uses only one source of data: hourly satellite images of GHI, which are easily accessible for different locations. Finally, the method needs no tuning, meaning that it can be easily applied to forecast the irradiance over sites in different locations, with different climatic conditions.

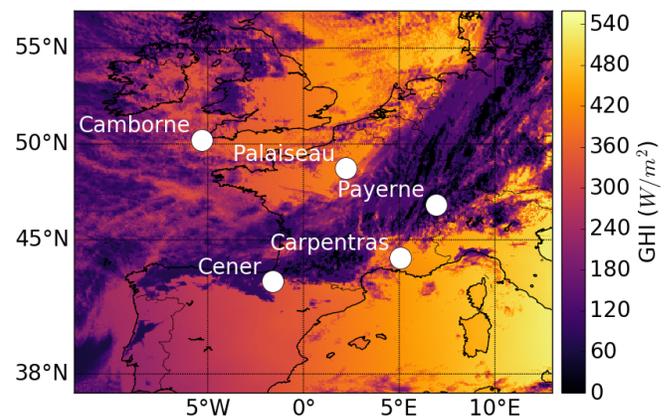


Fig. 1: OSI SAF satellite GHI image with selected BSRN stations in Europe on July 2<sup>nd</sup>, 2016.

Corresponding author: A. Ayet, alex.ayet@ens.fr <sup>1</sup>Elum Energy, Paris, France <sup>2</sup>École Normale Supérieure, Paris <sup>3</sup>Institut Mines-Télécom Atlantique, Brest, France

## II. DATA AND SETUP

To demonstrate the method, we use an archive of 18,521 GHI images obtained from the geostationary satellite Meteosat and processed by the Ocean and Sea Ice Satellite Application Facility (OSI SAF, [9]) covering western Europe and Africa. The images are remapped on a regular grid of  $0.05^\circ$ , and interpolated to produce hourly maps. The archive extends from Sept. 6<sup>th</sup>, 2011 to Dec. 31<sup>st</sup>, 2016. We use the 2016 year as a test year and the rest of the archive as the training set. The method is tested at the location of five stations of the Baseline Surface Radiation Network (BSRN, see [10]) where in-situ pyrgeometer data is available. The stations, shown in Fig. 1, cover a wide range of climatic situations (oceanic, mountain, continental) and constitute a good framework to test the robustness of the method.

The variability of GHI, hereinafter noted  $G$ , is due both to the daily and seasonal solar cycle (the clear sky contribution  $G_{clr}$ ) and to the cloudiness. Since clouds only reduce GHI, the clear sky at a location  $(x, y)$  is obtained as

$$G_{clr}(t, x, y) = \max_{t' \in \mathcal{S}(t)} G(t', x, y), \quad (1)$$

with  $\mathcal{S}(t)$  a 3-month interval around time  $t$ , with constant hour. This is a general expression that does not require any clear sky model (e.g. [11]). The cloud index  $c$  (between zero and one) is then defined as

$$G = (1 - c)G_{clr}. \quad (2)$$

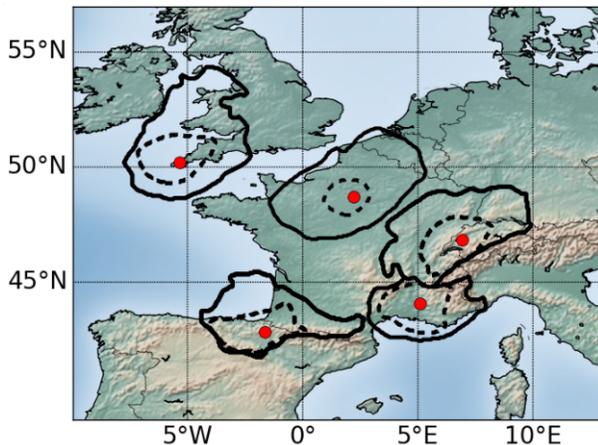


Fig. 2: Correlation masks for different BSRN sites (red dots) for Jan. 1<sup>st</sup> (full line) and July 1<sup>st</sup> (dotted line).

## III. METHODOLOGY

The main GHI variability being due to change in cloudiness, the analog method presented herein forecasts the cloud index, which is then converted to GHI using Eq. (2).

### A. Correlation Mask

For a given site of coordinates  $(x_s, y_s)$ , it is crucial to automatically select the zone in which the analogs are looked for. A daily correlation map  $C^m$  (for a day  $d$ ) between the pixel of interest and the surrounding region is computed. We use a metric inspired by [12] that measures the average spatial extension of cloud structures around the site

$$C^m(d, x, y) = \frac{\overline{c(t, x_s, y_s)c(t, x, y)}^d}{\left[ \overline{c(t, x_s, y_s)^2}^d \overline{c(t, x, y)^2}^d \right]^{1/2}}, \quad (3)$$

where the averages  $\overline{\cdot}^d$  are temporal within a 3-month interval around  $d$ . The region where the correlation is higher than 0.9 is then selected. Examples of masks are presented in Fig. 2. In the following, all computations are performed considering only image pixels in the mask corresponding to the site and day of forecast.

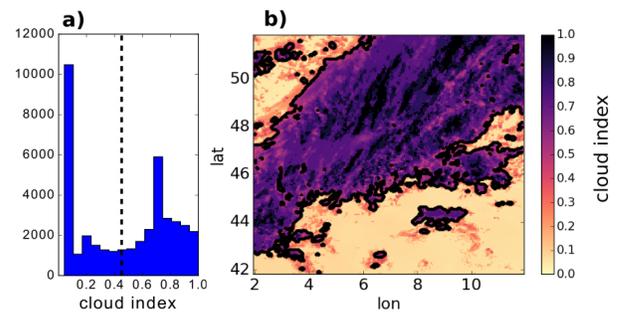


Fig. 3: (a) Histogram of cloud index from image (b) on the 2<sup>nd</sup> of July 2016 at Payerne. The threshold in (a) corresponds to the thick black line in (b).

### B. Analog forecasting

The forecast consists in two steps: the analogs selection and aggregation. To avoid overfitting and for computational efficiency, the analogs selection is done considering images compressed in a four dimensional space. For a given cloud index image, Otsu's method [13] (similar to a bimodal Fisher's discriminant analysis on a histogram of cloud index) is used to obtain a threshold

separating the clear sky and the cloud pixels ( $c$  below and above the threshold respectively, see Fig. 3). The images (observed and database) are compressed into four features between 0 and 1:

- 1) the *cloud fraction*: number of cloud pixels over the total number of pixels in the mask
- 2) the *cloud spread*: number of cloud pixels over the number of pixels in the convex hull of the clouds. It tends to one when there is only one cloud (its convex hull is nearly identical to the cloud itself) and to zero when there are many separate clouds
- 3) the *clear sky intensity*: the mean cloud index of the clear sky pixels
- 4) the *cloud intensity*: the mean cloud index of the cloud pixels.

Following [7], the database is first shrunk by considering images only within a time of the year (3-month window) and time of the day ( $\pm 3$ h) interval with respect to the date at which the forecast is to be issued. This increases the likelihood of finding similar convective and advective patterns. Then, the  $k$ -nearest neighbors of the observed image are selected using the Euclidean distance in the four-dimensional features space. For the BSRN sites of this study, the optimal number of neighbors is close to  $k = 90$ .

Next, the selected analogs are aggregated to produce a probabilistic forecast. For a given  $k$ -th analog, we determine the optimal spatial translation  $\delta$  such that the correlation  $C_k^s$  between the analog cloud index  $c_k^a$  and the observed cloud index  $c^o$

$$C_k^s(\mathcal{T}_\delta) = \frac{\langle c^o(x, y) \mathcal{T}_\delta c_k^a(x, y) \rangle}{[\langle c^o(x, y)^2 \rangle \langle \mathcal{T}_\delta c_k^a(x, y)^2 \rangle]^{1/2}} \quad (4)$$

is maximal ( $\mathcal{T}_\delta \cdot$  is the translation operator).

A local linear operator (see [14] or [15] for more details) is then applied on the translated images: it consists in fitting a linear regression between the translated analogs and successors, taking into account the weights computed in Eq. (4). The regression operator is then applied to the current observation to provide the analog nowcasts.

#### IV. EVALUATION

The deterministic forecast of GHI (the mean of the predicted Gaussian) is evaluated with the normalized Root

Mean Squared Error (RMSE) for a set of validation observations  $\mathcal{S}$

$$\text{RMSE} = \frac{\sqrt{\sum_{s \in \mathcal{S}} (c_s^o - \hat{c}_s)^2}}{\sum_{s \in \mathcal{S}} c_s^o} \quad (5)$$

with  $c_s^o$  the observed cloud index from a satellite image at the BSRN sites, and  $\hat{c}_s$  the corresponding forecast. The analog forecast is compared to an Eulerian persistence (keeping the last cloud index observation frozen), and a hourly climatology, obtained by taking the hourly average of GHI in the train dataset for days in a 2-weeks interval around the forecasted day. Results are given in Fig. 4 and indicate good performance and robustness to different locations. In all cases, the analog nowcasting procedure reaches better performances than the persistence and climatology method.

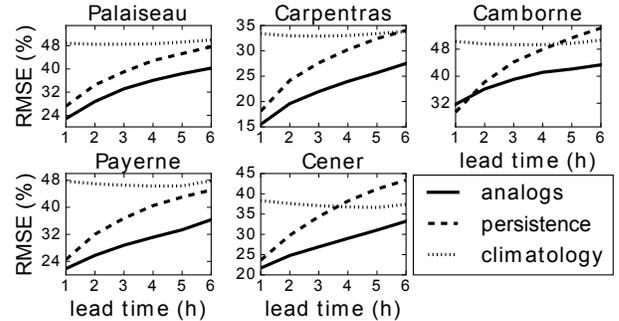


Fig. 4: Normalized RMSE for the five BSRN sites.

#### V. CONCLUSION AND PERSPECTIVES

We have presented a computationally efficient method for GHI analog nowcasting on a particular site. The method uses a  $k$ -nearest neighbors algorithm on a four-dimensional feature-space of cloud index to then apply a local regression between selected analogs and successors. The methodology has proven to be robust to different geographical locations, and requires no tuning, no in-situ data nor a numerical weather model.

The method will be extended by downscaling predictions on a particular site. The analogs times will be used to select historical in-situ data, using the BSRN data also available during the period 2011-2016. An aggregation operator will then be applied to forecast the in-situ production.

#### ACKNOWLEDGMENTS

The first author acknowledges the funding provided by Elum Energy as well as the supports from IMT-Atlantique and Elum R&D teams.

## REFERENCES

- [1] D. Heinemann, E. Lorenz, and M. Girodo, "Forecasting of solar radiation," *Solar energy resource management for electricity generation from local level to global scale*. Nova Science Publishers, New York, 2006.
- [2] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65–76, 2013.
- [3] A. Hammer, D. Heinemann, E. Lorenz, and B. Lücke, "Short-term forecasting of solar radiation: a statistical approach using satellite data," *Solar Energy*, vol. 67, no. 1, pp. 139–150, 1999.
- [4] H. Escrig, F. Batlles, J. Alonso, F. Baena, J. Bosch, I. Salbidegoitia, and J. Burgaleta, "Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast," *Energy*, vol. 55, pp. 853–859, 2013.
- [5] S. Alessandrini, L. Delle Monache, S. Sperati, and G. Cervone, "An analog ensemble for short-term probabilistic solar power forecast," *Applied energy*, vol. 157, pp. 95–110, 2015.
- [6] L. Panziera, U. Germann, M. Gabella, and P. V. Mandapaka, "Nora – nowcasting of orographic rainfall by means of analogues," *Q. J. R. Meteorol. Soc.*, vol. 137, no. 661, pp. 2106–2123, 2011.
- [7] A. Atencia and I. Zawadzki, "A comparison of two techniques for generating nowcasting ensembles. part ii: Analogs selection and comparison of techniques," *Mon. Weather Rev.*, vol. 143, no. 7, pp. 2890–2908, 2015.
- [8] E. N. Lorenz, "Atmospheric predictability as revealed by naturally occurring analogues," *Journal of the Atmospheric sciences*, vol. 26, no. 4, pp. 636–646, 1969.
- [9] P. Le Borgne, G. Legendre, and A. Marsouin, "Meteosat and goes-east imager visible channel calibration," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 11, pp. 1701–1709, 2004.
- [10] A. Ohmura, H. Gilgen, H. Hegner, G. Müller, M. Wild, E. G. Dutton, B. Forgan, C. Fröhlich, R. Philipona, A. Heimo, et al., "Baseline surface radiation network (bsrn/wcrp): New precision radiometry for climate research," *Bulletin of the American Meteorological Society*, vol. 79, no. 10, pp. 2115–2136, 1998.
- [11] D. Cano, J.-M. Monget, M. Albuissou, H. Guillard, N. Regas, and L. Wald, "A method for the determination of the global solar radiation from meteorological satellite data," *Solar Energy*, vol. 37, no. 1, pp. 31–39, 1986.
- [12] I. I. Zawadzki, "Statistical properties of precipitation patterns," *Journal of Applied Meteorology*, vol. 12, no. 3, pp. 459–472, 1973.
- [13] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [14] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American statistical association*, vol. 74, no. 368, pp. 829–836, 1979.
- [15] R. Lguensat, P. Tandeo, P. Ailliot, M. Pulido, and R. Fablet, "The analog data assimilation," *Monthly Weather Review*, 2017, published online, <https://doi.org/10.1175/MWR-D-16-0441.1>.