Solar Energy
April 2018, Volume 164 Pages 301-315
http://dx.doi.org/10.1016/j.solener.2018.02.068
http://archimer.ifremer.fr/doc/00431/54229/

Nowcasting solar irradiance using an analog method and geostationary satellite images

Ayet Alex [1, 2, *], Tandeo P. [3]

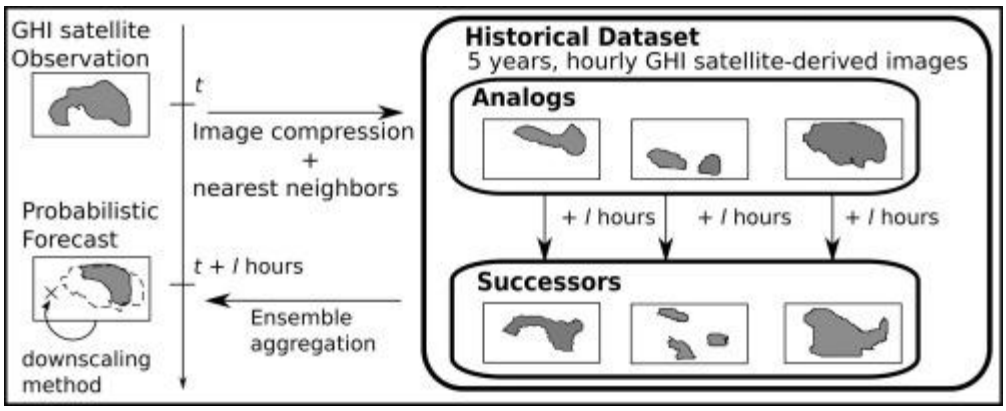
[1] Elum Energy, Paris, France
[2] Ifremer, CNRS, IRD, UBO/Laboratoire d'Océanographie Physique et Spatiale (LOPS), UMR 6523, IUEM, Plouzané, France
[3] IMT Atlantique, LabSTICC, UBL, 29238 Brest, France

* Corresponding author : Alex Ayet, email address : alex.ayet@ifremer.fr

Abstract :

Accurate forecasting of Global Horizontal Irradiance (GHI) is essential for the integration of the solar resource in an electrical grid. We present a novel data-driven method aimed at delivering up to 6 h hourly probabilistic forecasts of GHI on top of a localized solar energy source. The method does not require calibration to adapt to regional differences in cloud dynamics, and uses only one type of data, covering Europe and Africa. It is thus suited for applications that require a GHI forecast for solar energy sources at different locations with few ground measurements. Cloud dynamics are emulated using an analog method based on 5 years of hourly images of geostationary satellite-derived irradiance, without using any numerical prediction model. This database contains both the images to be compared to the current atmospheric observation and their successors at one or more hours of interval. The physics of the system is emulated statistically, and no numerical prediction model is used. The method is tested on one year of data and five locations in Europe with different climatic conditions. It is compared to persistence (keeping the last observation frozen), ensemble persistence (generating a probabilistic forecast using the last observations) and an adaptive first order vector autoregressive model. As an application, the model is downscaled using ground measurements. In both cases, the analog method outperforms the classical statistical approaches. Results demonstrate the skill of the method in emulating cloud dynamics, and its potential to be coupled with a forecasting algorithm using ground measurements for operational applications.

Graphical abstract



Highlights

► An novel statistical method to forecast Global Horizontal Irradiance is presented. ► Only one source of data is used: hourly geostationary satellite images. ► A framework to assess the quality of the analog method is proposed. ► The method is downscaled on a solar panel using on site measurements.

Keywords : Satellite-derived irradiance, Short term forecasting, Analog method, Geostationary satellite, PV

1. 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 its production makes its integration into an electrical grid challenging. Accurate forecasting of solar production is essential to ensure the stability of the grid and to optimize energy consumption. The main input for most solar power generation systems is Global Horizontal Irradiance (GHI), and its accurate probabilistic and deterministic forecasting is essential. The main source of variability in GHI is clouds, and thus most GHI forecasting algorithms aim at predicting cloud dynamics. The best performing method for GHI forecasting depends on the forecast

horizon (see the reviews Heinemann et al., 2006; Diagne et al., 2013). For intra-hour forecasting, machine learning methods (Marquez et al., 2013) or on site cloud tracking methods (Marquez and Coimbra, 2013) have been developed, while for more than six hours ahead and day ahead forecasts, Numerical Weather Prediction (NWP) forecasts are generally used as the primary source of information (Mathiesen et al., 2013; Thorey et al., 2015). Using satellite images provides information on horizontal cloud structures and has proven to be efficient for the intra-day horizon (see the recent review by Yang et al., 2018). The widely studied cloud motion vector methods (Hammer et al., 1999; Lorenz et al., 2004; Escrig et al., 2013) estimate a motion field from successive cloud satellite images and produce a forecast by advecting the clouds. The main drawback of these methods is the need for post-processing to take into account cloud dissipation and deformation. Numerical weather prediction models are another option, but require a well tuned regional atmospheric model (Mathiesen and Kleissl, 2011; Perez et al., 2010). However, to satisfy a forecasting demand for a large number of sites at different locations where ground observations are sparse, a robust and easy-to-use method is still needed.

The analog method, first introduced by Lorenz (1969), has gained renewed interest over the last few years, due to the availability of huge data sets and its computational efficiency. In particular, atmospheric analogs using ground radar data are used for precipitation nowcasting (Panziera et al., 2011; Atencia and Zawadzki, 2015). For a given observation of the atmosphere, analogs are past atmospheric states with an evolution which is assumed to be similar to that of the observed state. The forecast is issued by first running

a k-nearest neighbor algorithm on a historical data set of atmospheric states, finding the nearest neighbors to the observed state vector (the analogs). The past evolution of the analogs (called the successors) gives the prediction. The physics of the system is thus contained in the analog-successor pair, which is a real state of the atmosphere. Finding analogs with similar evolution to that of the current atmospheric state implies using a large historical data set (Van den Dool, 1994), and choosing an appropriate horizontal spatial domain over which analogs are compared to the observation (Root et al., 2007). For instance, a strong orographic forcing (as in Panziera et al. 2011), ensures a link between different atmospheric variables over a reduced region and helps identify good analogs by looking at cloud structures only in this domain. In general, the choice of the domain is crucial to constrain the problem such that only meaningful information is considered to identify analogs (see Lguensat et al. 2017 for a detailed discussion on local versus global analogs). Adding a temporal constraint to the analog selection, i.e. considering past states that are at the same phase of a diurnal and/or seasonal cycle as the observation (e.g. Atencia and Zawadzki 2015), is also an additional way to improve the quality of the analogs. Finally, the method is sensitive to the choice of the features on which the analogs are selected. Indeed, both for computational efficiency and to avoid overfitting, it is convenient to compress the state vectors in a space of reduced dimension (the feature space) before running the k-nearest neighbor algorithm. Features represent the information determining atmospheric evolution and are another way to constrain the problem. However, as pointed out by Atencia and Zawadzki (2015), it is difficult to evaluate the quality of a given choice of features other than by

testing different configurations.

In the context of GHI forecasting, analog methods have been developed using a combination of ground measurements and NWP outputs. The aim of the methods is not to emulate atmospheric dynamics, as for precipitation nowcasting methods, but rather to apply a post-processing to NWP forecasts using a historical data set. The Pattern Sequence-based Forecasting method, introduced by Alvarez et al. (2011) and further developed by Wang et al. (2017) uses the aforementioned information as an input of a clustering algorithm. A unique label is assigned to each day of the data set. The last few observed labels define a temporal pattern, and similar past occurrences of this pattern are retrieved from the historical data set, providing a forecast for the next day. Hourly forecasts of GHI are performed in (Alessandrini et al., 2015) using past NWP outputs as features to identify analogs to a current NWP forecast. The concurrent past observations and successors are then used as an ensemble, providing a probabilistic forecast of GHI in place of the NWP forecast. Note that this method has been further combined with a neural network (Cervone et al., 2017), also used in the context of nowcasting by Aguiar et al. (2015).

Both GHI and precipitation prediction methods require cloud forecasting. There are, however, two fundamental differences between the analog methods used for precipitation nowcasting (Panziera et al., 2011; Atencia and Zawadzki, 2015) and for GHI forecasting (Alvarez et al., 2011; Alessandrini et al., 2015). First, the type of data is different. Global Horizontal Irradiance analog methods use NWP outputs to identify analogs, as opposed to only measurements for precipitation nowcasting. Outputs from NWP have

the advantage of providing different variables apart from GHI (e.g. ground temperature, pressure, wind speed) that can be used as features to identify analog situations. However, they require a well tuned regional atmospheric model over the region of interest, i.e. which contains all the relevant physical processes governing GHI variability over a given region. Apart from the computational cost of tuning and running such a model, a correct representation of the physics can be challenging in specific regions, for instance close to a mountain range. Precipitation nowcasting methods, by using only measurements to identify analogs, emulate the physics governing GHI variability with a statistical model (the k-nearest neighbor algorithm), and do not rely on any physical parametrization of the atmosphere. Using outputs from a well tuned NWP model usually guarantees better performances at longer lead times (beyond 6h).

The second difference is the type of information used to identify analog situations. The GHI forecasting methods presented above are based on temporal patterns: temporal sequences of NWP outputs above a given site are used, and additional information is added by using the different variables of the NWP models. On the other hand, using ground radar data or satellite-derived irradiance maps means using mainly spatial patterns to find similar physical situations (even-though temporal information can also be included by looking at sequences of images). Note that spatial information has been used for GHI forecasting using a gridded output of an NWP model by Davò et al. (2016).

In this paper, we present a methodology based on the analog method to forecast GHI over a solar energy source. Analogs are identified by compar-

ing spatial cloud patterns from hourly images of satellite-derived irradiance. These are easily available over Europe and Africa, and no additional data are used. The method needs no calibration of physical parameters to adapt to different climates. It is inspired from precipitation nowcasting analog methods, while carefully addressing the points highlighted above: the choice of the spatial domain is automatic and physical, and the choice of the features is carefully tested within a novel framework. In Section 2, we present the satellite and ground data. The analog algorithm is described in detail in Section 3. Section 4 presents a novel framework to evaluate the validity of the selected analogs, justifying some of the choices made in the methodology. Three reference statistical methods are then presented in Section 5, to which the analog method is compared in terms of standard forecasting scores in Section 6. Finally, a short overview of a simple downscaling method using additional ground measurements is presented in Section 7. Conclusions and perspectives are presented in Section 8.

The outputs of the algorithm and the code used to generate the figures are available freely on https://github.com/AAyet/analog-solar-forecasting.

2. Data pre-processing

In this Section, we present the satellite and ground data used in this study. We then introduce the variable of interest, the cloud index, that is forecasted by the analog method.

Nomenclature

G global horizontal irradiance (W m^{-2})
G_{clr} clear sky irradiance (W m^{-2})
c cloud index
t time (one hour resolution)
d day of year
h hour of day
c cloud index
(x, y) latitude and longitude of satellite-image pixels
(x_s, y_s) latitude and longitude of a site satellite-image pixel
DVI daily variability index
DCI daily clearness index
C^m daily temporal correlation map of cloud index
C_n^p spatial correlation between n-th analog and observation
C_n^* optimal spatial correlation for the n-th analog
w_n analog weights
c^o observed cloud index
$\hat{c}$ forecasted cloud index
$\mathbf{A}$ analog design matrix
$\mathbf{S}$ successor feature matrix
$\mathbf{W}$ analog weights matrix
$\mathbf{B}$ linear regression coefficients
r_n linear regression residuals
σ standard deviation of the forecast
$\bar{p}_{(s,d)}(l)$ average probability of the most probable cluster transitions
$\mathbf{X}$ VAR(1) method design matrix
$\boldsymbol{y}$ VAR(1) method endogenous variable
$\mathbf{B}^{\mathrm{ar}}$ VAR(1) method linear regression coefficients

2.1. Satellite and ground data

The Ocean and Sea Ice Satellite Application Facility (OSI SAF), almost in real time, develops, processes and distributes products related to key parameters of the ocean-atmosphere interface, that can be accessed free of charge

(http://www.osi-saf.org). To demonstrate the algorithm, we used an archive of $18,521$ images of satellite-derived GHI obtained from the Meteosat Second Generation geostationary satellites covering Europe and Africa (Brisson et al., 1999; Le Borgne et al., 2004, 2011). The images were remapped on a regular grid of $0.05°$ and interpolated to produce hourly maps. The archive extended from Sept. 6^{th}, 2011 to Dec. 31^{st}, 2016. We used the year 2016 as a test year and the rest of the archive as the training set. Aerosols are the main source of errors in satellite observations. This is corrected by OSI SAF with a deterministic global correction varying in latitude and period of the year, without accounting for Saharan dusts. There is no site adaptation and the correction is not parameterized by the soil albedo, which varies mainly with orography. Moreover, systematic errors occur in the mid-latitudes. Although satellite information has good spatio-temporal sampling, ground measurements are needed to avoid specific atmospheric contamination.

The method was tested at the location of five stations of the Baseline Surface Radiation Network (BSRN, Ohmura et al., 1998) where pyrgeometer measurements are available. The stations are shown in Fig. 1 (a) and are described in detail in Table 1. They cover the wide range of climates required to test the robustness of the method to local variations in cloud dynamics. The climatic difference between the sites is investigated further in Sec. 2.3. The algorithm is tested considering the observed satellite value of GHI to be the reference truth. In Section 7 the ground measurements are also used to present a simple downscaling method as a simple extension. Note that in the following, the Payerne ground data will not be used, due to a lack of historical measurements.

Station	Latitude	Longitude	Altitude (m)	Climate
Palaiseau (France)	48.713	2.208	156	Continental
Carpentras (France)	44.083	5.059	100	Mediterranean
Camborne (England)	50.217	-5.317	88	Oceanic
Payerne (Switzerland)	46.815	6.944	491	Semi-continental
Cener (Spain)	42.816	-1.601	471	Oceanic

Table 1: Selected BSRN stations with corresponding latitude, longitude, altitude and climate.

2.2. Cloud index and clear sky irradiance

Satellite-derived GHI, noted as G (see Fig. 1 (a) for an example), can be decomposed into two contributions: clear sky irradiance G_{clr}, the radiation received by the ground in the absence of clouds, with a deterministic diurnal and seasonal cycle, and cloud cover, a negative contribution which is the main source of unpredictability of GHI. Cloud cover is represented by the cloud index c varying between zero and one such that

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

Clear sky irradiance G_{clr} is usually computed using a clear sky model (e.g. the Ineichen and Perez model (Perez et al., 2002), the Heliosat method (Cano et al., 1986; Hammer et al., 2003), or the Frouin and Chertock (1992) parametrization, used by OSI SAF). It uses a deterministic formula accounting for the diurnal and seasonal variations in the solar zenith angle together with an aerosol climatology. In this paper, for simplicity, we do not use such deterministic formulations for G_{clr}. Instead, for a given day and hour (d, h),

G_{clr} is computed by taking the maximum GHI in the satellite database for data within a 3-month interval $\mathbb{S}(d,h)$ around (d,h) such that

$$G_{clr}(d,h,x,y) = \max_{t \in \mathbb{S}(d,h)} \mathrm{G}(t,x,y). \tag{2}$$

An example of a cloud index map obtained with this clear sky model is presented in Fig. 1 (b).

This simplistic model omits some features of the real atmospheric irradiance, such as cloud enhancement events (Inman et al., 2016), where GHI is enhanced as compared to a cloud-free situation due to the presence of clouds. These events induce an overestimation of the clear sky irradiance in Eq. (2). An evaluation of the quality of the simplistic clear sky model was done by comparing it to the Ineichen and Perez model, implemented with a monthly climatology of Linke turbidity (describing the optical thickness of the atmosphere due to both the absorption by water vapor and the absorption and scattering by aerosol particles relative to a dry and clean atmosphere). The clear sky model (2) was trained over the historical data set of satellite-derived irradiance as defined in the previous subsection (from 2011 to 2015), and it was compared to the prediction of the Ineichen and Perez model over the year 2016 (the test data set). We used the Mean Bias Error and the Root Mean Squared Error, defined in Sec. 6.2. Results are presented in Table 2. The simplistic clear sky model has a negative bias, which is low with respect to the RMSE as a result of the cloud enhancement events.

Such a simplistic model was used since the main goal of the paper is to assess the skill of the analogs to emulate cloud dynamics, i.e. at forecasting the cloud index. The scores presented in Section 6 have been computed

Station	MBE	RMSE
Palaiseau	-72	124
Carpentras	- 78	121
Camborne	- 70	124
Payerne	- 35	98
Cener	- 76	117

Table 2: Mean Bias Error (MBE) and Root Mean Squared Error (RMSE) in W m^{-2} between the estimated clear sky value from Eq. (2) and the Perez and Ineichen model.

in terms of GHI, but similar results have been observed when computed in terms of cloud index, revealing that their interpretation is not biased by the clear sky model errors. This model thus provides a sufficiently good data set of cloud index maps to test the validity of the proposed method.

2.3. Climatic description of the sites

One of the main goals of the present work is to assess the robustness of the analog method to different climatic situations. To further investigate the difference in climate between the different BSRN sites, we used the simplified framework presented in Huang et al. (2014) and originally introduced in Stein et al. (2012).

Two daily indices are computed over the whole data set, the Daily Variability Index (DVI)

$$\mathrm{DVI} = \frac{\sum_{i=2}^{n} |G(t_i, x_s, y_s) - G(t_{i-1}, x_s, y_s)|}{\sum_{i=2}^{n} |G_{clr}(t_i, x_s, y_s) - G_{clr}(t_{i-1}, x_s, y_s)|} \tag{3}$$

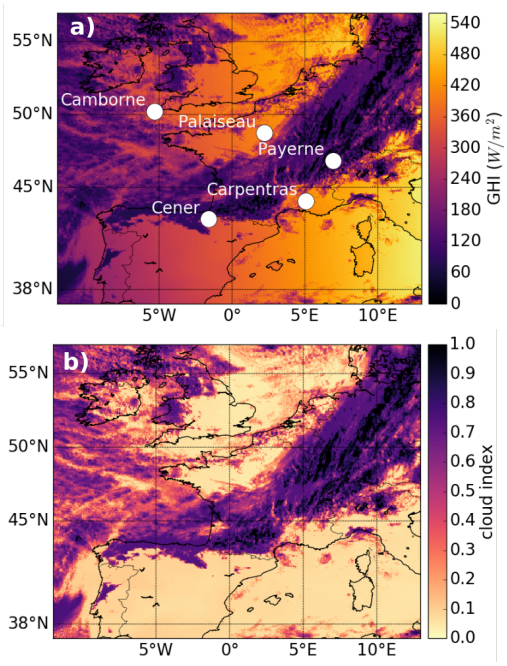


Figure 1: (a) OSI SAF Eumetsat satellite-derived GHI image with selected BSRN stations; (b) cloud index obtained after applying Eq. (1) to (a), on July 2$^{\text{nd}}$, 2016.

and the Daily Clearness Index (DCI)

$$\text{DCI} = \frac{\sum_{i=1}^{n} G(t_i, x_s, y_s)}{\sum_{i=1}^{n} G_{clr}(t_i, x_s, y_s)}, \tag{4}$$

where $(t_i)_{i\in\{1,\dots,n\}}$ is the set of times in a given day (with an hourly temporal resolution). Note that one DVI and one DCI value are computed per day in the data set, as opposed to one per day of one year in the case of clear sky irradiance.

Daily Clearness Index measures the average GHI relative to clear sky for a given day, and is between zero and one, while the DVI measures its variability, and is positive. The combination of both indices allows us to distinguish between three types of days: overcast days, with a DCI close to zero and a

DVI lower than one, intermittent days, with a higher DCI and a DVI higher than one, and clear days, with a DCI and a DVI close to one (see Fig. 2 of Huang et al. (2014) for more details).

The Probability Density Function (PDF) of both indices is then estimated using a Gaussian kernel density estimation method, where the bandwidth is estimated using Scott's rule (Scott, 2015). The resulting PDFs are presented in Fig. 2, using ground measurements or satellite-derived irradiance to compute the indices.

There is a clear difference between the different sites, that is visible both from satellite-derived irradiance (Figs. 2 (a) and (b)) and ground measurements (Figs. 2 (c) and (d)). Palaiseau and Camborne exhibit similar PDFs, indicating a high proportion of variable and overcast days. Carpentras and Cener show an opposite behavior, with a higher proportion of clear days, even-though the DCI of those sites is different, showing that Cener has a higher proportion of overcast days than Carpentras. The Payerne site (for which ground measurements are unavailable), shows in-between characteristics. Overall, the five different sites have a different proportion of overcast, intermittent and clear days, indicating different climates.

3. The analog method

3.1. Overview of the method

In this section we present the analog method as implemented in this work.. To forecast the cloud index over a given site (a pixel of a satellite-derived

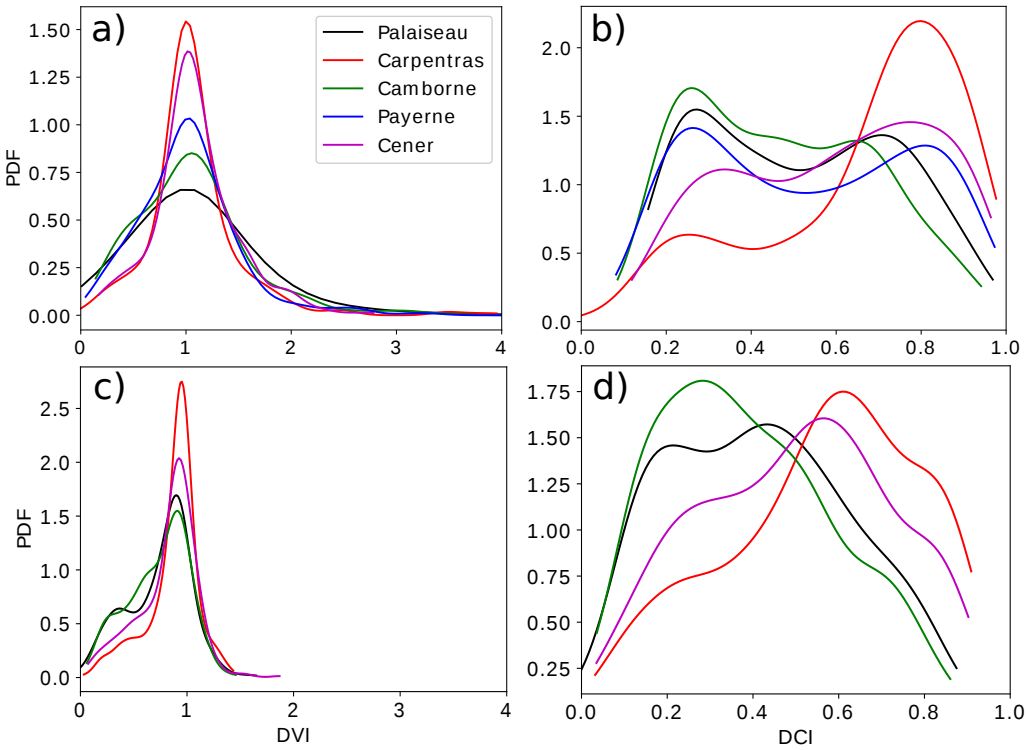


Figure 2: Estimated probability distribution function (PDF) of [(a), (c)] DVI and [(b), (d)] DCI for the selected BSRN sites; (a) and (b) are from one pixel satellite-derived irradiance while (c) and (d) are from ground pyrogeometer measurements.

irradiance map), the method requires a current observation of the cloud structures from which the forecast is made (a cloud index map around the pixel of interest) and a historical or training data set (a set of past cloud index maps around the pixel of interest). The method is divided into three steps.

The first step is the selection of the closest analogs to the current observation, through a k-nearest neighbor algorithm. This implies two sub-steps. Given a site over which the forecast is made, a set of neighboring pixels is first

constructed. These pixels are those with maximal information explaining the variability in the cloud index of the considered site. They define a correlation mask around the considered site, the size and shape of which depend on the day of the year under consideration (due to the seasonal variation of the cloud structures). To forecast the cloud index over a site for a given day of a year, only the pixels within the mask are considered, both for observation and for the training data set.

The resulting reduced images are then compressed into four features, that represent physical attributes of the cloud structures. The k-nearest neighbors (the analogs) to the observation are then retrieved: in the four-dimensional feature space, they are the closest images to the observation in terms of Euclidean distance. The compression into four features aims at both avoiding the "curse of dimensionality" due to the high number of pixels contained in the masks, and at including physical reasoning in the algorithm. The choice of the features can seem arbitrary, and is thus further investigated in Section 4.

The second step is the generation of a prediction ensemble from the selected analogs. It consists of a set of possible forecasts, weighted according to their reliability. Since the aim of the method is to forecast the cloud index over a precise pixel, the analogs are first spatially translated to match the observation as closely as possible. The resulting translated images are then weighted as a function of their correlation with the observation: a higher correlation is interpreted as a more reliable analog and thus has a higher weight. The ensemble of predictions is thus the ensemble of successors (the images that were observed l hours after the analogs, l being the lead time of

the forecast), translated and weighted in the same way as the corresponding analogs.

The last step is to aggregate the ensemble in order to estimate a PDF of the forecast. Under the assumption that the PDF is Gaussian, the estimation is made through a weighted linear regression between the analogs and the successors. The aggregation method, called local linear regression in the context of analog methods, is known for its robustness to small data sets and for handling non-linearities.

The method has one critical parameter, which is the number k of analogs selected in the first step. The determination of the optimal number of analogs is presented in Sec. 6.1. In the rest of the section, each of the steps of the method is explained in more detail.

3.2. 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 sought. In the case of precipitation forecasts (Atencia and Zawadzki, 2015; Panziera et al., 2011), a rectangular window is selected manually according to the scale of the structures to be forecasted and the bottom boundary forcing (mainly orographic). It is, however, essential for the method to have an automatic zone selection. In Dambreville et al. (2014), inter-correlation maps between the pixel of interest and the surrounding region are computed. It is shown that using them to select the pixels to be used in a spatio-temporal auto-regressive model improved the forecasting skills as compared to a simple squared zone selection. Here, we

proceed similarly, computing a daily temporal correlation map $\mathcal{C}^m$ (for a day d) between the pixel of interest (x_s, y_s) and the surrounding region using a metric inspired by Zawadzki (1973)

$$\mathcal{C}^m(d,x,y) = \frac{\overline{c(t,x_s,y_s)c(t,x,y)}^d}{\left[\overline{c(t,x_s,y_s)^2}^d \overline{c(t,x,y)^2}^d\right]^{1/2}}, \tag{5}$$

where the averages $\overline{\cdot}^d$ are temporal within a 3-month interval around d. An example of such a map is shown in Fig. 3 (a).

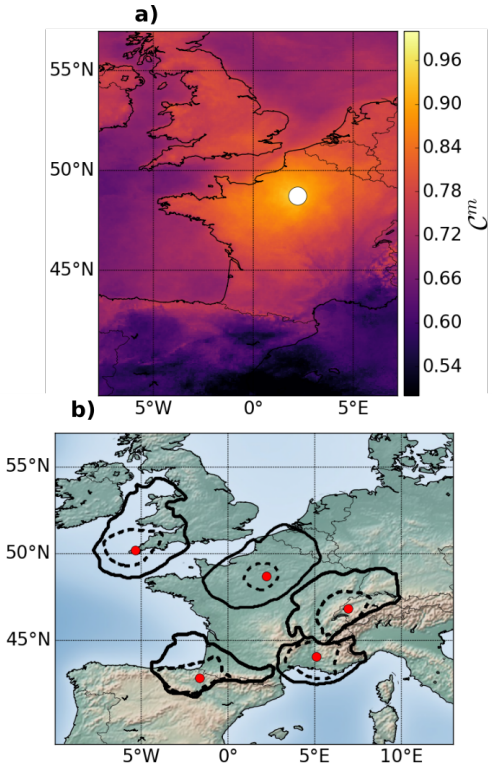


Figure 3: (a) Example of a correlation map for the Palaiseau site (white dot); (b) correlation masks for the different BSRN sites (red dots) for Jan. 1st (full line) and Jul. 1st (dashed lines).

The $\mathcal{C}^m$ metric is more suited to cloud forecasting than a standard Pearson

correlation (where the mean cloud index is subtracted in each of the factors), since it accounts for only the cloud structures (i.e. pixels where the cloud index is non zero) in the correlation. It gives a measurement of the mean geographical extension of the structures around the site pixel. For a given map, the region where the correlation is higher than 0.9 is selected. However, this can lead to unrealistic masks, where geographically distant regions can be artificially selected together. A segmentation algorithm (based on the watershed algorithm, e.g. Soille and Ansoult, 1990) is applied to select a connected component containing the site. Overall, we also set a minimal number of pixels that must be contained in the mask. Fig. 3 (b) presents the masks for the different BSRN stations, for two different days of the year. The summer masks tend to be smaller than the winter masks, reflecting a change in synoptic regimes. Note also the link between the mask contour and the orography for the Cener, Carpentras and Payerne sites.

3.3. Analog selection

Selecting correct analogs means identifying past atmospheric situations where cloud structures are similar to the current observation, and the evolution of which in the past (the successors) is representative of the future evolution of the hindcast. The aim is thus to find similar cloud regimes, i.e. cloud structures that evolve in a similar way. Since the identification of similar cloud regimes does not rely on the details of the cloud structure at a given time, the analog selection is performed considering images compressed in a four dimensional space (feature space). This also has the advantage of avoiding overfitting and being computationally efficient. We thus first describe

how the features are defined, and then give details on the k-nearest neighbor algorithm used to select the analogs.

3.3.1. Feature extraction

For a given cloud index image in a given daily mask, its cloud index histogram contains crucial information on the structure of the clouds. It is often bimodal, with its lower peak corresponding to "clear sky" pixels, and its higher peak to "cloud" pixels. Here, we make the assumption that histograms discriminate between different cloud regimes. This assumption requires that the dynamics of the clouds observed in the correlation mask are constrained enough, hence the importance of having a correct mask that selects only meaningful information. The validity of the assumption also depends on how constrained the weather for a given site is, for instance by orographic forcing and/or predictable synoptic weather patterns.

The four features described below are thus meant to differentiate images with different cloud histograms. The first step in the definition of the features for a given cloud index image is to separate the clear sky from the cloud pixels by finding a cloud index threshold separating the two modes (e.g. Fig. 4). This is done by using Otsu's method (Otsu 1979, similar to a bimodal Fisher's discriminant analysis of a histogram of cloud index). The image is then compressed into four features the values of which range from 0 to 1:

1. the cloud fraction, or number of cloud pixels over the total number of pixels in the mask
2. the cloud spread, or number of cloud pixels over the number of pixels in

the convex hull of the clouds (one when only one cloud and zero when many separate clouds)

3. the clear sky intensity, or mean cloud index of the clear sky pixels
4. the cloud intensity, or mean cloud index of the cloud pixels.

Note that a principal component analysis has been performed on the historical data set, resulting in the cloud fraction feature being correlated with the first principal component. This consolidates a fact already mentioned in Panziera et al. (2011).

3.4. k-nearest neighbor algorithm

Given an observation for which analog situations are found, the full historical data set of potential analogs is first reduced by imposing a temporal constraint. Following the method used by Atencia and Zawadzki (2015), only analogs within a time of the year (3-month window) and time of the day ($\pm$3h) interval are considered. A minimal 24h lag between two chosen analogs is also imposed. This increases the likelihood of finding similar convective and advective patterns. It also increases the robustness of the method to different geographical locations. Then, the k-nearest neighbors are selected, based on a Euclidean distance in the four dimensional feature space, in which the images are mapped. As described in Section 6, for the BSRN sites of this study, the optimal number of neighbors is close to 80.

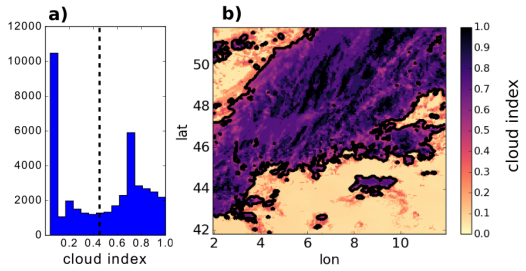


Figure 4: (a) Histogram of cloud index from image (b) on the 2nd of July 2016 at Payerne. The cloud threshold in full black line in (b) corresponds to the dashed line in (a).

3.5. Ensemble generation

Given a set of analog-successor pairs corresponding to an observed cloud index image, the next step is to create a forecast ensemble, i.e. a set of possible outcomes, weighted by their reliability.

The analogs have been found on the basis of features that do not depend on the details of the observed cloud structures. To improve the accuracy of the single-pixel forecast, the first step is to match the cloud structures of the analog cloud irradiance maps with those observed. The k analogs are thus spatially translated to match the observation. This translation is performed by maximizing a correlation metric between the analogs and the current observation. For a translation $(\Delta x, \Delta y)$ of the analog image, the spatial correlation $\mathcal{C}_n^p$ between the n-th translated analog and the observation is

$$\mathcal{C}_n^p(\Delta x, \Delta y) = \frac{< c^o(x,y) c_n^a(x+\Delta x, y+\Delta y) >}{[< c^o(x,y)^2 >< c_n^a(x+\Delta x, y+\Delta y)^2 >]^{1/2}}, \tag{6}$$

where c^o and $\{c_n^a\}_{n\in\{1,..,k\}}$ are respectively the observed and analog cloud index images (in the mask) and $< \cdot >$ is a spatial average over the mask. We use an optimization procedure to find the translation that maximizes

$\mathcal{C}_n^p$ defined in Eq. 6. The optimal translation is then used to compute the maximum correlation noted as $\mathcal{C}_n^*$. Note that the procedure does not allow information that was initially outside the mask to be translated to the site pixel (x_s, y_s).

Since the translation is performed within the correlation mask, we can assume that the translated cloud structures will evolve similarly to the original structures. The successors are thus translated with the same optimal displacement as their corresponding analogs. The forecast ensemble is then created by considering the translated successors as a set of possible outcomes, their reliability being measured by $\mathcal{C}_n^*$. The more a translated analog is correlated with the truth, the higher the weight of the analog-successor pair in the forecast ensemble. A simple exponential kernel (e.g. Lguensat et al. 2017) is chosen for the weights

$$w_n \propto \exp\left(\frac{\mathcal{C}_n^*}{\lambda}\right), \tag{7}$$

where λ is the median of the $\mathcal{C}_n^*$. In Fig. 5, we compute the mean over the test year of the correlation given in Eq. (6) between the observed field and the successors at different lead times of: the best analog after the k-nearest neighbor selection, after translation, and after reordering the analogs with the weights defined in Eq. (7). It shows that each of the steps increases the correlation of the analogs with the truth.

3.6. Aggregation of the ensemble

The previous step provides a weighted ensemble of forecasts with weights w_n. The ensemble is then aggregated to estimate the PDF of the forecast.

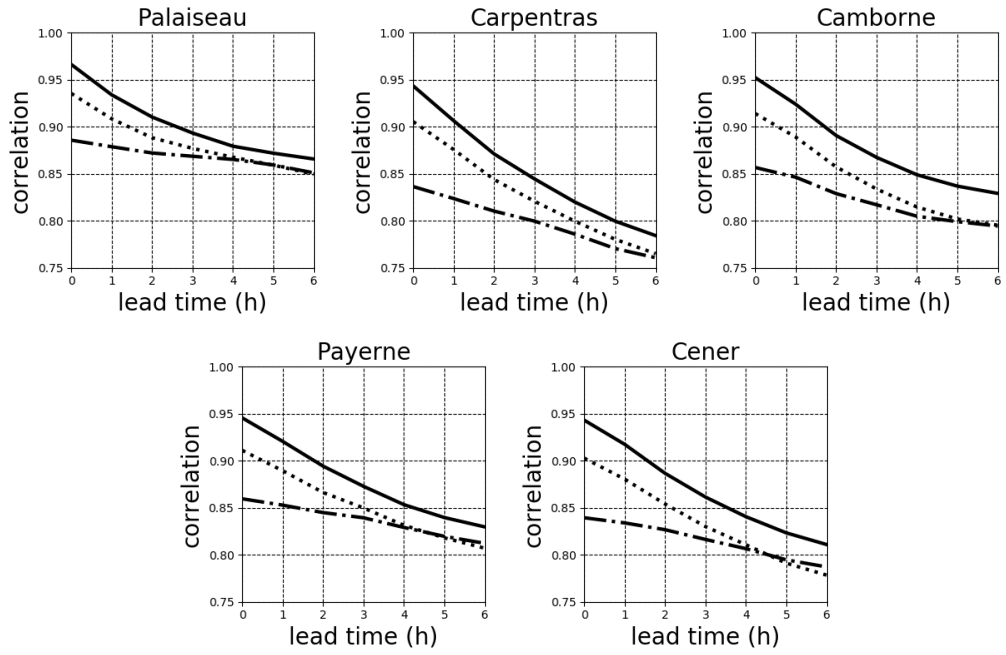


Figure 5: Mean correlation over the test data set between current observations and best analog-successor for different lead times. Three strategies are studied: the best analog without translation (dotted-dashed), the best analog with translation (dotted) and the best translated analog with respect to the weights given in Eq. (7) (full line).

We use a local linear aggregation operator (Cleveland, 1979) to estimate the mean and the standard deviation of the forecasted PDF, assuming that it is Gaussian. This operator makes an efficient use of small data sets and reduces biases. Fig. 6 presents a schematic of the forecasting operation:

1. a weighted least square regression $\mathbf{B}$ is fitted between the analogs c_n^a and the successors c_n^s. The dimension of the analogs is first reduced by using a Principal Component Analysis (PCA) on the set of analog images, as prescribed by Lguensat et al. (2017), to avoid overfitting. Principal Component Analysis finds a set of orthogonal vectors (set

to five in the present study) that explains most of the variance in the considered images, on which the data is projected.

2. the regression operator is then applied to the current observation c^o (compressed in the same space as the analogs) to obtain the mean of the estimated PDF $\hat{c}$

$$\hat{c} = \mathbf{B}c^o \tag{8}$$

3. the standard deviation of the PDF is given considering the weighted standard deviation of the residuals $r_n = c_n^s - \mathbf{B}c_n^a$

$$\sigma^2 = \frac{1}{1 - \sum_{n=1}^{k} w_n^2} \sum_{n=1}^{k} w_n \left(r_n - \overline{r}\right)^2 \tag{9}$$

where $\overline{r}$ is the weighted mean of the residuals.

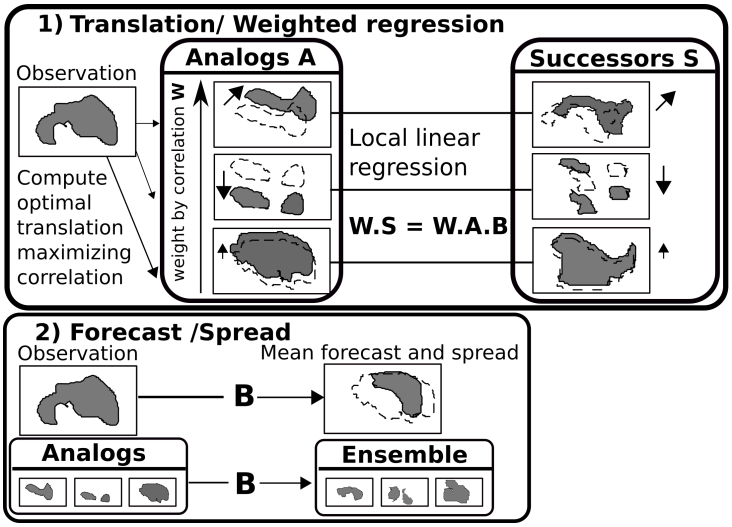


Figure 6: Schematic of the aggregation procedure: probabilistic forecast using a local linear regression $\mathbf{B}$ on selected analogs and successors.

For comparison, we have also used (not shown) a locally constant operator, a special case of the local linear operator where the weighted mean and standard deviation of the successors is considered, without any linear regression. This simple procedure is described in detail in Lguensat et al. (2017). The local linear operator has proven to be the most efficient, based on the performance criteria described in Section 6.

4. Evaluation of the quality of the analogs

The choice of the features and distance used to select the analogs is crucial for the performances of the algorithm. In this work, the constraint of using only one source of data with sufficient history in many geographical locations (maps of satellite-derived irradiance), has led to the heuristic choice of features presented in Section 3. An automated selection of the features could also have been performed, using for instance principal component analysis on maps of satellite-derived irradiance as in Davò et al. (2016) or Foresti et al. (2015). In the present method, a physics-based choice of features has been preferred over a data-based choice.

Using a time-lagged embedded state vector (i.e. considering images at previous time steps in addition to the current observation) to determine cloud regimes is another classic choice. Embedded vectors have been recently used in the frame of the Nonlinear Laplacian Spectral Analysis (NLSA) to forecast the Madden Julian Oscillation index by means of analogs (Alexander et al., 2017). However, since GHI images are available only for daylight hours, constructing embedded states is not optimal. In addition, the low temporal

resolution of the images decreases the precision of any method estimating a temporal change in cloud structures. Still, the algorithm has been tested considering an estimation on the mean displacement of clouds from two successive images as additional features. The average score of the algorithm decreased on different sites with respect to the algorithm presented above (not shown). These additional features made the method less robust to different sites, since the estimation of cloud motion is also difficult due to cloud deformation, especially in the presence of a strong orographic forcing.

As mentioned in the introduction, it is difficult to assess the quality of a given set of features and distance used to select the analogs. A simple test, suggested in Van den Dool (1994), considers analogs as valid if they are closer to the observation than a climatology (in the four dimensional features space). This was done (not shown), with an hourly climatology, yielding positive results. Another quality test is proposed herein, by providing some insight into the physical meaning of the features. It aims at quantifying if temporal patterns of cloud structures can be easily discriminated in the four-dimensional feature space. If the features and the distance are well chosen, initial images of similar temporal patterns should be grouped in well defined clusters in feature space (with respect to the chosen distance). Cloud regimes (similar-evolving cloud index fields) can then be identified in relation to the clusters.

The methodology used to define clusters is similar to that used in Wang et al. (2017) in the context of the Pattern Sequence-based Forecasting method. We first define a finite set of clusters using a k-mean algorithm (Lloyd 1982) on the historical data set of features for a given site. It is defined for each day

of the year, which corresponds to a different mask, and thus different values of the features. The optimal number of clusters is obtained by maximizing the silhouette score (Rousseeuw, 1987). It is defined as the average of $(b - a)/\max(a, b)$ over all the points in the feature space, with a the mean distance between a point and all other points in the same cluster, and b the mean distance between a point and all other points in the next nearest cluster. The score is bounded between -1 for incorrect clustering and +1 for highly dense clustering. Scores around zero indicate overlapping clusters.

An example of clustering for the Palaiseau site is given in Fig. 7, for Jan. 1^{st}. The optimal number of clusters for this particular day is five. We can interpret the different clusters by their cloud structures: red for small and isolated clouds, cyan for large and dense clouds, green and blue for intermediary situations, yellow when the whole mask is covered by dense clouds. Table 3 presents the mean (over all the days of the year) optimal number of clusters and the associated mean silhouette score for the different sites. Palaiseau and Camborne display similar characteristics, with a low mean silhouette score and a high number of clusters.

The remaining question is to link the clusters to cloud regimes. More precisely, we look at how the states in a given cluster evolve in time looking at the transitions between the different clusters. For a given site s and day d to which corresponds an optimal set of clusters, we compute the transition matrix between the clusters, i.e. the set of probabilities of an image being in a cluster j knowing that it was in a cluster i, l hours before, denoted

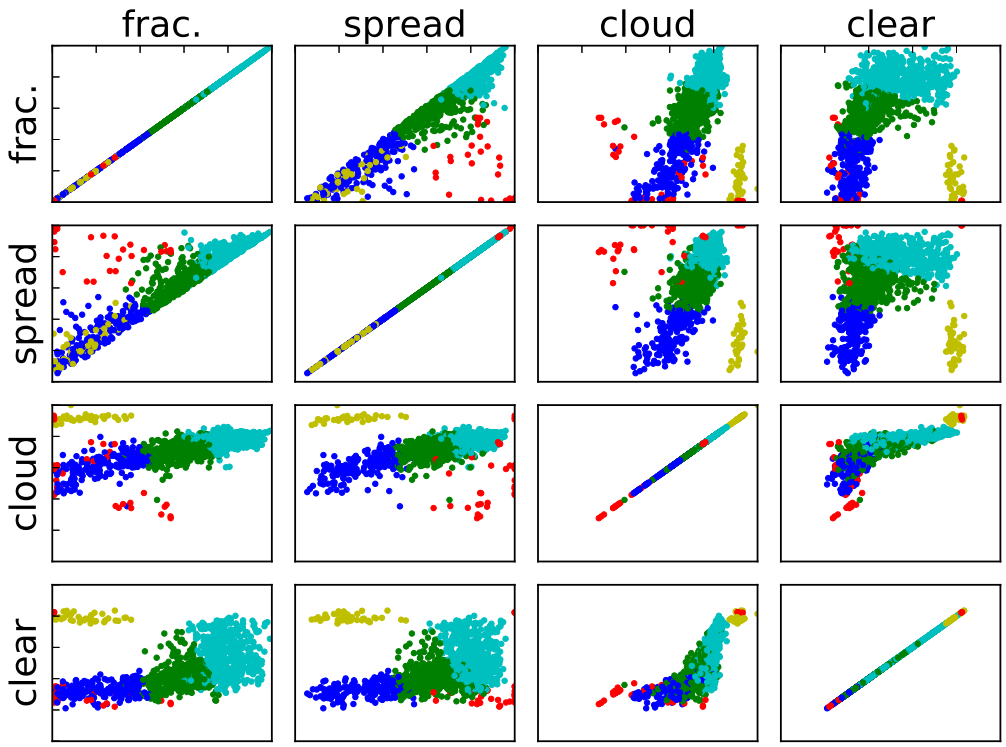


Figure 7: Scatter plots of the different features, with colors corresponding to five clusters obtained with a k-mean algorithm, for the Palaiseau site, on Jan. 1st. All the axes range between zero and one.

$p_{(s,d)}(j|i,l)$. We then obtain the mode of this distribution by

$$p^*_{(s,d)}(i,l) = \max_j p_{(s,d)}(j|i,l) \tag{10}$$

and compute the weighted average of the modes defined in Eq. (10) over all the clusters and for a given lead time l as

$$\bar{p}_{(s,d)}(l) = \sum_i p_{(s,d)}(i) p^*_{(s,d)}(i,l). \tag{11}$$

where $p_{(s,d)}(i)$ is the probability of being in cluster i. Table 4 presents the averages of $\bar{p}_{(s,d)}(l)$ over all the days of the year d, along with its standard

Station	Number of clusters	Score
Palaiseau	5.0 (0.9)	0.38 (0.02)
Carpentras	3.5 (0.5)	0.46 (0.02)
Camborne	4.2 (1.0)	0.37 (0.03)
Payerne	3.8 (1.1)	0.44 (0.03)
Cener	3.1 (0.4)	0.41 (0.03)

Table 3: Mean optimal number of clusters and mean silhouette score for the BSRN sites. The standard deviation is between brackets.

deviations, for the different sites. We see again that the Palaiseau and Camborne sites have the worse scores, meaning that there is no clear most probable transition between clusters. On the contrary, the mean probabilities for the Cener site are higher with a low confidence interval. Thus, we can expect better performances for this site.

Station	+1h	+2h	+3h	+4h	+5h	+6h
Palaiseau	0.78 (0.07)	0.67 (0.08)	0.59 (0.08)	0.55 (0.07)	0.52 (0.08)	0.48 (0.16)
Carpentras	0.82 (0.04)	0.74 (0.06)	0.69 (0.07)	0.65 (0.08)	0.55 (0.13)	0.58 (0.16)
Camborne	0.81 (0.05)	0.70 (0.07)	0.63 (0.08)	0.58 (0.08)	0.55 (0.13)	0.49 (0.17)
Payerne	0.85 (0.04)	0.77 (0.06)	0.71 (0.07)	0.66 (0.08)	0.62 (0.09)	0.57 (0.16)
Cener	0.84 (0.03)	0.76 (0.04)	0.70 (0.05)	0.66 (0.05)	0.63 (0.05)	0.61 (0.06)

Table 4: Mean probabilities and associated standard deviation (between brackets) of states that follow the most probable path between clusters in the training data set.

5. Reference methods

The analog method is compared to three reference statistical methods, presented in this section.

5.1. Persistence

The first method is a clear-sky adjusted persistence. The last observed cloud index is used as a forecast. It is converted back to GHI using the clear sky irradiance predicted at the time of the forecast. This method provides a deterministic forecast and is used as the reference method to evaluate deterministic forecasts.

5.2. Persistence ensemble (PeEn)

The second method is an extension of the persistence method which is commonly used (e.g. Alessandrini et al., 2015) to generate probabilistic forecasts: the persistence ensemble (PeEn) method. The last 20 observations for each site from the past 20 days at the same diurnal hour are considered as an ensemble of forecasts. A cumulative distribution function of the forecast $\hat{c}$ can then be estimated from the ensemble c_i^e

$$P(\hat{c} < c) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(c_i^e < c) \tag{12}$$

with $n = 20$ the size of the ensemble and $\mathbf{1}$ the indicator function.

5.3. Adaptive vector-autoregressive model (VAR(1))

The third method is an "adaptive" order-one vector auto-regressive model, denoted as VAR(1). Non adaptive VAR models (Box et al., 1970) are commonly used in the literature for GHI forecasting (e.g. Dambreville et al. 2014; Alessandrini et al. 2015). The version presented here is more sophisticated and is a special case of the analog method, in which the local linear regression is performed considering the whole reduced historical database (without analog selection or weighting).

The VAR method predicts a given time series by means of a linear regression. For a given day d, the exogenous variables are the values of GHI in the pixels contained in the day d mask. To build the design matrix $\mathbf{X}(d)$, only the images in the training data set within a $\pm$ 45 day window around d are considered

$$\mathbf{X}_{i,j}(d) = c(t_i, x_j, y_j) \tag{13}$$

with (x_j, y_j) the set of coordinates in the day d mask, and t_i the set of times in the $\pm$ 45 day window.

The endogenous variable y is the cloud index value observed at the pixel containing the site of interest (x_s, y_s), l hours after the initial observation

$$\boldsymbol{y}_i(d, l) = c(t_i + l, x_s, y_s) \tag{14}$$

The model is an ensemble of $2,190$ linear regressions $\mathbf{B}^{\mathrm{ar}}(d, l)$: one per day and per lead time, such that

$$\|\boldsymbol{y}(d, l) - \mathbf{X}(d)\mathbf{B}^{\mathrm{ar}}(d, l)\| \tag{15}$$

is minimized, where $\|\cdot\|$ is the Frobenious norm. A "global" Auto-Regressive model has also been tested, fitting a unique model for each location and for each lead time, totaling 6 linear regressions per site (as in the literature). Results are worse than for the adaptive VAR(1) model, especially at low lead times. These are not shown here for clarity.

The VAR(1) model also provides a probabilistic forecast, by computing a variance in the forecasted index using the residuals, as presented in Eq. (9) for the local linear operator. This method will thus be used as a reference for the evaluation of probabilistic forecasts in the following.

6. Numerical experiment

In this section, we evaluate the performance of the analog nowcasting method using different classic scores, and comparing it to statistical methods. We also discuss the optimal number of analogs for each site. The method is also improved by a simple bias correction of the forecast.

6.1. Experiment set-up

As mentioned in Section 2.1, a data set of $18,521$ hourly images of satellite-derived GHI (from Sept. 6^{th}, 2011 to Dec. 31^{st}, 2016) is split into two parts. The training set extends from Sept. 6^{th}, 2011 to Dec. 31^{st} 2015. The year 2016 is the test set, over which the performances of the algorithms presented below are evaluated.

A key parameter in the analog algorithm is the number of analogs k. Its optimization is done by cross-validation on the training data set. The Root

Mean Squared Error (presented below) is then computed as a function of the number of analogs. Results (not shown) indicate an optimal value close to $k = 80$ analogs for the five sites. The performance of the algorithm is robust to slight changes in RMSE, and one can safely choose the number of analogs to be 80 for different sites. Note that also in other papers on analog methods (Panziera et al., 2011; Atencia and Zawadzki, 2015; Alessandrini et al., 2015), the number of analogs is empirically chosen depending on the situation of interest.

6.2. Reference scores

Hereinafter, $\hat{c}$ corresponds to the statistical forecast (for instance, from the analog method) and c^o is the observed cloud index from a satellite image at the location of the considered site, considered to be the reference truth. We evaluate the performance of the deterministic forecast of our methodology over a set of validation observations $\mathcal{S}$ of cardinality $|\mathcal{S}|$, using the Mean Biased Error (MBE)

$$\text{MBE} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \hat{c}_i - c_i^o \tag{16}$$

the Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} (c_i^o - \hat{c}_i)^2} \tag{17}$$

and the Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} |c_i^o - \hat{c}_i|. \tag{18}$$

We also compute the forecast skill (FS) in terms of RMSE, to measure the improvement of the forecast model with respect to the persistence model

$$\mathrm{FS} = 1 - \frac{\mathrm{RMSE}}{\mathrm{RMSE}_p}. \tag{19}$$

The RMSE is more severe in the evaluation of the performances of the algorithm since it is a quadratic score, but, as advocated by Alessandrini et al. (2015), we also present the MAE here, since penalties paid by the solar energy producers are usually proportional to the imbalances in their production.

The evaluation of the probabilistic forecast is carried out by computing the Brier Skill Score (BSS), defined from the Brier Score (BS) as the improvement over the VAR(1) model

$$\mathrm{BSS} = 1 - \frac{\mathrm{BS}}{\mathrm{BS}_{\mathrm{VAR}}}. \tag{20}$$

The BS is the equivalent of mean square error for probabilistic forecasts and is defined as

$$\mathrm{BS} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{T}} (\hat{p}_{i,j} - T_{i,j})^2 \tag{21}$$

where $\mathcal{T}$ is an ensemble of possible categories in which the observation can fall, in the present case intervals of 10 $W.m^{-2}$ starting from 0 up to 1000 $W.m^{-2}$. We define $T_{i,j} = 1$ if the observation c_i^0 falls within the interval j and $T_{i,j} = 0$ otherwise. The predicted probability for the j-th interval $\hat{p}_{i,j}$ is computed from the cumulative distribution function (CDF) predicted by the model $P(\hat{c}_i \leq c)$.

The Continuous Ranked Probability Score (CRPS) (Hersbach, 2000), an equivalent of MAE for a probabilistic forecast is also computed. It is given

by

$$\mathrm{CRPS} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \int [P(\hat{c}_i \leq c) - P(c_i^o \leq c)]^2 dc \tag{22}$$

where $P(\hat{c}_i^o \leq c)$ is the CDF of the observation, considered here to be a step function. We recall here that both the analog and the VAR(1) methods produce Gaussian forecasts, so that the computation of the CDF is straightforward.

All together, these five scores provide a standard evaluation of the various aspects of the proposed method. The comparison with standard methods is given in Section 6.3.

6.3. Comparison with reference methods and bias correction

Our main concern in this section is to assess the robustness of the methods to different geographical locations, i.e. their capacity to forecast the cloud index for different climatic situations.

The first evaluation score is the MBE, presented in Fig. 8 as a function of lead time. The analog method shows a bias peaking for lead times of 3 to 4 hours. The bias is not present in the VAR(1) method. Fig. 9 presents a scatter plot of the forecasted value versus the observation for the analog methods. It shows that a linear relation between the bias and the forecasted intensity can be clearly inferred.

This bias is thus corrected by applying a simple post-processing method to the forecast. We fit 6 linear regressions per site (one per lead time) between the observed bias and the forecasted value from the analog prediction.

These linear regressions are then applied to the forecasted GHI value to obtain a "post-processed" analog method (called p-analog in the following). Fig. 8 shows that the p-analog method is globally less biased than the analog method, except for the Payerne and Cener sites. This is a very simplistic post-processing method, and more sophisticated model output statistics could be used, such as Kalman filtering (Diagne et al. 2014).

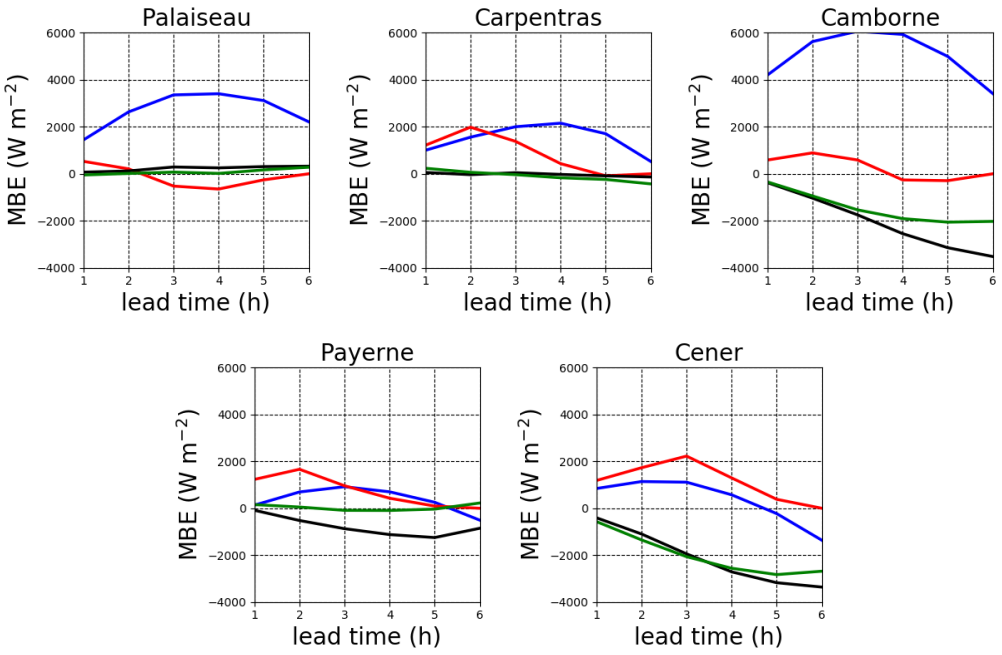


Figure 8: MBE as a function of lead time for the analog method (blue), the post-processed analog method (red), the persistence method (black), and the adaptive VAR(1) model (green).

Figs. 10 and 11 show respectively the MAE and the RMSE for the different methods. A quantitative evaluation of the methods in terms of RMSE skill score is presented in Table 5. Table 6 and Fig. 12 present respectively the BSS and the CRPS as a function of lead time for the different methods. We

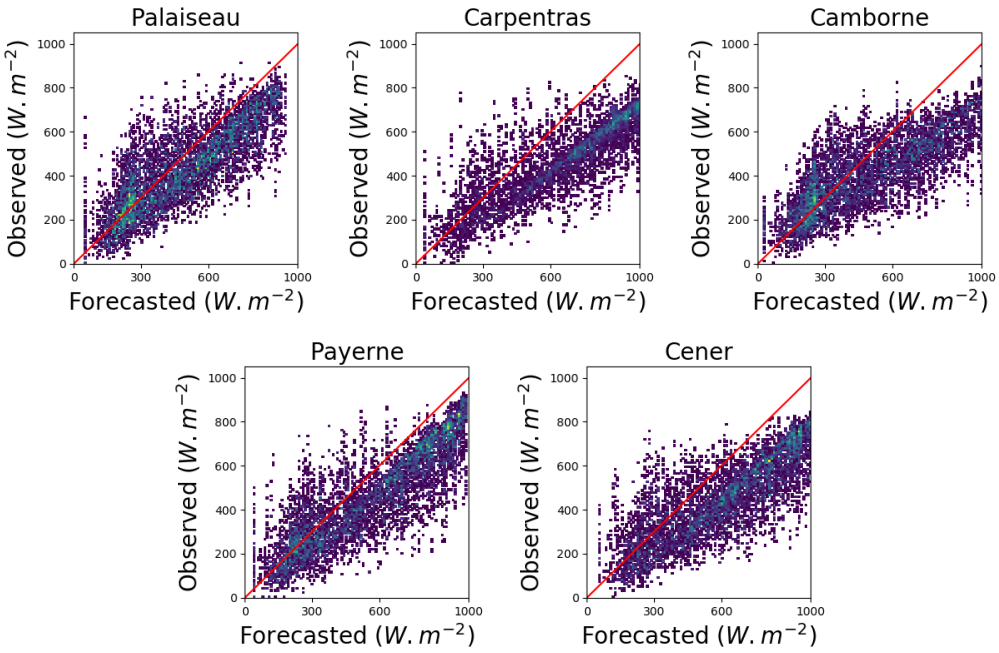


Figure 9: Scatter plot between the observed and predicted GHI values. The red line corresponds to the centered and indentity slope line.

can first note that even-though the p-analog method is more biased than the analog method for certain sites, both its probabilistic and deterministic performances are similar or better than the analog method. Note also the consistency of the results with the climatic analysis of the different sites in Fig. 2: on sites with a higher proportion of intermittent days (Palaiseau and Camborne) the analog and p-analog methods have a worse score as compared to the persistence method, since those days are more difficult to forecast.

When comparing the analog method between sites, results are coherent with the probabilities given in Table 4: a small error corresponds to a large

percentage of images following the most probable transition paths in the database, i.e. where the selected analogs exhibit the same cloud regimes as the truth. Additional evidence of this statement is found by looking at the improvement in RMSE and MAE of the analog method with respect to the VAR(1) model only, which shows the improvement of the forecasting skill due to the analog selection step. The largest improvement also occurs in the sites where the probability defined in Table 4 is high. Comparing the performances of the analog method with the p-analog method, we see that the bias correction dramatically increases the performances on the sites where the analog method does not beat the VAR(1) method. We can thus infer that the inappropriate choice of analogs on the sites where cloud regimes are not well defined is the source of the bias of the method.

Bootstrap confidence intervals have also been computed for each of the scores. Overall, they indicate statistically significant difference scores between the persistence (or PeEn) method and the three other methods. However, the distinction between the analog and p-analog forecast skill is more unclear, and has been further investigated using the Diebold-Mariano (DM) test, introduced in econometry by Diebold and Mariano (1995) and then used in meteorology by Gilleland and Roux (2015). The DM test evaluates whether the difference between two forecasts, quantified by a given forecast score, is statistically significant. The test (not shown) was performed based on the mean squared error score. Results indicate that most of the time the difference between the analog and p-analog methods is significant.

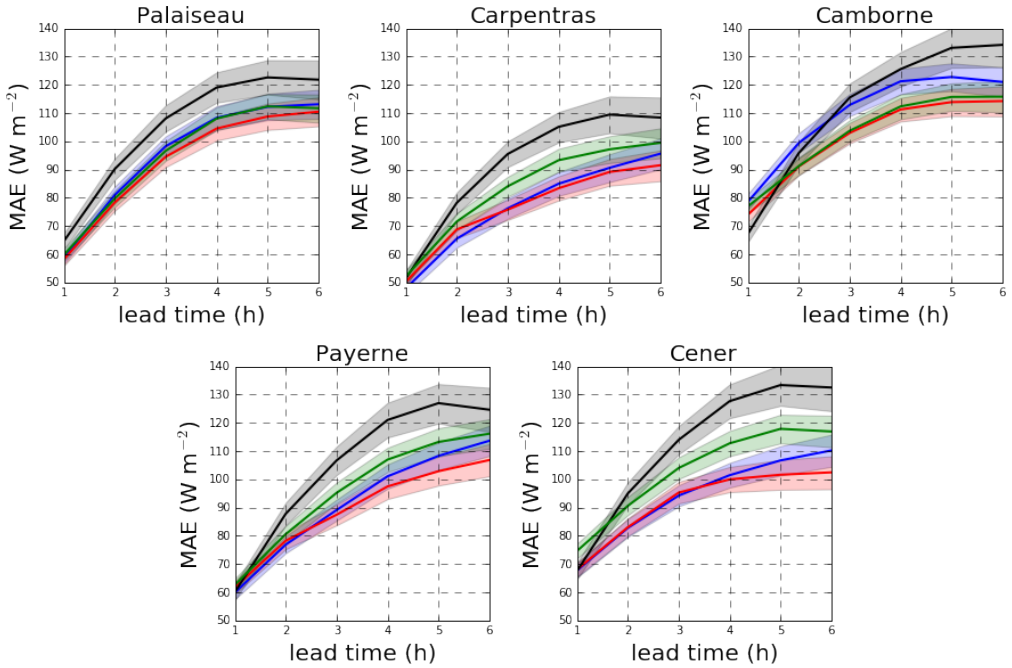


Figure 10: Mean MAE and corresponding 95% bootstrap confidence interval as a function of lead time for the analog method (blue), the post-processed analog method (red), the persistence method (black), and the adaptive VAR(1) model (green).

7. Statistical downscaling

In the previous Section, GHI has been forecasted using only geostationary satellite images. Here, we extend this procedure by showing a simple example of statistical downscaling between satellite and ground data on different BSRN sites. This application shows the potential of the method for operational forecasting.

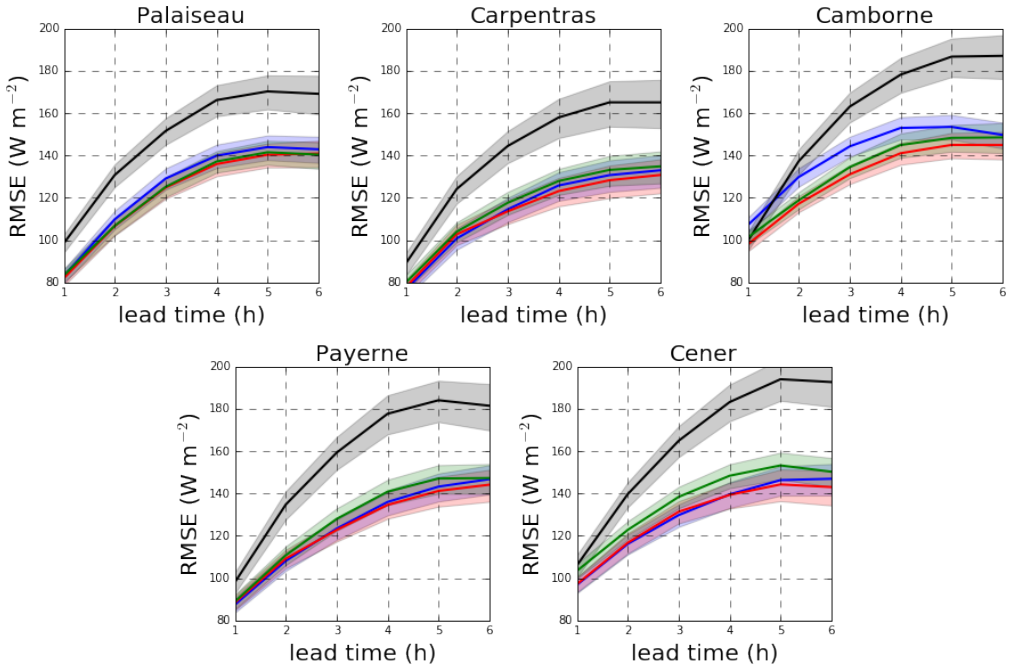


Figure 11: Mean RMSE and corresponding 95% bootstrap confidence interval as a function of lead time for the analog method (blue), the post-processed analog method (red), the persistence method (black), and the adaptive VAR(1) model (green).

7.1. Description of the method

The aim of the present extension is to forecast the cloud index over a local site corresponding to the location of a solar energy source, where a historical database of ground GHI measurements is available.

The principle of the method, first introduced in Zorita and Von Storch (1999), is to use cloud index maps as a measure of the weather regime in which the solar energy source is. Each historical cloud index image is paired with its concurrent (at the same date) ground measurement of GHI. Given an observed cloud index image from which the forecast is made, an estimation

Station	Method	+1h	+2h	+3h	+4h	+5h	+6h
Palaiseau	VAR(1)	0.154	0.183	0.174	0.174	0.169	0.17
	analog	0.157	0.16	0.148	0.158	0.154	0.155
	p-analog	0.168	0.184	0.176	0.182	0.176	0.167
Carpentras	VAR(1)	0.102	0.162	0.186	0.19	0.193	0.183
	analog	0.139	0.188	0.207	0.203	0.208	0.194
	p-analog	0.125	0.172	0.214	0.22	0.223	0.208
Camborne	VAR(1)	-0.016	0.134	0.175	0.186	0.205	0.206
	analog	-0.074	0.054	0.118	0.143	0.179	0.202
	p-analog	0.016	0.145	0.198	0.209	0.225	0.229
Payerne	VAR(1)	0.093	0.178	0.196	0.207	0.201	0.188
	analog	0.103	0.193	0.221	0.232	0.224	0.193
	p-analog	0.093	0.185	0.225	0.239	0.234	0.205
Cener	VAR(1)	0.024	0.121	0.161	0.19	0.21	0.22
	analog	0.087	0.168	0.213	0.238	0.246	0.237
	p-analog	0.084	0.164	0.203	0.239	0.256	0.257

Table 5: Skill RMSE relative to the persistence method.

of the weather regime is performed by selecting analogs, and weighting them as a function of their reliability. The concurrent ground measurements are then used as a forecast ensemble, with the same weights as the corresponding satellite-derived image.

A similar methodology has been developed in Alessandrini et al. (2015) under the name of the analog ensemble method. In this article, the weather regime is estimated using NWP forecasts: past forecasts, analog to the cur-

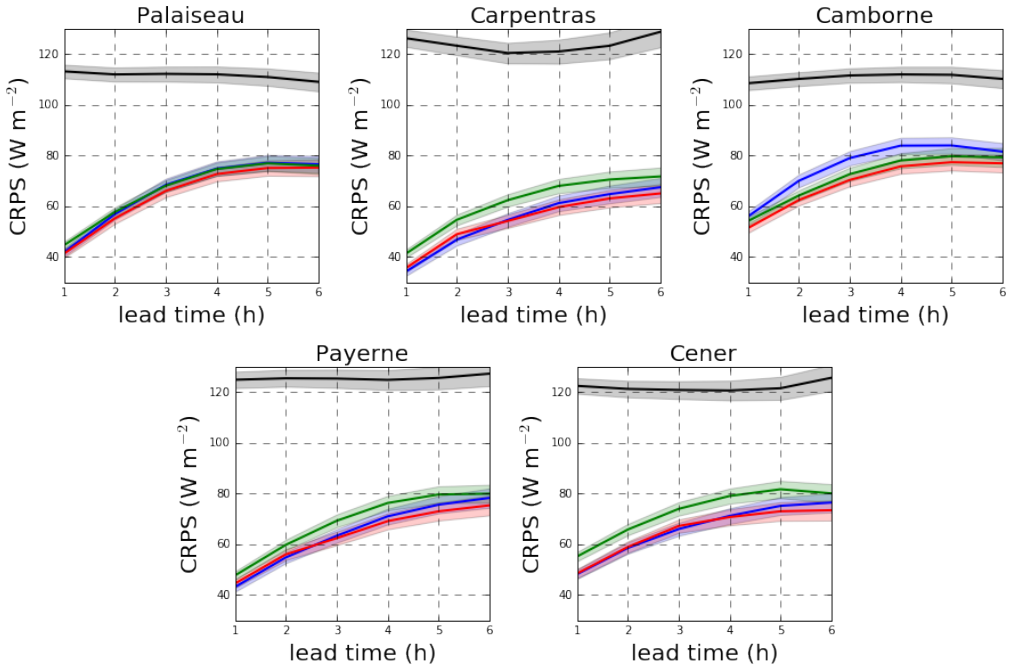


Figure 12: Mean CRPS and corresponding 95% bootstrap confidence interval as a function of lead time for the analog method (blue), the post-processed analog method (red), the PeEn (black), and the adaptive VAR(1) model (green).

rent one are found, and the concurrent past observations are then used as a forecast ensemble. Apart from the fact that it uses a different type of data, there is a fundamental difference in the assumption made by both methods. The analog ensemble method finds similar past NWP forecasts to the current forecast. The underlying assumption is thus that "if similar past forecasts are found, their errors will likely be similar to the errors of the current forecast, which can be inferred from theirs". In the present case, the hypothesis underlying the statistical downscaling method is that similar weather structures as defined by satellite-derived maps yield a similar local evolution in the cloud index. The analog ensemble method thus aims at forecasting an NWP

Station	Method	+1h	+2h	+3h	+ 4h	+ 5h	+ 6h
Palaiseau	analog	0.059	0.024	0.008	0.005	0.006	0.002
	p-analog	0.062	0.028	0.009	0.003	0.002	0.001
Carpentras	analog	0.181	0.123	0.093	0.069	0.049	0.032
	p-analog	0.151	0.093	0.093	0.074	0.055	0.045
Camborne	analog	0.014	-0.012	-0.014	-0.01	-0.009	-0.007
	p-analog	0.029	0.015	0.012	0.005	0.008	0.008
Payerne	analog	0.093	0.062	0.046	0.033	0.019	0.012
	p-analog	0.065	0.047	0.053	0.033	0.026	0.016
Cener	analog	0.086	0.056	0.041	0.033	0.026	0.016
	p-analog	0.081	0.051	0.033	0.034	0.03	0.021

Table 6: BSS relative to the adaptive VAR(1) method.

model error, while in the present case the aim is to forecast the evolution of measured GHI from a global observation.

The algorithm consists in the following steps:

1. given an observed satellite cloud index image, find k analogs, using the Euclidean distance in the previously defined four dimensional feature space

2. select the k concurrent ground measurements, corresponding to the times of the k satellite analogs. These are called "local" analogs. The corresponding "local" successors are the values that have been measured l hours after the "local" analogs, l being the lead time of the forecast

3. weight the "local" analog-successor pairs with the exponential kernel given in Eq. (7) where C_n^* is now the correlation between the analog satellite image and the current observation without any translation

4. use a local-linear operator between the "local" analogs and "local" successors to obtain a probabilistic forecast

7.2. Numerical experiment

We now evaluate the proposed downscaling method on the BSRN sites. The experiment set-up is the same as in Sec. 6. A local clear sky model is first defined using ground measured values of GHI, which defines ground values of cloud index. The optimal number of analogs is 80 for all the sites. It has been calculated by cross-validation on the training dataset.

For the sake of clarity, we present here only the ground RMSE, as defined in Eq. (17), but using the ground measurements of cloud index as the truth. The analog method is compared to ground clear-sky adjusted persistence, defined in the same way as in Section 6. An adaptive VAR(1) is also built, fitting linear regressions between the ground GHI measurements and their successors. Results are presented in Fig. 13 and show an improvement of +1.4%, +3.1% in RMSE with respect to the adaptive VAR(1) and the persistence method. The results of this downscaling method are also consistent with Table 4, and the VAR(1) method has worse scores than previously as compared to the analog method. The results of a post-processed analog method are also presented, using the same methodology as in the previous section. On the sites where the cloud regimes are not well defined, the post-

processing again brings an improvement. This shows that the bias corrected by the post-processing model is due to a bad evaluation of local cloud conditions, and is not linked to the type of data that is used. It gives additional evidence of the validity of the analysis presented in Section 4.

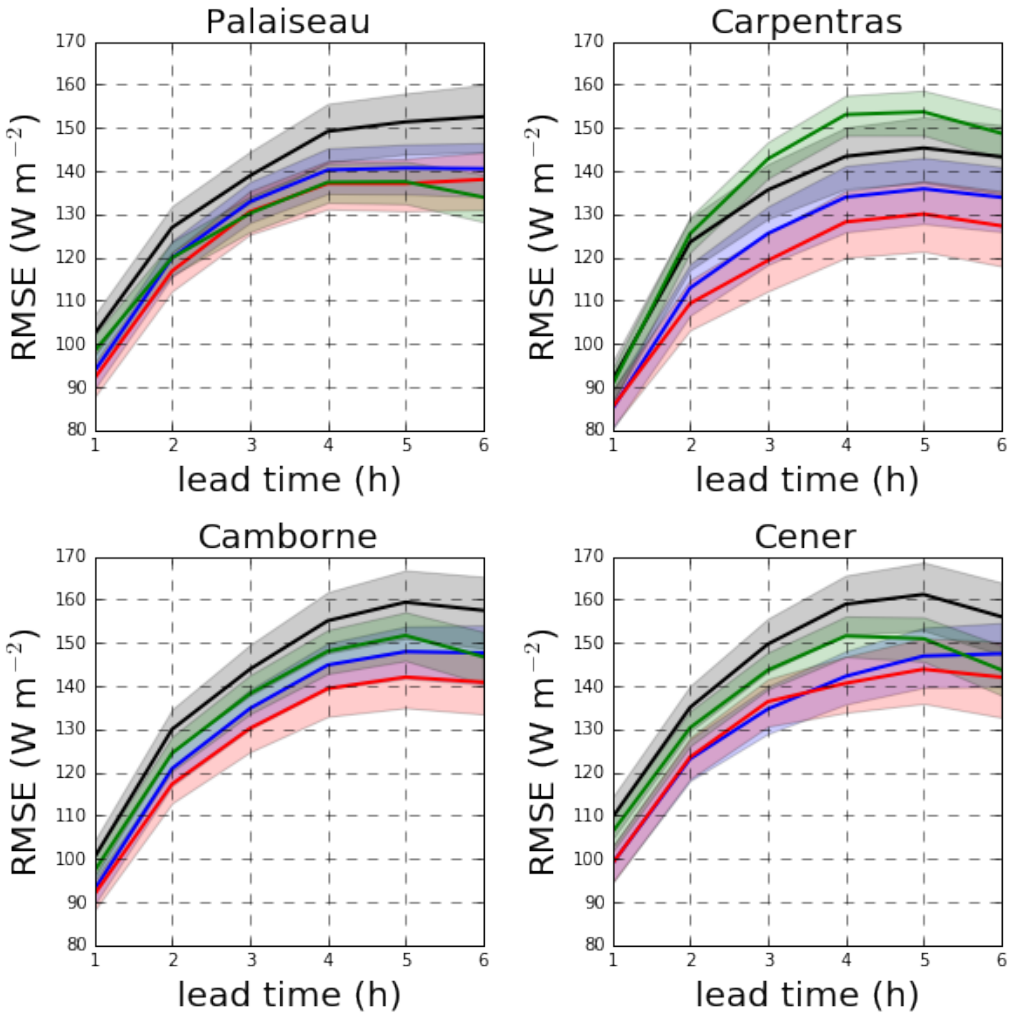


Figure 13: Normalized "ground" RMSE and corresponding 95% bootstrap confidence interval as a function of lead time for the analog method (blue), the post-processed analog method (red), the persistence method (black), and the adaptive VAR(1) model (green).

8. Conclusion

We have presented a computationally efficient method for GHI nowcasting on a given solar energy source, tested on sites representing different climatic conditions in Europe. The method uses a k-nearest neighbor algorithm on a four-dimensional feature space obtained from cloud index images to then apply a linear regression between selected analogs and successors. The methodology has proven to be robust to different geographical locations, and requires little tuning, no ground measurements nor a numerical weather model. We have also presented a framework to assess the performance of a given metric and set of features to choose the analogs, based on the analysis of their clustering performances. It provides a tool for evaluating the potential performance of the method on a location without having to run extensive numerical tests, and can be used when building an analog method.

The method has also been extended with a simple downscaling algorithm, when ground GHI measurements are also available. In both cases, the analog method shows a bias, which could be interpreted within the clustering-performance framework. A simple post-processing bias correction has been suggested, effectively improving the performances of the algorithm.

The method has proven to have potential for operational applications and future works will go in three directions. We first plan to use recent developments in information geometry for the analog-successor selection (as an alternative to the heuristic features). This would allow us to improve the identification of large scale cloud conditions, and thus reduce the intrinsic bias of the method.

The analog algorithm is a flexible method, designed to be easily combined with other algorithms in the context of operational forecasting. The combination of different forecast methods has been recently identified as an essential trend in solar forecasting (Yang et al., 2018). The downscaling procedure presented herein is a simple example of such a combination. It will be improved by nesting a statistical model forecasting local GHI values with a hidden Markov chain, defined with the satellite image analogs. Other machine learning methods (e.g., random forests, neural networks) will be tested with the same setup to better assess the analog method strengths.

Acknowledgments

The first author acknowledges the funding provided by Elum Energy and IMT Atlantique for this work. The authors also wish to thank the Elum Energy R&D team and Pr. Jordi Badosa for the valuable discussions. Finally, the authors would like to thank the three anonymous reviewers for their valuable comments that contributed to greatly improve the manuscript.

Aguiar, L. M., Pereira, B., David, M., Diaz, F., Lauret, P., 2015. Use of satellite data to improve solar radiation forecasting with bayesian artificial neural networks. Solar Energy 122, 1309–1324.

Alessandrini, S., Delle Monache, L., Sperati, S., Cervone, G., 2015. An analog ensemble for short-term probabilistic solar power forecast. Applied energy 157, 95–110.

Alexander, R., Zhao, Z., Székely, E., Giannakis, D., 2017. Kernel analog fore-

casting of tropical intraseasonal oscillations. Journal of the Atmospheric Sciences 74 (4), 1321–1342.

Alvarez, F. M., Troncoso, A., Riquelme, J. C., Ruiz, J. S. A., 2011. Energy time series forecasting based on pattern sequence similarity. IEEE Transactions on Knowledge and Data Engineering 23 (8), 1230–1243.

Atencia, A., Zawadzki, I., 2015. A comparison of two techniques for generating nowcasting ensembles. part ii: Analogs selection and comparison of techniques. Mon. Weather Rev. 143 (7), 2890–2908.

Box, G. E., Jenkins, G. M., Reinsel, G., 1970. Forecasting and control. Time Series Analysis 3, 75.

Brisson, A., Le Borgne, P., Marsouin, A., 1999. Development of algorithms for surface solar irradiance retrieval at osi saf low and mid latitudes. Eumetsat Ocean and Sea Ice SAF internal project team report.

Cano, D., Monget, J.-M., Albuisson, M., Guillard, H., Regas, N., Wald, L., 1986. A method for the determination of the global solar radiation from meteorological satellite data. Solar Energy 37 (1), 31–39.

Cervone, G., Clemente-Harding, L., Alessandrini, S., Delle Monache, L., 2017. Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble. Renewable Energy 108, 274–286.

Cleveland, W. S., 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American statistical association 74 (368), 829–836.

Dambreville, R., Blanc, P., Chanussot, J., Boldo, D., 2014. Very short term forecasting of the global horizontal irradiance using a spatio-temporal autoregressive model. Renewable Energy 72, 291–300.

Davò, F., Alessandrini, S., Sperati, S., Monache, L. D., Airoldi, D., Vespucci, M. T., 2016. Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting. Solar Energy 134 (Supplement C), 327 – 338.

Diagne, M., David, M., Boland, J., Schmutz, N., Lauret, P., 2014. Post-processing of solar irradiance forecasts from wrf model at reunion island. Solar Energy 105 (Supplement C), 99 – 108.

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, 65–76.

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. Journal of Business & economic statistics 13 (3), 253–263.

Escrig, H., Batlles, F., Alonso, J., Baena, F., Bosch, J., Salbidegoitia, I., Burgaleta, J., 2013. Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast. Energy 55, 853–859.

Foresti, L., Panziera, L., Mandapaka, P. V., Germann, U., Seed, A., 2015. Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall. Meteorological Applications 22 (2), 141–155.

Frouin, R., Chertock, B., 1992. A technique for global monitoring of net solar irradiance at the ocean surface. part i: Model. Journal of Applied meteorology 31 (9), 1056–1066.

Gilleland, E., Roux, G., 2015. A new approach to testing forecast predictive accuracy. Meteorological Applications 22 (3), 534–543.

Hammer, A., Heinemann, D., Hoyer, C., Kuhlemann, R., Lorenz, E., Müller, R., Beyer, H. G., 2003. Solar energy assessment using remote sensing technologies. Remote Sensing of Environment 86 (3), 423–432.

Hammer, A., Heinemann, D., Lorenz, E., Lückehe, B., 1999. Short-term forecasting of solar radiation: a statistical approach using satellite data. Solar Energy 67 (1), 139–150.

Heinemann, D., Lorenz, E., Girodo, M., 2006. Forecasting of solar radiation. Solar energy resource management for electricity generation from local level to global scale. Nova Science Publishers, New York.

Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather and Forecasting 15 (5), 559–570.

Huang, J., Troccoli, A., Coppin, P., 2014. An analytical comparison of four approaches to modelling the daily variability of solar irradiance using meteorological records. Renewable Energy 72, 195–202.

Inman, R. H., Chu, Y., Coimbra, C. F., 2016. Cloud enhancement of global horizontal irradiance in california and hawaii. Solar Energy 130, 128–138.

Le Borgne, P., Legendre, G., Marsouin, A., 2004. Meteosat and goes-east imager visible channel calibration. Journal of Atmospheric and Oceanic Technology 21 (11), 1701–1709.

Le Borgne, P., Legendre, G., Marsouin, A., Péré, S., Philippe, S., 2011. Meteosat and goes-e radiative fluxes validation report (products osi-303, 304, 305, 306).

Lguensat, R., Tandeo, P., Ailliot, P., Pulido, M., Fablet, R., 2017. The analog data assimilation. Monthly Weather Review.

Lloyd, S., 1982. Least squares quantization in pcm. IEEE transactions on information theory 28 (2), 129–137.

Lorenz, E., Hammer, A., Heinemann, D., et al., 2004. Short term forecasting of solar radiation based on satellite data. In: EUROSUN2004 (ISES Europe Solar Congress). pp. 841–848.

Lorenz, E. N., 1969. Atmospheric predictability as revealed by naturally occurring analogues. Journal of the Atmospheric sciences 26 (4), 636–646.

Marquez, R., Coimbra, C. F., 2013. Intra-hour dni forecasting based on cloud tracking image analysis. Solar Energy 91, 327–336.

Marquez, R., Pedro, H. T., Coimbra, C. F., 2013. Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to anns. Solar Energy 92, 176–188.

Mathiesen, P., Collier, C., Kleissl, J., 2013. A high-resolution, cloud-

assimilating numerical weather prediction model for solar irradiance forecasting. Solar Energy 92, 47–61.

Mathiesen, P., Kleissl, J., 2011. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental united states. Solar Energy 85 (5), 967–977.

Ohmura, A., Gilgen, H., Hegner, H., Müller, G., Wild, M., Dutton, E. G., Forgan, B., Fröhlich, C., Philipona, R., Heimo, A., et al., 1998. Baseline surface radiation network (bsrn/wcrp): New precision radiometry for climate research. Bulletin of the American Meteorological Society 79 (10), 2115–2136.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics 9 (1), 62–66.

Panziera, L., Germann, U., Gabella, M., Mandapaka, P. V., 2011. Nora – nowcasting of orographic rainfall by means of analogues. Q. J. R. Meteorol. Soc. 137 (661), 2106–2123.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., Vignola, F., 2002. A new operational model for satellite-derived irradiances: description and validation. Solar Energy 73 (5), 307–317.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., Hoff, T. E., 2010. Validation of short and medium term operational solar radiation forecasts in the us. Solar Energy 84 (12), 2161–2172.

Root, B., Knight, P., Young, G., Greybush, S., Grumm, R., Holmes, R., Ross,

J., 2007. A fingerprinting technique for major weather events. Journal of applied meteorology and climatology 46 (7), 1053–1066.

Rousseeuw, P. J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics 20, 53–65.

Scott, D. W., 2015. Multivariate density estimation: theory, practice, and visualization. John Wiley & Sons.

Soille, P. J., Ansoult, M. M., 1990. Automated basin delineation from digital elevation models using mathematical morphology. Signal Processing 20 (2), 171–182.

Stein, J. S., Hansen, C. W., Reno, M. J., 2012. The variability index: A new and novel metric for quantifying irradiance and pv output variability. In: World Renewable Energy Forum. pp. 13–17.

Thorey, J., Mallet, V., Chaussin, C., Descamps, L., Blanc, P., 2015. Ensemble forecast of solar radiation using tigge weather forecasts and helioclim database. Solar Energy 120, 232–243.

Van den Dool, H., 1994. Searching for analogues, how long must we wait? Tellus A 46 (3), 314–324.

Wang, Z., Koprinska, I., Rana, M., 2017. Solar power forecasting using pattern sequences. In: International Conference on Artificial Neural Networks. Springer, pp. 486–494.

Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T., Coimbra, C. F., 2018. History and trends in solar irradiance and pv power forecasting: A preliminary assessment and review using text mining. Solar Energy.

Zawadzki, I. I., 1973. Statistical properties of precipitation patterns. Journal of Applied Meteorology 12 (3), 459–472.

Zorita, E., Von Storch, H., 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. Journal of climate 12 (8), 2474–2489.