On the Use of the Coefficient of Variation to measure Spatial and Temporal correlation of Global Solar Radiation

Rudy Calif, Ted Soubdhan

On the Use of the Coefficient of Variation to measure Spatial and Temporal correlation of Global Solar Radiation

Rudy CALIF, Ted SOUBDHAN*

Laboratory LARGE, University of Antilles and Guiana, Guadeloupe, France

*Corresponding author. Address: University of Antilles and Guiana, 97159 Pointe à Pitre cedex, France. E-mail address: tsoubdha@univ-ag.fr

Abstract :

In this work we perform a statistical analysis of global solar radiation measured at two sites, 38 km distant from, in Guadeloupe, French West Indies. We have established a correlation model based on the coefficient of variation assuming a time scale separation. The coefficient of variation is calculated on 10 minutes interval with data measured at 1 Hz. This analysis highlight the dynamic correlation that can occur between measurements from two different sites with a time step of one second. From this results, knowing the coefficient of variation on a site we have established a new correlation model on this parameter for another site. A diagram linking the standard deviation on both sites, for a given coefficient of variation is proposed for correlated and non-correlated cases. More over this analysis put in evidence the existence of a threshold time under which there is no significant correlation.

This model can be useful in choosing new sites of production not only for establishing correlation from on site to another, but also in optimizing the geographical distribution of the solar farms within the territory.

Keywords : Solar radiation, cross-correlation, time scale separation, coefficient of variation

1.Introduction

1.1 Context

Because of clouds, solar radiation is a fluctuating data especially under tropical climate. Indeed, rapid changes in the local meteorological condition as observed in tropical climate can provoke large variation of solar radiation. The amplitude of these variations can reach up to 700 W/m². Moreover, these solar radiation variations can occur within short time interval (from few seconds to few minutes), depending on the clouds size, speed and number. The typical time scales associated with these solar radiation variations also vary significantly with the geographical location.

Studies of solar energy systems are traditionally performed using daily or hourly data [1] [2]. These data do not take into account the fluctuations mentioned previously. It has been shown that the fractional time distribution for instantaneous radiation differs significantly from that obtained with daily values.

Since solar energy systems are sensitive to instantaneous radiation fluctuations, simulations of these systems with daily or hourly data can lead to significant error especially under tropical climate. Indeed rapid variation of solar energy induces rapid and large variation of the output of such systems. For example solar cells (photovoltaics), used for electrical production, have very short response time and their electrical output will follow almost instantaneously the variations of solar radiation.

With a high density of photovoltaics generation in a power distribution grid, rapid fluctuations of the produced electrical power can lead to unpredictable variations of node voltage and power. In small grids as they exist on islands (such a Guadeloupe, FWI) such fluctuations can cause instabilities in case of intermediate power shortages with insufficient back-up capacity available. To manage the electrical network and the alternative power sources requires a better identification of these small

time scales variations. This stresses the need for power system operators to develop real time estimation tools for such disturbances.

Over the last years, some work were dedicated in establishing correlation between sites as a function of distance [3] and in parameterization of short term irradiance variability [4]. There has been resurgent interest in modeling and quantifying sub-hourly variability using different methods and metrics. In [4] the authors uses and frequency analysis based on spectral and coherence functions. This study showed high correlation for time scale greater than 3 hours :24h, 12h , 6h and 4h which are the main component in the irradiance signal.

Other studies have quantified the solar radiation variability with a prediction goal. Methods such as wavelets [6], fractal parameters [7], [8] or hybrid models [9] were used for a given site.

The Marquez formula [10] is also used to measure the variability, it is defined with the variation of the clear sky index CSI, $diff = CSI(t) - CSI(t-1)$ and $V = \sqrt{mean(diff^2)}$.

An extension of this metric (called P) proposed by Perez et al. [5] is based on the dispersion of the quantity diff $P = std(diff)$. Four metrics are used to characterize intra-hourly variability, including the standard deviation of the global irradiance clear sky index, and the mean index change from one time interval to the next, as well as the maximum and standard deviation of the latter. Different time scales are studied from 20 seconds to 15 minutes.

In the present work we compare and analyze the statistical moments of the solar radiation parameters measured at two sites in the Guadeloupean archipelago with a time step of one second. We have established a correlation model for the coefficient of variation defined over a ten minutes period.

The paper is organized as follow: the theoretical frame is exposed in section 2,

This model can be useful in choosing new sites of production not only for establishing correlation from on site to another, but also in optimizing the geographical distribution of the solar farms within the territory to favor a smoothing of the overall available solar electricity production. Indeed this choice should be done not only on the basis of the annual potentiality of a site but also from the stability of its production.

1.2 Methodology

In the present work, we compare and analyze the solar global radiation signals G_1, G_2, measured at two sites in the Guadeloupean Archipelago. This study put in evidence the existence of a dynamical correlation between G_1 and G_2. To do so, we calculate for each signal, the global solar radiation fluctuation $G'_{[T_1;\,T_2]}$ for time scales T= 10 min, 20 min, 30 min, …, 3 hours. The cross-correlation coefficient $C_{G_1G_2}$ is used for each time scale T. From this result, we determine the time scale threshold T_{trs}: for$T < T_{trs}$, the two signals are decorrelated and for $T > T_{trs}$, the two signals are correlated. To evaluate the scattering effect of two sites, we determine the coefficient of variation for the sum signal G_{1+2}.

Finally we propose, analytical relationship of coefficient of variation I_{1+2} for two cases i) two signals statistically independent with different momentum, ii) two signals correlated with different momentum.

2.Theoretical frame

2.1 Decomposition of solar radiation signal using moving average filter

The moving average of the instantaneous solar radiation G, at a given instant t for a given averaging time T, is defined as(Papoulis):

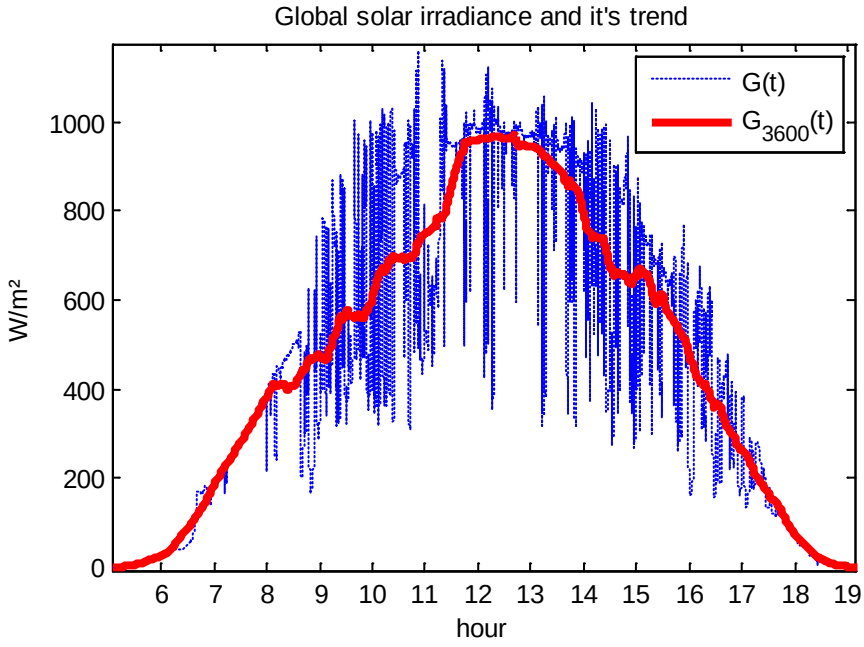
$$\overline{G_{\mathrm{T}}}(t) = \frac{1}{N}\sum_{i=t-\frac{T-1}{2}}^{i=t+\frac{T+1}{2}} G(i) \qquad (1)$$

93 The instantaneous solar radiation G can then be expressed as (ref):

94 $$G(t) = \overline{G_T}(t) + G'_{\mathrm{T}}(t) \qquad (2)$$

95 Where $G'_{\mathrm{T}}(t)$ is the solar radiation fluctuation for time scales smaller then T while the moving
96 average $\overline{G_T}(t)$ gives the solar radiation evolution for time scales larger than T. This decomposition
97 is illustrated in figure (1), where $\overline{G_T}(t)$ is superimposed to $G(t)$ for a measurement duration of one
98 day and for $T = 3600s$. Figure (1.a) illustrates the fluctuations obtained from equation (2).

99



100

101 Figure1: An example of instantaneous solar radiation signal (red line) and it's trend (black line).

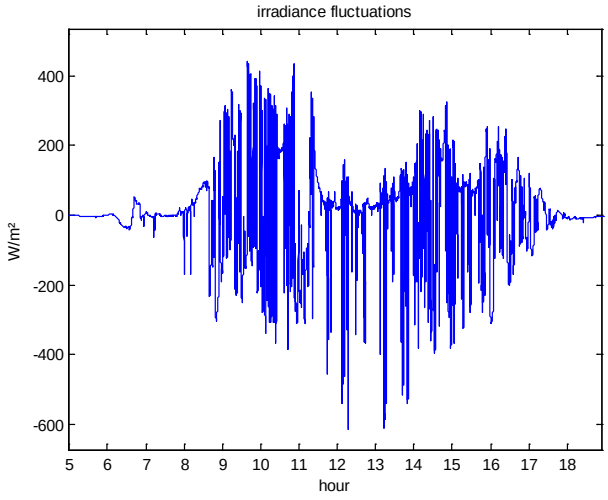


Figure 1a: global solar radiation fluctuations obtained from moving average decomposition

Besides, the difference between two moving average having respectively averaging time T_1 and T_2, gives the fluctuation for time scales ranging on $[T_1, T_2]$:

$$G'_{T_1-T_2}(t) = \overline{G_{T_1}}(t) - \overline{G_{T_2}}(t) \qquad (3)$$

2.2 Coefficient of variation defined for two sites.

In this study we consider solar radiation signal of 10 min length defined arbitrary. Let consider $G_1 \begin{cases} \bar{G}_1 \\ \sigma_1 \end{cases}$ and $G_2 \begin{cases} \bar{G}_2 \\ \sigma_2 \end{cases}$ two signals measured at site 1 and 2 respectively. We can define, for each G_1 and G_2, from the two momentums $\bar{G}$ and σ, the coefficient of variation, a normalized measurement of dispersion of a probability distribution as:

$$I = \frac{\sigma}{\bar{G}}$$

This parameter is a second order statistic moment measuring the degree of variability of a given process. In the turbulence and the wind energy communities, it's called turbulence intensity and is classically used. Turbulence intensity is a metric characterizing turbulence expressed as a percent. For an illustration, an idealized flow of air with absolutely no fluctuations in air speed would have a turbulence intensity value of 0%. This idealized cases never occurs on earth. However, the values greater than 100% are possible. This can happen, for example, when the average air speed is small and there are large fluctuations present. The higher the coefficient is, the more turbulent the flow.

This parameter is computed for the first time with global solar radiation data. This could be an interest for the solar energy community. It could be a simple way in order to qualify and classify the variability level of sites.

Let consider the sum G_{1+2} $is\ defined\ by$ $G_{1+2}\begin{cases}\bar{G}_{1+2}\\ \sigma_{1+2}\end{cases}$ with G_{1+2}= G_1+ $G_2$and $\overline{G_{1+2}}$ = $\overline{G_1}$+ $\overline{G_2}$

Let us define σ_{1+2}. By definition of standard deviation

$$\sigma_G = \sqrt{\overline{(\mathrm{G}-\bar{G})^2}}$$

thus,

$$\sigma_{1+2} = \sqrt{\overline{(G_1+G_2)^2 + (\overline{G_1+G_2})^2 - 2(G_1+G_2)(\overline{G_1+G_2})}}$$

This equation becomes

$$\sigma_{1+2} = \sqrt{\overline{G_1}^2 + \overline{G_2}^2 + 2\overline{G_1G_2} + \overline{G_1}^2 + \overline{G_2}^2 + 2\overline{\mathrm{G}_1}\,\overline{\mathrm{G}_1} - 2\,\overline{\mathrm{G}_1\overline{\mathrm{G}_1}} - 2\,\overline{\mathrm{G}_1\overline{\mathrm{G}_2}} - 2\,\overline{\mathrm{G}_2\overline{\mathrm{G}_1}} - 2\,\overline{\mathrm{G}_2\overline{\mathrm{G}_2}}}$$

A) Assuming that G_1 and G_2 are statistically independent and have same moment order:

We have

$\overline{G_1 G_2} = \overline{G_1}\,\overline{G_2}$

And

$\overline{G_1}=\overline{G_2}$

With this assumption equation (2) becomes:

$$\sigma_{1+2} = \sqrt{2(\overline{G_1^{\,2}} - \overline{G_1}^2)}$$

Hence we can define the coefficient of variation of G_{1+2} as:

$$I_{1+2} = \frac{\sigma_{1+2}}{\overline{G_{1+2}}} = \frac{\sqrt{2(\overline{G_1^{\,2}} - \overline{G_1}^2)}}{2\overline{G_1}}$$

$$I_{1+2} = \frac{1}{\sqrt{2}}\frac{\sigma_1}{\overline{G_1}} \qquad (4)$$

This result can be extended to n independent signals having the same moments:

$$I_n = \frac{\sqrt{(\sum_{i=1}^{n} \sigma_i{}^2)}}{\sum_{i=1}^{n} \overline{G_i}} \qquad (5)$$

B) Let us now consider G_1 and G_2 two signals statistically independent with different momentum.

We can decompose each signal as $G = \overline{G} + G'$

Were G' is the fluctuation in the signal around the mean value $\overline{G}$.

Thus the sum is

152 $G_{1+2}= \overline{G_1}+ G_1{}'+\overline{G_2}+ G_2{}'$

153 $G_{1+2}=\overline{G_1} + \overline{G_2}+ G_1{}'+G_2{}'$

154 $G_{1+2}=\overline{\overline{G_{1+2}}} + G_{1+2}{}'$

155 With

156 $\overline{\overline{G_{1+2}}}=\overline{G_1} + \overline{G_2}$

157 and $G_{1+2}{}' = G_1{}'+G_2{}'$

158 The standard deviation of G_{1+2} is

159 $$\sigma_{1+2} = \sqrt{\overline{(G_{1+2} - \overline{G_{1+2}})^2}}$$

160 Elsewhere $\sigma_{1+2} = \sqrt{\overline{(G_1{}' + G_2{}')^2}} =\sqrt{\overline{(G_1{}'^2 + 2G_1{}'G_2{}' + G_2{}'^2)}}$

161

162 As the signals are considered statistically independent and decorrelated, we have

163 Corr $(G_1{}', G_2{}') = c = \frac{\sqrt{\overline{G_1{}'G_2{}'}}}{\sigma_{1\sigma_2}}= 0$

164 Where *c* is the normalized correlation coefficient

165 Thus $\overline{G_1{}'G_2{}'}= 0$

166 And

167 $\sigma_{1+2} = \sqrt{\overline{(G_1{}'^2 + G_2{}'^2)}} = \sqrt{\overline{(\sigma_1{}^2 + \sigma_2{}^2)}}$

168 And the variation coefficient is

$$I_{1+2} = \frac{\sqrt{\overline{\sigma_1{}^2 + \sigma_2{}^2}}}{\overline{G_1} + \overline{G_2}} \qquad (6)$$

C) Now let us consider two signals correlated with different momentum as

$$c = \frac{\sqrt{\overline{G_1{}' G_2{}'}}}{\sigma_{1\sigma_2}} \text{ and } \sigma_{1+2} = \sqrt{\overline{(G_1{}'^2 + 2G_1{}' G_2{}' + G_2{}'^2)}}$$

$$\sigma_{1+2} = \sqrt{\overline{\sigma_1{}^2 + 2(c\,\sigma_1\sigma_2)^2 + \sigma_2{}^2)}}$$

As said previously c is the normalized correlation coefficient and the variation coefficient is

$$I_{1+2} = \frac{\sqrt{\overline{\sigma_1^2 + 2(c\sigma_1\sigma_2)^2 + \sigma_2^2}}}{\overline{G_1} + \overline{G_2}} \qquad (7)$$

Consequently, using the analytical relation given above, considering two signals decorrelated, c=0, we found the analytical expression given in equation 6. Besides, if we consider two signals correlated, c=1, the analytical relation given 7 becomes:

$$I_{1+2} = \frac{\sqrt{\overline{\sigma_1^2 + 2(\sigma_1\sigma_2)^2 + \sigma_2^2}}}{\overline{G_1} + \overline{G_2}}$$

3 Solar radiation measurements

In this study, the solar radiation measurements sampled at $1hz$ were measured during one year, from January 2006 to January 2007, at two different sites in Guadeloupe. The daily average for the solar load on a horizontal surface is around 5 kWh/m². A constant sunshine combined with the thermal inertia of the ocean makes the air temperature variation quite weak, between 17°C and 33°C with an average of 25°C to 26°C. Relative humidity ranges from 70% to 80% and the trade winds are relatively constant all along the year. Two main regimes of cloudiness are superposed:

the clouds driven by the synoptic conditions over the Atlantic Ocean and the orographic cloud layer generated by the local reliefs.

4 Results

4.1 Time scale separation

To put in evidence an eventually dynamical correlation between the two sites, the cross-correlation coefficient $C_{G'_1 G'_2}$ is defined as [3]:

$$R_{G'_1 G'_2}(\tau) = \begin{cases} \sum_{n=0}^{T-\tau-1} {G'}_{1n+\tau} {G'}_2 y_n & \tau \geq 0, \\ R_{G'_1 G'_2}(-\tau) & \tau < 0 \end{cases}$$

$$C_{G'_1 G'_2}(\tau) = \frac{R_{G'_1 G'_2}(\tau)}{\sigma_{G'_1} \sigma_{G'_2}}$$

Where G'_1, G'_2 are zero-mean stochastic variables, $\sigma_{G'1}$ and $\sigma_{G'2}$ are the corresponding standard deviation . For each day we have calculated this coefficient $C_{G'_1 G'_2}$ for different time scales.

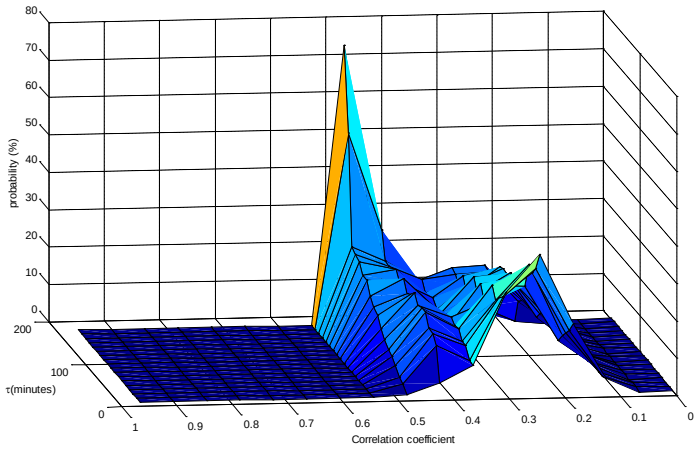
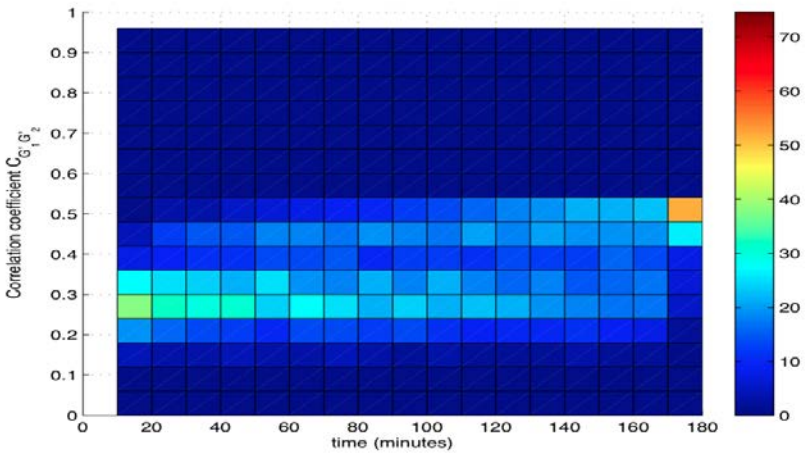


Figure 2: The conditional probability of the cross correlation.



199

200 Figure 3: percentage of correlation values for different τ

201 For the highlighting of a threshold time T_{trs} under which the fluctuations of two sites
202 are not correlated, we determine the correlation coefficient $C_{G'_1 G'_2}(\tau)$ for the fluctuations
203 $G'_{T_1-T_2}(t)$ defined in equation 3, with T_1=10j minutes and T_2=T_1+10 minutes, for j between
204 1 and 18. The figure illustrates the conditional probability $P(C_{xy}|\Delta t)$, with Δt = T_2 – T_1. We
205 observe that the maximum value of the probability increase 0.25 to 0.5 for time range
206 [10;170] minutes. We observe a significant correlation for time [160; 170] and $C_{G'_1 G'_2}$= 0,5.
207 This correlation keeps increasing for higher time scale. Thus the threshold time is estimated
208 to be $T_{trs} = 170\ min$.

209

4.2 Time scales for uncorrelated signals: T<T_{trs}.

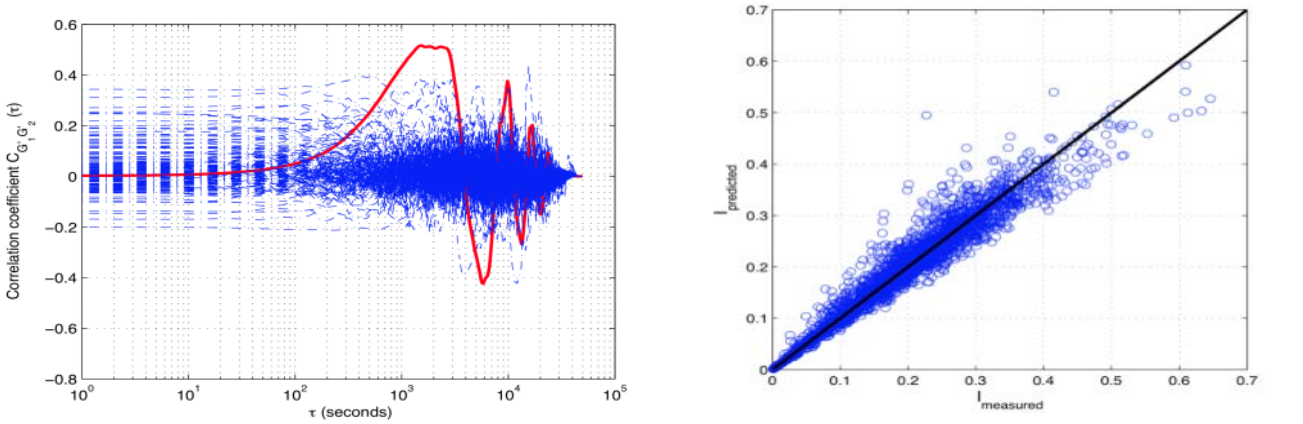


Figure 4: a (left) The cross correlation function in semi-log diagram. The red curve correspond to the cross correlation for an example of a marginal event, and the blue curve represent the cross correlation for the majority of the events. b(right) the predicted coefficient of variation versus the experimental coefficient of variation.

In this section, the figure 3a presents the cross correlation coefficient of the two sites, for time scales T< T_{trs}. In 97% of cases, the value of $C_{G'_1G'_2}(\tau)$ is smaller than $0{,}4$. For time going to [10; 50] minutes the maximum correlation coefficient is $C_{G'_1G'_2}$ =0,25.

A weak value of the cross correlation coefficient for two measurement sites indicates that the solar radiation fluctuations are statistically independent. This shows that the fluctuations are not correlated on these time scales.

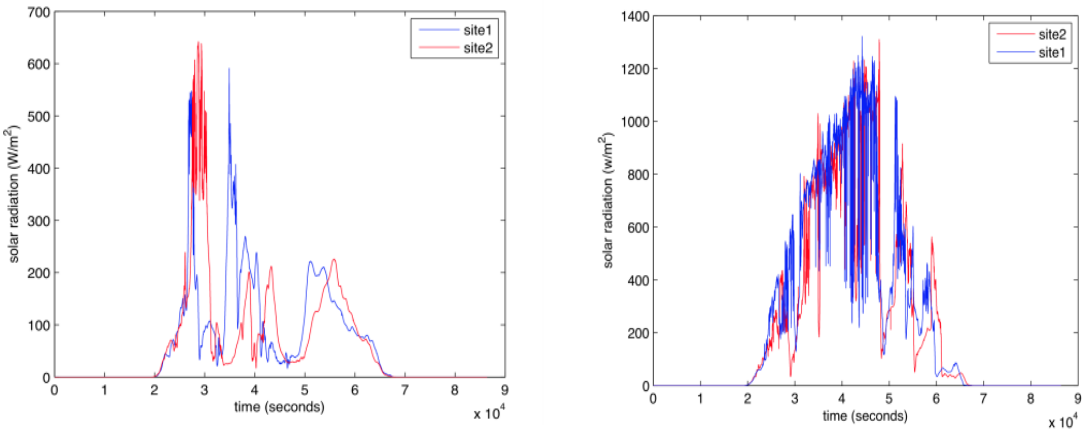


Figure 5: Examples of correlated marginal events for small time scales.

However, in 2% of cases we observe a dynamical correlation for C_{xy} values superior to 0,5. These marginal events correspond to extreme meteorological situations (clear sky or cloudy sky) as illustrated in figure 4.

Besides, for these time scales, in figure 3b, we have plotted the experimental coefficient variation $\mathrm{I}_{1+2\mathrm{exp}} = \frac{\sigma_{1+2}}{\overline{G_{1+2}}}$ versus the analytical relationship given in equation 6, $I_{1+2\mathrm{theo}} = \frac{\sqrt{\overline{\sigma_1}^2+\overline{\sigma_2}^2}}{\overline{G_1}+\overline{G_2}}$. A good agreement is observed between the predicted coefficient variation and the measured coefficient variation, confirming a non-synchronous variability of the two sites for times scales $T<T_{Trs}$.

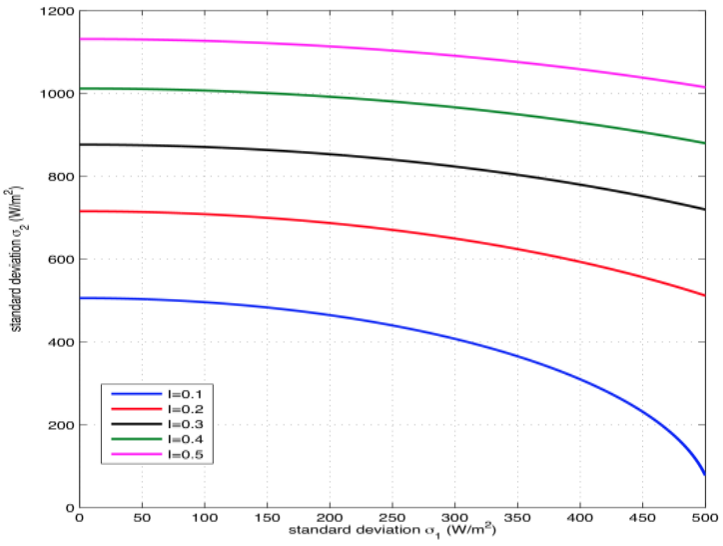


Figure 6: Diagram of the predicted coefficient of variation determined with the standard deviations of two sites, uncorrelated cases.

Moreover, in figure 6, we give the evolution of the predicted coefficient variation versus the given standard deviation σ_1 and σ_2, respectively for a site 1 and site 2.

4.3 Time scales for correlated signals: T>T_{trs}

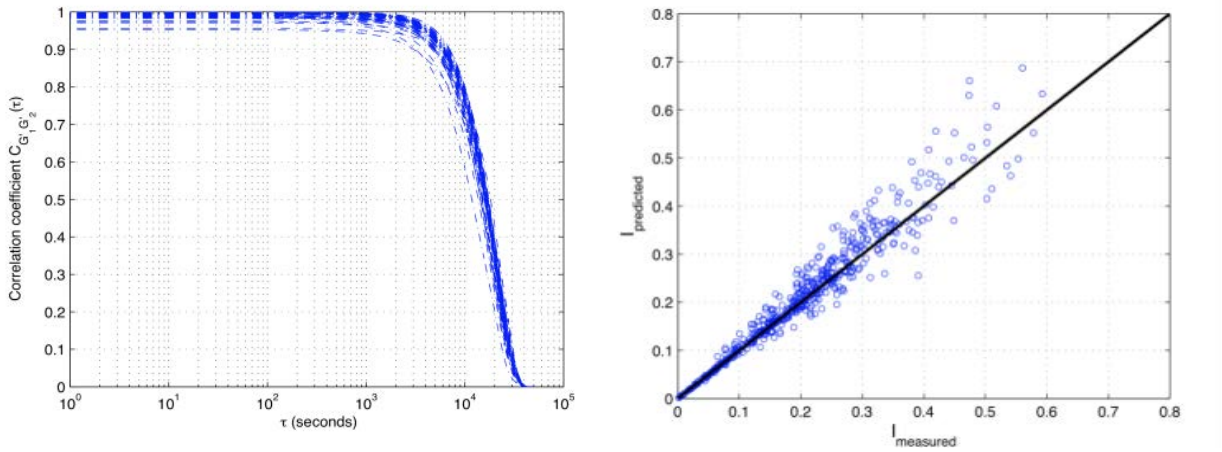


Figure 7: a (left) the cross correlation coefficient in semi-log diagram. b(right) the predicted coefficient of variation versus the experimental coefficient of variation.

The figure 7a displays the cross correlation of the two sites, for time scales T> T_{trs}. We can observe that the cross correlation coefficients almost are close to 1 obtained for a time delay equal to 0.

For these time scales, the characteristic length of the convective structure, in the atmospheric sub- layer, is of the order of several hundreds of kilometers (ref); we recall that the distance between the two sites is around to 40 kilometers.

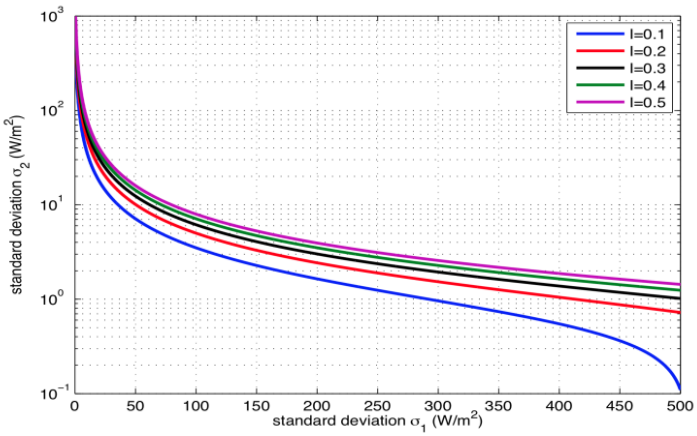


Figure 8: Diagram of the predicted coefficient of variation determined with the standard deviations of two sites, correlated cases.

For these time scales, we illustrated in figure 7b, the predicted coefficient of variation $I_{1+2\text{theo}}$ given in equation 7, is plotted versus the experimental coefficient of variation, indicating a good agreement between the predicted values and the measured values. The figure 8 presents the evolution of the predicted coefficient of variation versus the given standard deviation σ_1 and σ_2, respectively for a site 1 and site 2.

5. Discussions

In this study, we attempt to quantify the variability and the coupling of the global solar radiation for two sites in Guadeloupe, through the coefficient of variation, on ten minutes interval. This coefficient is defined from the mean and standard deviation of solar irradiance over the considered interval.

We have established a model for the coefficient of variation assuming a time scale separation. Two cases can be distinguished according to the correlation of the first order moment (mean).

The analysis of the cross correlation coefficient put in evidence à threshold under which there is no significant correlations of the global solar radiation fluctuations. According to this assumption we have established a model for the coefficient of variation defined for a pair of site, each one being characterized by the first and second order momentums $\bar{G}$ and σ. This is given by the equation (6).

When the signal are significantly correlated, the coefficient of variation of a pair of site is given by equation (7).

Hence according to this model we can determine the coefficient of variation of a pair of site knowing the mean and standard deviation on one of the sites.

The model was tested on experimental data for both cases, correlated and non-correlated. In both cases the predicted values by the model are in good agreement with the calculated coefficient of variation.

For both cases a diagram is proposed to determine the coefficient of variation of a pair of site. This model can be useful to determine in a very simple and easy way the amount of energy and the associated fluctuations that may occur on a pair of site knowing the conditions on one site.

When considering the PV production on different sites, it becomes possible to estimate the smoothing effect of the sum knowing the PV power at one site.

This study cans also be helpful in choosing the best places for PV plants to take advantage of the smoothing effect due to the connection of different sites.

References

[1] G. Notton, C. Cristofari, P. Poggi, M. Muselli, Calculation of solar irradiance profiles from hourly data to simulate energy systems behaviour. Renewable Energy 27 123–142, 2002.

[2] G. Notton, M. Muselli and A. Louche, Two estimation methods for monthly mean hourly total irradiation on tilted surfaces from monthly mean dayly horizontal irradiation from solar radiation data of Ajaccio. Solar Energy Vol. 57, No. 2, pp. 141-153. 1996

[3] S. Papoulis et S. Unnikrishna Pillai. Probability, Random variables and stochastic processes. Mc Graw Hill, 2002.

[4] R. Perez, J. Schlemmer, S.Kivalov, K. Hemker Jr, T. Hoff, Short term irradiance station pair correlation as a function of distance, American Solar Energy Society, ASES Conference, Raleigh, 2011.

[5] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, T.Hoff, Parameterization of site-specific short-term irradiance variability, Solar Energy Volume 86, Issue 8, August 2012, Pages 2170–2176.

[6] M. Lave, J. Kleissl, Solar variability of four sites across the state of Colorado, Renewable Energy 35 (2010), Pages 2867-2873.

[8] C. Monteiro, L. Fernandez-Jimenez, I. Ramirez-Rosado, A. Munoz-Jimenez, and P. Lara-Santillan, "Short-term forecasting models for photovoltaic plants: Analytical versus soft-computing techniques," Mathematical Problems in Engineering, vol. 2013, 2013.

[9] M. Bouzerdoum, A. Mellit, and A. Massi-Pavan, "A hybrid model (sarima–svm) for short-term power forecasting of a small-scale grid-connected photovoltaic plant," Solar Energy, vol. 98, pp. 226–235, 2013.

[10] Marquez R, Coimbra CFM. Proposed Metric for Evaluation of Solar Forecasting Models. J Sol Energy Eng 2012;135:011016–011016. doi:10.1115/1.4007496.

[11] P.Mandal, S. Madhira, J.Meng, and R. Pineda, "Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques," Procedia Computer Science, vol. 12, pp. 332–337, 2012.

[12] E. Lorenz, J. Hurka, D. Heinemann, and H. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 2, no. 1, pp. 2–10, 2009.

[12] J. Shi, W. Lee, Y. Liu, Y. Yang, and P. Wang, “Forecasting power output of photovoltaic systems based on weather classification and support vector machines,” Industry Applications, IEEE Transactions on, vol. 48, no. 3, pp. 1064–1069, 2012.