Analysis and synthesis of the variability of irradiance and PV power time series with the wavelet transform

O. Perpiñán[a,1,*], E. Lorenzo[b]

[a]*Electrical Engineering Department, EUITI-UPM, Ronda de Valencia 3, 28012 Madrid, Spain.*
[b]*Instituto de Energía Solar, UPM, Ciudad Universitaria s/n, 28040 Madrid, Spain.*

Abstract

The irradiance fluctuations and the subsequent variability of the power output of a PV system are analysed with some mathematical tools based on the wavelet transform. It can be shown that the irradiance and power time series are nonstationary process whose behaviour resembles that of a long memory process. Besides, the long memory spectral exponent α is a useful indicator of the fluctuation level of a irradiance time series. On the other side, a time series of global irradiance on the horizontal plane can be simulated by means of the wavestrapping technique on the clearness index and the fluctuation behaviour of this simulated time series correctly resembles the original series. Moreover, a time series of global irradiance on the inclined plane can be simulated with the wavestrapping procedure applied over a signal previously detrended by a partial reconstruction with a wavelet multiresolution analysis, and, once again, the fluctuation behaviour of this simulated time series is correct. This procedure is a suitable tool for the simulation of irradiance incident over a group of distant PV plants. Finally, a wavelet variance analysis and the long memory spectral exponent show that a PV plant behaves as a low-pass filter.

Keywords: irradiance fluctuations, power fluctuations, wavelet analysis, wavestrapping, nonstationary, long memory process.

Nomenclature

ACVS   Autocovariance sequence.

α Spectral exponent of a long memory process.

B Backward shift operator.

β Wavelet variance exponent of a long memory process.

$B_o(0)$ Extra-atmospheric irradiance on the horizontal plane.

CDF   Cumulative distribution function.

Δt Sampling time of a signal.

$\widetilde{\mathcal{D}}_j$ jth level detail of an MRA with a MODWT.

---

*Corresponding author
Email addresses:* `o.perpinan@upm.es` (O. Perpiñán), `lorenzo@ies-def.upm.es` (E. Lorenzo)
[1]ISES member

$\mathcal{D}_j$ jth level detail of an MRA with a DWT.

DWT   Discrete wavelet transform.

$G(0)$ Global irradiance on the horizontal plane.

G_{ef} Effective global irradiance incident on the generator.

G_{ef}^B Wavestrapped effective global irradiance time series.

G_{ef}^D Detrended effective global irradiance time series.

$G_{ef}^{D,B}$ Wavestrapped version of a detrended effective global irradiance time series.

G_{ef}^T Trend of a effective global irradiance time series.

$\mathcal{G}_{ef}(f)$ Fourier transform of a irradiance time series.

$\mathcal{H}(f)$ Fourier transform of the transfer function of a filter.

$\widetilde{\mathcal{H}}_j^D$ The squared gain function associated with the wavelet filter at scale j.

k_t Clearness index.

λ_{J_0} Scale J_0 of a WT.

MODWT   Maximal overlap discrete wavelet transform.

MRA   Multiresolution analysis

$\nu_{X,j}^2$ Wavelet variance of the scale τ_j.

$\nu_G^2(\tau_j)$ Wavelet variance at the j scale of the irradiance time series.

$\nu_P^2(\tau_j)$ Wavelet variance at the j scale of the power time series.

$\mathcal{P}(f)$ Fourier transform of a power time series.

SDF   Spectral density function.

$\overline{S_{G,j}}(f)$ Average SDF at the j scale of the irradiance time series.

σ_X^2 Variance of the time series X_t.

$\widetilde{\mathcal{S}}_{J_0}$ J_0th level smooth of an MRA with a MODWT.

$\mathcal{S}_{J_0}$ J_0th level smooth of an MRA with a DWT.

$S_j(f)$ Spectral density function of a time series of the MODWT coefficients at scale j.

$\overline{S_{P,j}}(f)$ Average SDF at the j scale of the power time series.

$S_X(f)$ Spectral density function of a time series X_t.

τ_j Scale j of a WT.

$\widetilde{\mathbf{V}}_{J_0}$ Vector of MODWT scaling coefficients associated with the scale λ_{J_0}.

$\mathbf{V}_{J_0}$ Vector of DWT scaling coefficients associated with the scale λ_{J_0}.

$\widetilde{\mathbf{W}}$ N dimensional vector of MODWT coefficients.

$\mathbf{W}$ N dimensional vector of DWT coefficients

$\widetilde{\mathcal{W}}$ $N \times N$ real-valued matrix defining the MODWT.

$\mathcal{W}$ $N \times N$ real-valued matrix defining the DWT.

$\widetilde{\mathbf{W}}_j$ Vector of MODWT coefficients associated with the scale τ_j.

$\mathbf{W}_j$ Vector of DWT coefficients associated with the scale τ_j.

WT   Wavelet transform

$\mathbf{X}$ N dimensional vector containing a real-valued time series.

X_t Real-valued time series.

1. Introduction

The integration of large PV plants in the conventional electrical grid must cope with several problems [1]. One of the most important is the variability of the power output of the system due to the irradiance fluctuations. This variability could force the grid operator to take decisions in order to guarantee the stability of the electrical system. Therefore it is important to quantify the intensity, duration and probability of occurrence of power fluctuations. This issue has been previously analysed in [2–5] with some important results:

- The power fluctuation level of a PV system depends on its size. A PV system smooths the fluctuation in a manner inversely proportional to the square root of its area.

- Moreover, the power fluctuation level in a zone where a group of PV systems are connected depends on the number of plants of the set, following a decreasing exponential relation.

These smoothing effects depend on the characteristics of the group of PV systems (size of the individual systems and geographical relations) but also on the meteorology of the location. The irradiance profiles incident on the PV plant of the set evolve differently due to the movement and transformation of clouds. It is plausible to assume that these different profiles own the same statistical properties if they are not too far from one another. Thus, a tool for the simulation of collections of time series of irradiance may be valuable in order to analyse the power fluctuations from a hypothetic group of PV systems.

Several procedures have been proposed for the synthesis of *daily* time series of irradiation [6]. These procedures are mainly based on the premises of stationarity and short memory of the clearness index [7]. However, these conditions do not hold for clearness index with short sampling times. If the simulation procedure is constructed with classical methods of time series analysis (AR, MA, or ARMA models, for example) it is necessary to constrain the analysis to temporal windows where the process can be regarded as stationary. Thus, the first step of the desired procedure of simulation is to cut the day in suitable periods where the stationary conditions are valid. Moreover, as it will be shown later, the *instantaneous* clearness index is a long memory process, so several previous samples are needed in the regression formula for the estimation of a new sample.

On the other hand, a non-parametric analysis with the wavelet transform (WT) do not impose such restrictions. The WT carries out a *multiresolution analysis* (MRA) of an time series, where each of the decompositions of the analysis is a representation of the original signal with a different temporal scale. Besides, the combination of bootstrap methods with the wavelet decomposition is a powerful procedure to obtain new replicas of a original irradiance signal with the same statistical properties.

This paper is organised as follows: a summary of WT and its application to irradiance time series analysis is included in section 2.1; the wavelet variance and its relation with spectral analysis is explained in section 2.2; the long memory processes and their application to the irradiance time series are shown in section 2.3; the wavestrapping technique (combination of bootstrap with the WT) is explained in section 2.4; finally, a procedure for the simulation of time series of irradiance on the horizontal plane (section 3.1) and on the inclined plane (section 3.2) is developed. Conclusions and future actions are summarised in section 4.

The irradiance and power output data analysed in this paper belongs to the Milagro PV plant located at Navarra, Spain. This PV system, installed by Acciona Solar, was started up in 2007. It is composed of more than 800 azimuthal trackers summing a generator power of 9,5 MWp[2]. The database comprises a year of time series of irradiance, power output, wind speed, temperature and some other variables with one sample per second.

2. Wavelets for irradiance time series

The analysis of meteorological time series with the wavelet transform can be easily found in a variety of research papers. One of the first comprehensive analysis was published in 1995 [8]. Later, a practical guide was proposed in 1998 [9], with more recent review papers published in 2005 [10, 11]. The wavelet transform has also been applied to the analysis of solar irradiance, for example in [3, 12–14].

The theoretical development of the next section follows the book of Percival and Walden [15].

2.1. The Wavelet Transform

Let $\mathbf{X}$ be an N dimensional vector containing the real-valued time series $X_t : t = 0, \ldots, N - 1$. The discrete wavelet transform (DWT) of level J_0 (with $N = 2^{J_0}$ to obtain a full transform) of $\mathbf{X}$ is an orthonormal transform given by $\mathbf{W} = \mathcal{W}\mathbf{X}$, where $\mathbf{W}$ is an N dimensional vector of DWT coefficients, and $\mathcal{W}$ is an $N \times N$ real-valued matrix defining the DWT. Both the DWT coefficients and matrix can be partitioned such that $\mathbf{W}_j = \mathcal{W}_j\mathbf{X}$ and $\mathbf{V}_{J_0} = \mathcal{V}_{J_0}\mathbf{X}$. Here $\mathbf{W}_j$ is an $N_j \equiv N/2$ dimensional vector of wavelet coefficients associated with *changes* on the scale $\tau_j \equiv 2^{j-1}$ and $\mathbf{V}_{J_0}$ is an N_{J_0} dimensional vector of scaling coefficients associated with *averages* on the scale $\lambda_{J_0} \equiv 2^{J_0}$. The vector $\mathbf{X}$ can be synthesised from $\mathbf{W}$ with an additive decomposition in terms of the vectors $\mathcal{D}_j \equiv \mathcal{W}_j^T\mathbf{W}_j$ and $\mathcal{S}_{J_0} \equiv \mathcal{V}_{J_0}^T\mathbf{V}_{J_0}$, the jth level detail and the J_0th level smooth respectively.

However, the DWT is sensitive to the starting position of the time series and depends on whether a given event resides within a wavelet averaging window. In other words, the DWT is not translation-invariant, so it is not best suited for the analysis of time series. A modified version of the DWT is better adapted to this task. This modified version has received several names such as undecimated or stationary DWT. However, we will follow [15, 16] using the term "maximal overlap DWT" (MODWT).

The MODWT of $\mathbf{X}$ is a transform given by $\widetilde{\mathbf{W}} = \widetilde{\mathcal{W}}\mathbf{X}$. It is similar to the DWT in that both are linear filtering operations producing a set of time-dependent wavelet and scaling coefficients which are related to variations over a set of scales. The MODWT differs from the DWT in that is a highly redundant, nonortoghonal transform:

- The MODWT retains downsampled values at each level of the decomposition that would be otherwise discarded by the DWT. Thus, while the DWT restricts the sample size of the series to an integer multiple of 2^{J_0}, the MODWT is well defined for any sample size (with the consequent computational cost increment). Now, the wavelet and scaling coefficients ($\widetilde{\mathbf{W}}_j$ and $\widetilde{\mathbf{V}}_{J_0}$, respectively), and the details and smooth ($\widetilde{\mathcal{D}}_j$ and $\widetilde{\mathcal{S}}_{J_0}$, respectively) are N dimensional vectors.

- In contrast to the DWT, the details and smooth of the MODWT are shifted by a corresponding amount if the time series is shifted. Moreover, the MODWT details and smooths are easily lined up in an MRA with the original time series.

- The MODWT wavelet and scaling coefficients can be used to form an energy decomposition of $\mathbf{X}$. However, the MODWT is *not* an orthonormal transform, and this energy decomposition is not true for the MODWT details and smooths.

The time series $\mathbf{X}$ can be recovered from the MOWDT with a multiresolution analysis:

$$\mathbf{X} = \widetilde{\mathcal{W}}^T \widetilde{\mathbf{W}} = \sum_{j=1}^{J_0} \widetilde{\mathcal{W}}_j^T \widetilde{\mathbf{W}}_j + \widetilde{\mathcal{V}}_{J_0}^T \widetilde{\mathbf{V}}_{J_0} \equiv \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0} \tag{1}$$

Besides, the energy decomposition of $\mathbf{X}$ can now be defined *only* in terms of the wavelet and scaling coefficients:

$$\|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2 \tag{2}$$

This property will be detailed in the next section with the definition of the wavelet variance.

As an illustrative example, a MODWT is applied to the irradiance time series registered in the Milagro PV plant on the 13th of October 2009. A suitable graphical structure for the display of the MRA is the *horizon graph*. This kind of graphics comprises the vertical information (Y-axis) encoding differences in magnitude as differences in colour intensity, which we naturally perceive quantitatively (darker is greater and lighter is less) and encoding positive or negative values with different colour hues. Thus, with the technique of "small multiples" described in [17], a horizon plot allows spotting extraordinary behaviours and predominant patterns, view changes, analyse each of the series independently from the others and making comparisons between the series [18, 19].

For example, figure 1 uses blue for positive values and red for negative values. The irradiance fluctuation near the sunrise (around 8:00h) is clearly detected in the physical scales of 1m20s, 2m40s and 5m20s, the irradiance fluctuation before the midday (around 11:00h) is most important in physical scales of 10m40s and 21m20s, while the fluctuation at midday is mainly detected in physical scales of 20 and 40 seconds.

2.2. The wavelet variance and the spectral density function

The spectral analysis with the spectral density function, designed to work with stationary processes, cannot be directly used on nonstationary processes [20]. As an alternative, the wavelet variance, defined both for stationary and nonstationary processes with dth order stationary
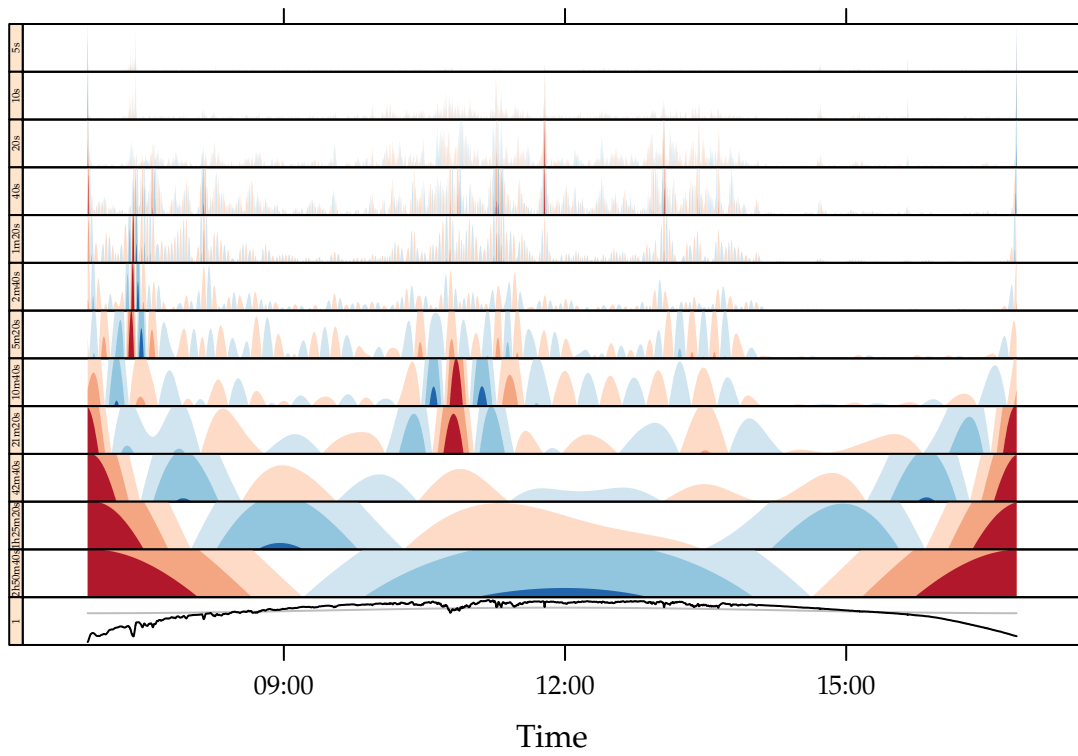
Figure 1: Horizon plot of the wavelet multiresolution analysis of the irradiance registered in the Milagro PV plant on the 13rd of October of 2009. Each row includes a different detail of the MRA while the last row shows the original irradiance time series with its smooth decomposition. The y-axis (not shown) encodes irradiance values.

backward differences, provides a consistent summary of the information contained in the spectral density function (SDF) on a octave band basis [15, 21]. This tool will be useful for the analysis of irradiance and power time series since they are nonstationary processes.

Using the energy decomposition of the MODWT (equation 2), the contribution of each scale to the variance of the time series X_t, $\sigma_X^2 = \|\mathbf{X}\|^2 - \overline{X}^2$ can be determined with[3] $\sigma_X^2 = \frac{1}{N} \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_{\mathbf{j}}^{\mathbf{X}}\|^2$, where $\|\widetilde{\mathbf{V}}_{\mathbf{J_0}}^{\mathbf{X}}\|^2 \simeq \overline{X}^2$.

Suppose that X_t is a stochastic process whose dth order backward difference, $Y_t = (1-B)^d X_t$, is a stationary process (where d is a nonnegative integer and B is the backward shift operator[4] defined by $BX_t = X_{t-1}$ and $B^k X_t = X_{t-k}$). Then, the result of the MODWT of X_t, if an adequate wavelet filter is used[5], is a set of coefficients time series, $\widetilde{W}_{j,t}$, which can be regarded as stationary processes whose SDF can be defined by:

$$S_j(f) = \widetilde{\mathcal{H}}_j^D(f) S_X(f) \tag{3}$$

where $\widetilde{\mathcal{H}}_j^D$ is the squared gain function associated with the wavelet filter at scale j [15]. If we denote with $\nu_{X,j}^2$ the wavelet variance of the scale τ_j, since the variance of a stationary process is equal to the integral of its SDF, then:

$$\nu_{X,j}^2 = \int_{-1/2}^{1/2} S_j(f) \mathrm{df} = \int_{-1/2}^{1/2} \widetilde{\mathcal{H}}_j^D S_X(f) \mathrm{df} \tag{4}$$

Therefore, the variance of $\mathbf{X}$ is

$$\sigma_X^2 = \sum_{j=1}^{J_0} \nu_{X,j}^2 \tag{5}$$

The wavelet variance of each scale represents the contribution to the SDF in the correspondent octave. The width of the octave of the scale τ_j is $1/(2^{j+1}\Delta t)$, where Δt is the sampling time of the signal, and the frequency band of this scale is $1/(2^{j+1}\Delta t) \leq f \leq 1/(2^j \Delta t)$. The average value of SDF over this interval is:

$$\overline{S_{X,j}} = 2^{j+1} \Delta t \int_{\frac{1}{2^{j+1}\Delta t}}^{\frac{1}{2^j \Delta t}} S_X(f) \mathrm{df} \tag{6}$$

Thus, the wavelet variance can be used as an estimator of the average value of the SDF [15, 21]:

$$\nu_{X,j}^2 = \frac{\overline{S_{X,j}}}{2^j \Delta t} \tag{7}$$

For example, we calculate and compare the wavelet variance of irradiance time series of several consecutive days of October 2009. Previously, a moving average of 5 seconds is applied. Moreover, the time series are scaled subtracting the daily mean and dividing with the STC

---

[3]The limits of this sum must be corrected to include only those coefficients which are not subject to to reflection conditions and then obtain an unbiased estimator of the variance. However, for ease of exposition, the equation we are showing includes the whole set of coefficients.

[4]Although B is commonly used to denote the direct component of irradiance, in the context of this investigation the risk of confusion is nonexistent, so we chose to follow the nomenclature conventions of the time series analysis literature.

[5]For example, a Daubechies wavelet filter whose width L fulfills the condition $L \geq 2d$

irradiance value. Finally, the wavelet variance is calculated for each day with the *Daubechies least asymmetric* LA(8) (or *symmlet* S(8)) filters[6] . The figure 2 shows the time evolution of the irradiance and power and the figure 3 displays their wavelet variances with their confidence intervals. For example, the higher fluctuation levels of the 14th day are converted to higher variance values over the whole set of scales.

2.3. Long memory process

Suppose now that X_t is an *stationary* process with a SDF denoted by S_X. X_t is an stationary long memory process if $S_X(f) \simeq C_S|f|^\alpha$, where C_S and α are constants such that $C_S > 0$ and $-1 < \alpha < 0$, with the approximation improving as f approaches zero. Therefore, the autocovariance sequence (ACVS) $s_{X,\tau}$ of this process is $s_{X,\tau} \simeq C_s\tau^\beta$ where $\beta = -\alpha - 1$. Standard time series models such as stationary autoregressive processes have ACVSs such that $s_{X,\tau} \simeq C\phi^\tau$. In both cases $s_{X,\tau} \to 0$ as $\tau \to \infty$, but the rate of decay toward zero is much slower for a long memory process, implying that current observations retain *memory* of the distant past [15, 22]. Under this model, the process has inertia and it is less likely to show abrupt changes. This description is consistent with the formulation of the probability of fluctuation level with a decaying exponential model published in [2].

When $\alpha \leq -1$, X_t can be interpreted as a nonstationary pure power law process that can be turned into a stationary process, Y_t, through an appropriate differencing operation, $Y_t = (1 - B)^d X_t$. It is interesting to note that wavelet variance plots can help to assess when a time series is adequately modelled as a nonstationary process with stationary backward differences. The approximation $S_X(f) \propto |f|^\alpha$ is equivalent to $\nu_X^2(\tau_j) \propto \tau_j^\beta$. When $\log(\nu_X^2(\tau_j))$ is plotted versus $\log(\tau_j)$ the exponents become slopes, with $\beta = -1 - \alpha \geq 0$ corresponding to nonstationarity.

The markedly positive slopes of the first two days of figure 3 for the lowest scales show that the irradiance is a nonstationary long memory process. The value of the slopes are different for each day due to their fluctuation behaviour. The value of β (α) is 2.16 (-3.16) for the 13th October and 1.77(-2.77) for 14th October, while it decreases (increases) to -0.017 (-0.98) for 15th October and 0.66 (-1.66) for 16th October.

Thus, the value of the β (or α) exponent may be a useful indicator of the fluctuation level in lower scales. Only one number could be enough to summarise the information of a whole day, and form the basis for a clustering analysis of a irradiance database or for prediction of fluctuation levels. Besides, there is a variety of models for nonstationary long memory processes such as the pure power law process or the fractional differenced process [15, 22], which could be useful for the simulation of irradiance time series.

2.4. Wavestrapping

Bootstrapping is the practice of estimating properties of an estimator by measuring those properties when sampling from an approximating distribution. One standard choice for an approximating distribution is the empirical distribution of the observed data. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of resamples of the observed dataset (and of equal size to the observed dataset), each of which is obtained by random sampling with replacement from the original dataset. The combination of the bootstrap with the WT is called *wavestrapping* [23]. The idea behind wavestrapping is that for stationary processes, the DWT acts as a decorrelating transform for time series, so the wavelet coefficients can be regarded as

---

[6] The length of a LA(8) filter is 8, so it is able to handle adequately a nonstationary process with 4th order stationary backward differences
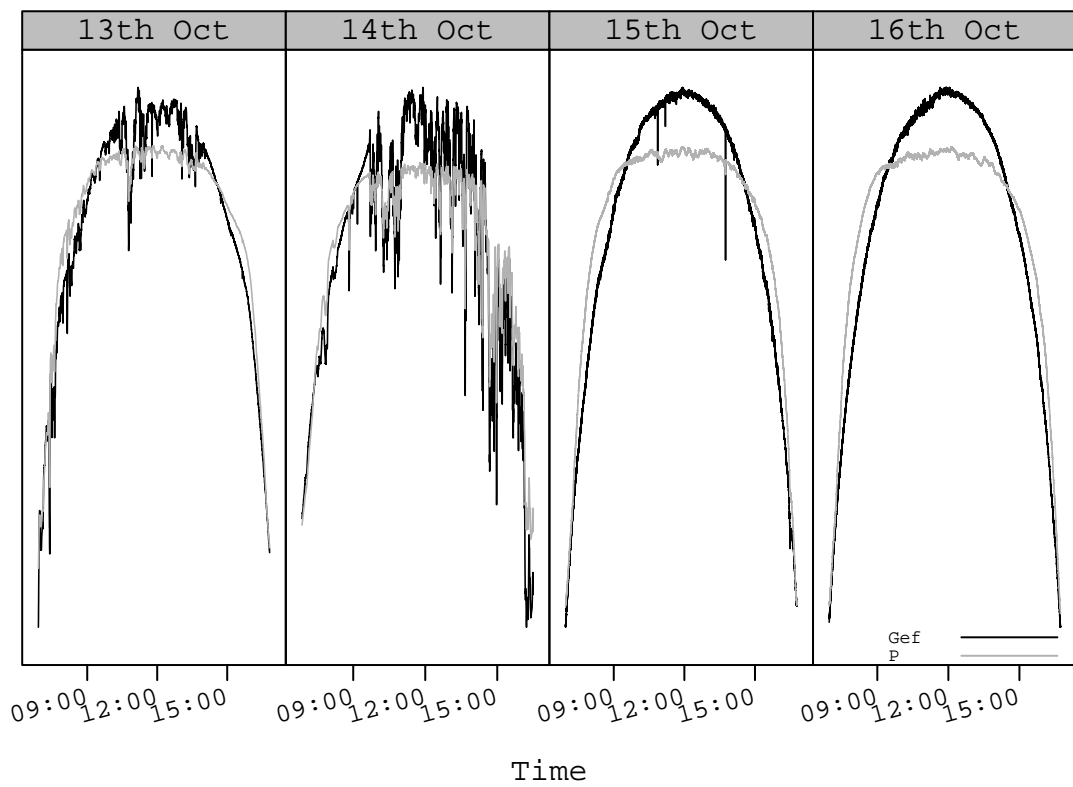
Figure 2: Time evolution of the effective irradiance and power output of several consecutive days of October 2009. Both time series have been centred and scaled for ease of comparison (y-axis is not shown).
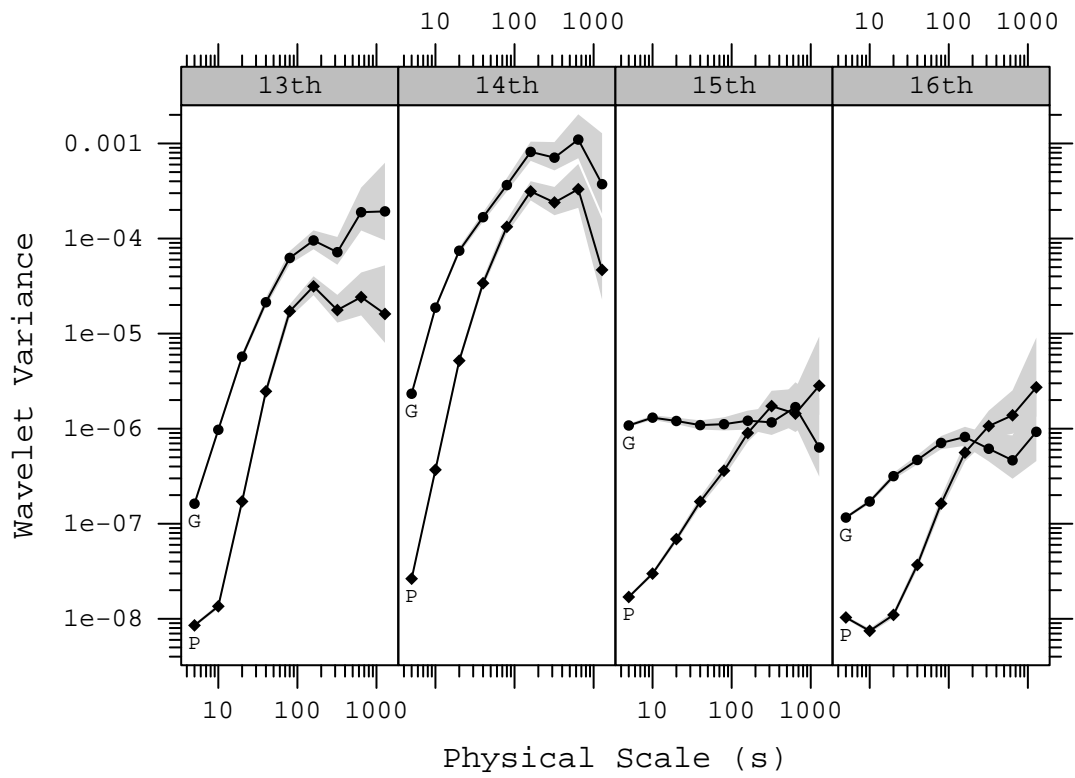
Figure 3: Wavelet variance (and confidence intervals) of the effective irradiance and power output of the Milagro PV plant on several days of October 2009.

uncorrelated and suitable for bootstrapping techniques. The clear advantage of this method is that it allows the synthesis of a time series in a non-parametric way if stationary criteria apply.

Although the irradiance is a nonstationary process, the next sections will propose how to extract information with a quasi-stationary behaviour from the original time series. This quasi-stationary process will be the source for the wavestrapping methods.

3. Synthesis of temporal series of irradiance

The solar irradiance incident on a surface outside the terrestrial atmosphere, denoted as the extra-atmospheric irradiance, can be assumed as a deterministic time process and can be estimated by means of geometric considerations. However, the complex process of reflection, attenuation and dispersion due to the components of the atmosphere transforms the extra-atmospheric irradiance into an stochastic process, the global irradiance incident on a surface.

3.1. Irradiance on the horizontal plane

The global irradiance on the horizontal plane, $G(0)$ is commonly described as a deterministic process (the extra-atmospheric irradiance on the horizontal plane, $B_o(0)$) modified with a multiplicative stochastic effect (the well-known clearness index, k_t) summarising the behaviour of the atmosphere. For the time series analysis, it is recommended to remove the seasonal and trend component of the series [24]. This is one of the advantages of using the clearness index, since the diurnal trend due to the terrestrial movement is effectively removed. Although the *instantaneous* clearness index is not strictly stationary, the departure from stationarity is small enough to use the wavestrapping technique.

From the data collected at the Milagro PV plant, the clearness index is calculated and a new time series of clearness index is simulated with the wavestrapping procedure. Finally, a new time series of global irradiance on the horizontal plane is obtained multiplying the simulated clearness index with the calculated extra-atmospheric irradiance. The result is shown in figure 4, where the original and the simulated irradiance time series corresponding to the 13/10/2009 are displayed together.

The main objective of this generator is to reproduce the fluctuations of the original time series. This result can be assessed comparing the cumulative distribution function (CDF) of the fluctuations of both time series. This fluctuation has been calculated with the first-order backward differences, $(1 - B)G_t(0) = G_t(0) - G_{t-1}(0)$. Figure 5 shows that the wavestrapped series contains a higher probability of fluctuation, but this method imitates accurately the fluctuation behaviour of the original series.

This same procedure has been tested with different days obtaining similar satisfactory results.

3.2. Irradiance on the inclined plane

The same conceptual approach could be suitable for the global irradiance incident on the plane of the PV generator, $G_{ef}(I)$. Again, the extra-atmospheric irradiance on the inclined plane, $B_o(I)$ can be calculated with only geometric considerations. A new clearness index on the inclined plane is defined as the ratio $k_t(I) = G_{ef}(I)/B_o(I)$. However, this clearness index does exhibit a trend possibly due to the anisotropic behaviour of the sky (the PV generator is not receiving energy from the whole sky). When its WT is resampled with the wavestrapping technique and the series is reconstructed, this diurnal trend is severely transformed. The global irradiance obtained multiplying this simulated clearness index by the extra-atmospheric irradiance is now incoherent with the original one.

Another approach is to take advantage of the filtering process of the WT, since the diurnal trend can be safely removed with a multiresolution analysis including the details from a subset
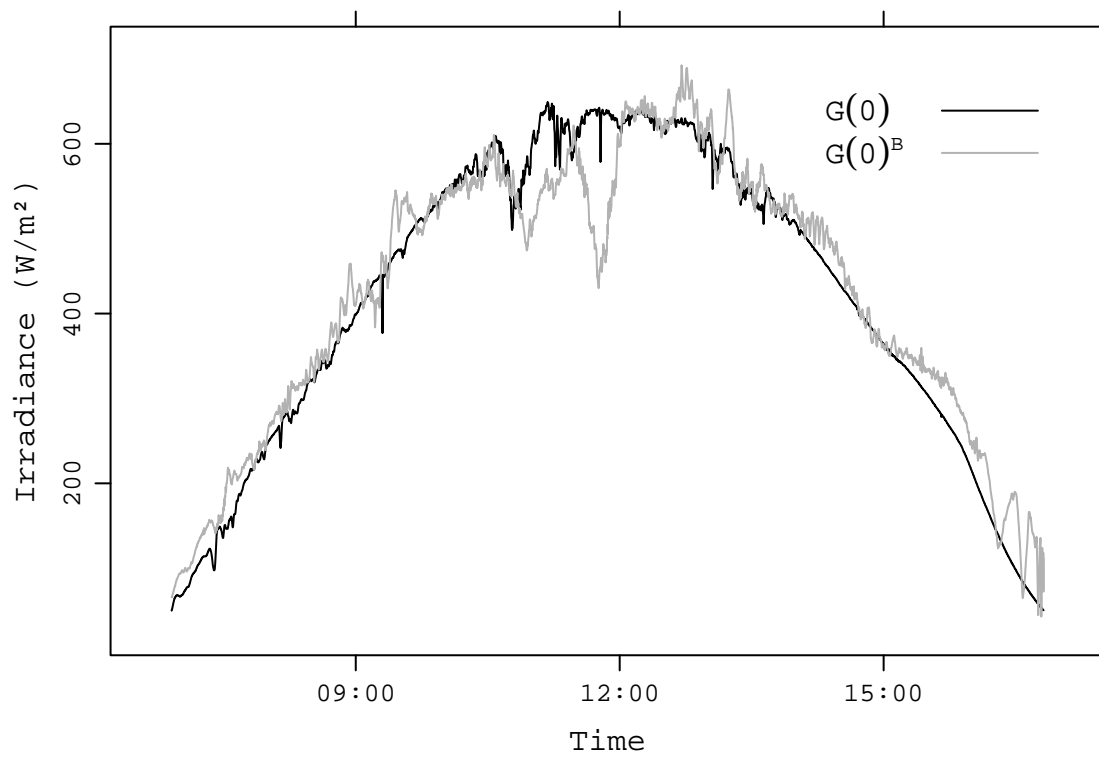
Figure 4: Simulation of a time series of horizontal global irradiance with wavestrapping techniques applied to the clearness index for the 13th October 2009. Both the original $G(0)$ and the wavestrapped $G(0)^B$ time series are shown.
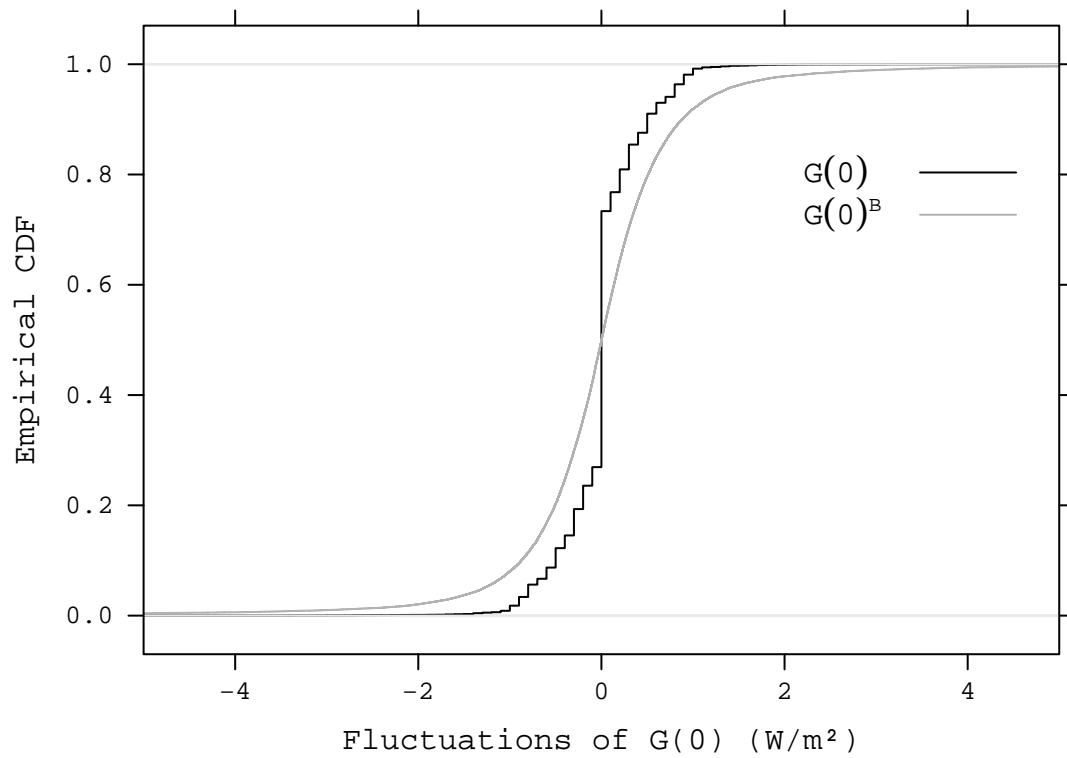
Figure 5: Comparative between the experimental cumulative distribution function of the fluctuations of the original time series, $G(0)$, and the wavestrapped one, $G(0)^B$.

of scales. For example, for a sampling time of 1 second, this trend is adequately captured if the scales 1 to 9 are taken away from the MRA. Therefore, the trend is:

$$G_{ef}^T = \sum_{i=10}^{J_0} \mathcal{D}_i + \mathcal{S} \tag{8}$$

and the detrended signal is

$$G_{ef}^D = \sum_{i=1}^{9} \mathcal{D}_i \tag{9}$$

This detrended signal is approximately stationary and it is now possible to obtain a new wavestrapped version of it, $G_{ef}^{D,B}$. This wavestrapped signal is added to the trend to get a simulated irradiance time series containing the same fluctuation behaviour as the original one.

$$G_{ef}^B = G_{ef}^T + G_{ef}^{D,B} \tag{10}$$

The evolution of the irradiance on the inclined plane with both the trend and the detrended signal is included in the figure 6. Moreover, the wavestrapped simulation of the detrended signal and the resulting simulated irradiance time series are shown for comparison.

Once again, the CDF evaluates the performance of this procedure for reproducing the fluctuation behaviour. The satisfactory results are shown in figure 7.

This procedure is useful for the simulation of irradiance series incident in a group of distant PV plants. It is likely that the geographic relation between them and the climatic conditions will determine the number of scales which are to wavestrapped. For example, if the plants are distant from one another, the sky appearance is likely to vary also in higher scales, and therefore the subset of scales to be wavestrapped should be larger than for near PV plants. The analysis of a database which includes several PV plants will help to adequately choose this subset. This analysis is postponed for future works.

3.3. Power output of a group of PV systems

The figure 3 shows the wavelet variance of the power output of the Milagro PV plant (previously normalised with the power output of the generator). Again, the power time series behaves as a long memory process for the first set of scales. The slope of the curve of the power output is higher than the curve of irradiance, confirming the low-pass filter behaviour of the system: the value of β (α) is 2.94 (-3.94) for the 13th October and 3.11(-4.11) for the 14th October, while it decreases (increases) to 1.13 (-2.13) for the 15th October and 1.02 (-2.02) for the 16th October.

A possible approach is to analyse the attenuation in the frequency domain through the SDFs of the input and output signals. If both processes were stationary their SDFs would be directly related through the squared gain of the transfer function of the system. However, the irradiance and power are nonstationary time processes, so this relation is not valid unless some conditions are fulfilled [20].

On the other hand, since their dth order backward difference are stationary processes, their wavelet variances are well defined (section 2.2). The comparison between them represents the behaviour of the PV plant on a scale basis. If $\overline{S_{G,j}}(f)$ and $\overline{S_{P,j}}(f)$ stand for the average SDFs at the j scale of the irradiance and power time series, respectively, while $\nu_G^2(\tau_j)$ and $\nu_P^2(\tau_j)$ denote the wavelet variance of scale j of these processes, then using equation (7):

$$\frac{\nu_{P,j}^2}{\nu_{G,j}^2} = \frac{\overline{S_{P,j}}}{\overline{S_{G,j}}} \tag{11}$$
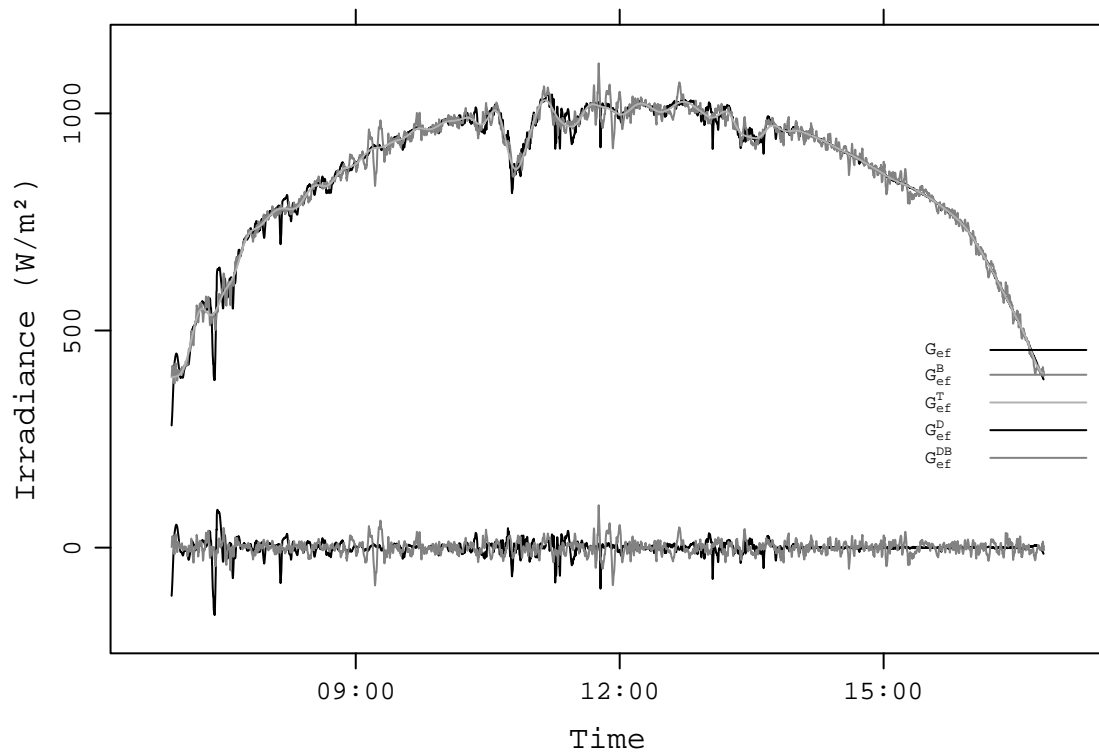
14

Figure 6: Evolution of the irradiance on the inclined plane, G_{ef} compared with the trend decomposition, G_{ef}^T, and the wavestrapped time series, G_{ef}^B. Besides, the detrended signal, G_{ef}^D, and its wavestrapped version, G_{ef}^{DB}, are shown at the bottom. These time series belong to the 13th October 2009.
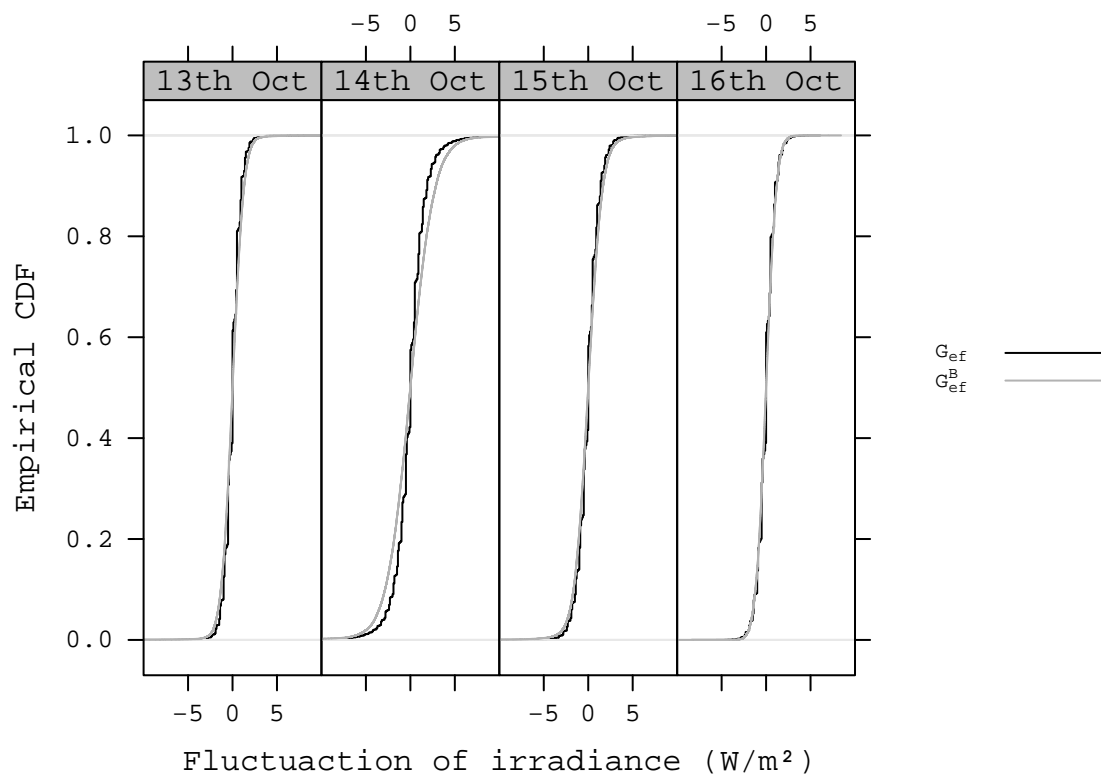
Figure 7: Comparative between the empirical cumulative distribution function of original fluctuations of the effective irradiance time series, G_{ef}, and the wavestrapped simulation, G_{ef}^{B} for several consecutive days of October 2009.

The figure 8 shows this ratio for the same four consecutive days of figure 3. The plant behaves as a low-pass filter with a similar slope during the lower scales for the four days. On higher scales ($\tau_j \Delta t \geq 160s$), the response is almost flat during the days whose fluctuation level is higher, and steep during the days with a low fluctuation level. It is important to note that, due to the confidence intervals over these scales (figure 3), it is advisable to be cautious about the comparisons on these octaves.

A provisional explanation of this behaviour is proposed. The output of a photovoltaic system depends on the incident irradiance through the equation:

$$P(t) = \eta_s \cdot G_{ef}(t) \tag{12}$$

where $P(t)$ is the power at the DC/AC inverter power, η_s is the system efficiency and $G_{ef}(t)$ is the effective irradiance incident on the PV generator plane. This equation works under the assumption that the incident irradiance is the same simultaneously over the whole PV generator plane. As the reference [2] has shown, this supposition is not true for large PV generators. A plausible approximation is to model the behaviour of the plant around two assumptions:

- The response of the PV system is invariable with time. In other words, the efficiency η_s has no time dependence[7].

- The irradiance incident on a point of the system in a instant is received with approximately no changes in another point located at a certain distance after a time interval. The length of this interval depends on the size of the plant and on the meteorological conditions, (mainly on the wind speed [3]).

Under such ideal circumstances, the PV plant carries out a moving average of the irradiance time series, and thus behaves as a low-pass filter. The impulse response, $h(t)$, of the equivalent filter, is characterised by its amplitude and length. The amplitude is the value of the system efficiency and its length, T, depends on, as previously stated, the size of the plant and the wind conditions. The output power is the convolution of this response with the irradiance time series.

$$P(t) = h(t) * G_{ef}(t) \tag{13}$$

The transfer function of an ideal low-pass filter is the $\text{sinc}(x)$ function, so the Fourier transform of $h(t)$ is:

$$H(f) = \eta_s \cdot \frac{\sin(\pi f T)}{\pi f} \tag{14}$$

Therefore, $\mathcal{P}(f) = \mathcal{H}(f)\mathcal{G}_{ef}(f)$, where $\mathcal{P}(f)$ and $\mathcal{G}(f)$ stand for the Fourier transforms of the power and irradiance time series, respectively.

The time length of $h(t)$ is the time spent by a irradiance profile to run across the PV plant in a longitudinal or transversal direction. If the plant is supposed to be a square whose side is L, and a constant wind speed v, this time is $T = L/v$. Thus, the larger the size of the plant (or the lower the wind speed), the wider the length of the time window (and the narrower the central lobe of the transfer function). Under this model, the irradiance fluctuations are likely to be attenuated when the combination between plant size and wind speed results in a time window wide enough (or, equivalently, in a narrow transfer function).

---

[7]This assumption implies that the efficiencies of the PV generator and inverter are independent of the irradiance level. Although this approximation is not acceptable for energy calculations, it is admissible for a simplified explanation of the fluctuations behaviour.
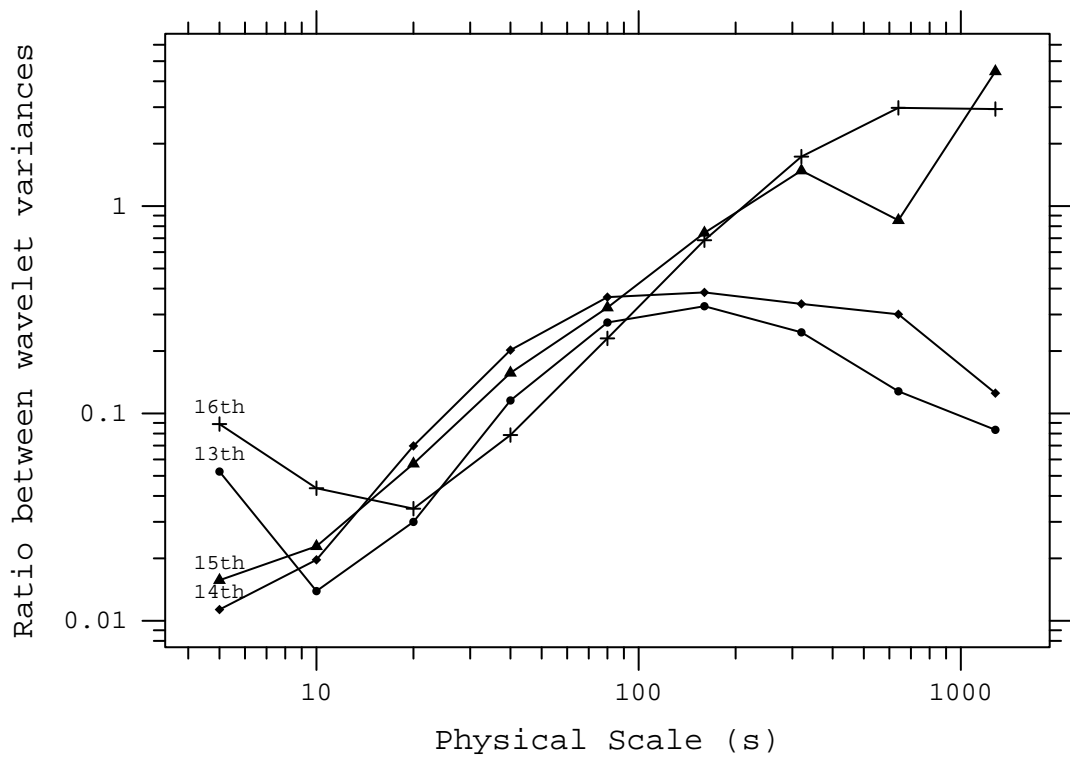
Figure 8: Ratio between the wavelet variances of power output and irradiance time series of the Milagro PV plant during several days of October 2009.

As an example, for the 14th October 2009 the ratio between the variances of the power and irradiance (equation (11)) has been calculated and normalised. This ratio is compared with moving averages of different widths (35 to 55 seconds), as it is shown in figure 9. From this figure, it is clear that the intrinsic variations of the PV plant, and the departure of its real behaviour from our ideal supposition, smooth the impulse response of the system and alter the ideal transfer function.

4. Conclusions

The irradiance fluctuations and the subsequent variability of the power output of a PV system have been analysed with some mathematical tools based on the wavelet transform. The most important findings and proposals are:

- The irradiance and power time series are nonstationary process whose behaviour resembles a long memory process.

- The long memory spectral exponent α is a useful indicator of the fluctuation level of a irradiance time series.

- A time series of global irradiance on the horizontal plane can be simulated by means of the wavestrapping technique on the clearness index. The fluctuation behaviour of this simulated time series correctly resembles the original series.

- A time series of global irradiance on the inclined plane can be simulated with the wavestrapping procedure applied over a signal previously detrended by a partial reconstruction with a wavelet multiresolution analysis. Once again, the fluctuation behaviour of this simulated time series is correct. This procedure is proposed as a suitable tool for the simulation of irradiance incident over a group of distant PV plants.

- A wavelet variance analysis and the long memory spectral exponent show that a PV plant behaves as a low-pass filter.

From this knowledge, there are several steps to follow:

- The long memory spectral exponent, as a summary of the fluctuation behaviour of a irradiance or power time series, could serve as an index for the classification of days according to its fluctuation level. Therefore, a irradiance/power database can be sorted with this index. Moreover, clustering techniques can take advantage of this summary, while the results of the classification methods may be useful for developing prediction procedures.

- The behaviour of a PV plant is determined by its size and the meteorology conditions of its location. The wavelet analysis of both irradiance and wind speed may help to obtain an accurate model of this performance in order to obtain power fluctuation levels from the irradiance fluctuation levels.

- Geostatistics and wavelet time series analysis may serve as a useful combination for the study of irradiance and power series of a group of PV plants and inside a large PV plant.
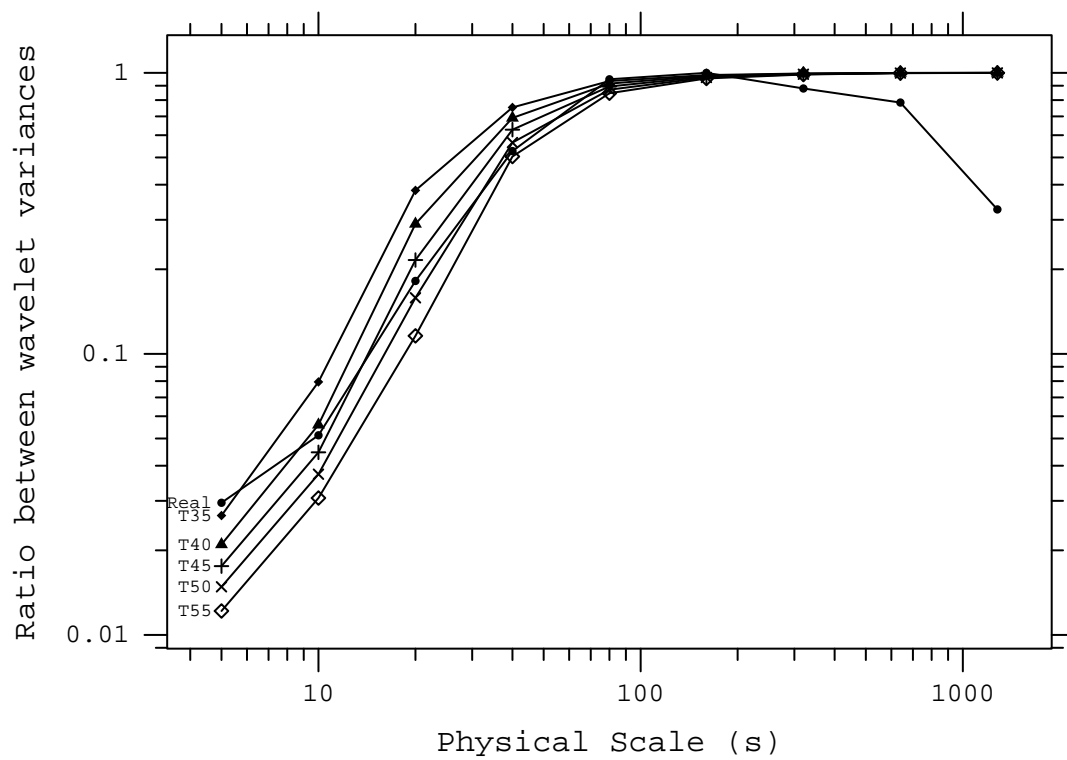
Figure 9: Comparison between the ratio of variances of power and irradiance ("Real") and the ratio of variances of moving averages with different widths (35 s to 55 s) and the original irradiance time series.

Acknowledgements

Software tools

The calculation procedures has been implemented with R, a free software environment for statistical computing and graphics [25]. The package `wmtsa` [26] includes the tools for the wavelet analysis, and the packages `solaR` [27] and `zoo` [28] provide with the functions for the solar and time series calculations, respectively. Finally the packages `lattice` [29] and `latticeExtra` [30] include a complete set of graphical utilities for the production of the set of figures of this paper.

The code implemented for this paper is available at http://procomun.wordpress.com/documentos/articulos/.

References

[1] E. Caamaño, H. Laukamp, M. Jantsch, T. Erge, J. Thornycroft, H. de Moor, S. Cobben, D. Suna, B. Gaiddon, Interaction between photovoltaic distributed generation and electricity networks, Progress in Photovoltaics: Research and Applications doi:10.1002/pip.845.

[2] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira, E. Izco, Power output fluctuations in large scale PV plants: One year observations with 1 second resolution and a derived analytic model., Progress in Photovoltaics.

[3] A. Woyte, R. Belmans, J. Nijs, Fluctuations in instantaneous clearness index: Analysis and statistics, Solar Energy 81 (2007) 195–206.

[4] J. Ramachandran, N. M. Pearsall, G. A. Putrus, Reduction in solar radiation fluctuation by spatial smoothing effect, in: 19th European Photovoltaic Solar Energy Conference, 2004, pp. 2900–2903.

[5] T. Tomson, G. Tamm, Short-term variability of solar radiation, Solar Energy 80 (2006) 600?606. doi:doi:10.1016/j.solener.2005.03.2009.

[6] R. J. Aguiar, M. Collares-Pereira, J. P. Conde, Simple procedure for generating sequences of daily radiation values using a library of markov transition matrices, Solar Energy 40 (3) (1998) 269–279.

[7] R. J. Aguiar, M. Collares-Pereira, Statistical properties of hourly global radiation, Solar Energy 48 (3) (1992) 157–167.

[8] K.-M. Lau, H. Weng, Climate signal detection using wavelet transform: How to make a time series sing, Bulletin of the American Meteorological Society 76 (12) (1995) 2391–2402.

[9] C. Torrence, G. P. Compo, A practical guide to wavelet analysis, Bulletin of the American Meteorological Society 79 (1) (1998) 61–78.

[10] M. O. Domingues, O. Mendes, A. M. da Costa, On wavelet techniques in atmospheric sciences, Fundamentals of Space Environment Science 35 (5) (2005) 831–842. `doi:10.1016/j.asr.2005.02.097`.

[11] D. Labat, Recent advances in wavelet analyses: Part 1. a review of concepts, Journal of Hydrology 314 (1-4) (2005) 275 – 288. `doi:DOI:10.1016/j.jhydrol.2005.04.003`.

[12] H. S. Oh, C. M. Ammann, P. Naveau, D. Nychka, B. L. Otto-Bliesner, Multi-resolution time series analysis applied to solar irradiance and climate reconstructions, Journal of Atmospheric and Solar-terrestrial Physics 65 (2) (2003) 191–201.

[13] N. Kawasaki, T. Oozeki, K. Otani, K. Kurokawa, An evaluation method of the fluctuation characteristics of photovoltaic systems by using frequency analysis, Solar Energy Materials & Solar Cells 90. `doi:10.1016/j.solmat.2006.02.034`.

[14] A. Mellit, M. Benghanem, S. A. Kalogirou, An adaptive wavelet-network model for forecasting daily total solar-radiation, Applied Energy 83.

[15] D. Percival, A. T. Walden, Wavelet Methods for Time Series Analysis, Cambridge University Press, 2006.

[16] C. R. Cornish, C. S. Bretherton, D. B. Percival, Maximal overlap wavelet statistical analysis with application to atmospheric turbulence, Boundary-layer Meteorology 119 (2) (2006) 339–374. `doi:10.1007/s10546-005-9011-y`.
URL `http://staff.washington.edu/dbp/PDFFILES/WaveTechAnalAtmosTurb.pdf`

[17] E. R. Tufte, The visual display of quantitative information, Graphic Press, 2001.

[18] J. Heer, N. Kong, M. Agrawala, Sizing the horizon: The effects of chart size and layering on the graphical perception of time series visualizations, in: ACM Human Factors in Computing Systems (CHI), 2009, pp. 1303 – 1312.
URL `http://vis.berkeley.edu/papers/horizon/2009-TimeSeries-CHI.pdf`

[19] S. Few, Time on the horizon, Tech. rep., Perceptual Edge (2008).
URL `http://www.perceptualedge.com/articles/visual_business_intelligence/time_on_the_horizon.pdf`

[20] M. B. Priestley, Evolutionary spectra and non-stationary processes, Journal of the Royal Statistical Society. Series B (Methodological) 27 (1965) 204–237.

[21] D. P. Percival, On estimation of the wavelet variance, Biometrika 82 (3) (1995) 619–631.
URL `http://staff.washington.edu/dbp/PDFFILES/wavevar.pdf`

[22] B. Whitcher, S. D. Byers, P. Guttorp, D. B. Percival, Testing for homogeneity of variance in time series: Long memory, wavelets, and the nile river, Water Resources Research 38 (5). `doi:10.1029/2001WR000509`.
URL `http://staff.washington.edu/dbp/PDFFILES/wrr.pdf`

[23] D. Percival, S. Sardy, A. Davison, Nonlinear and nonstationary signal processing, Cambridge University Press, 2000, Ch. Wavestrapping Time Series: Adaptive Wavelet-Based Bootstrapping, pp. 442–71.
URL `http://www.unige.ch/math/folks/sardy/Papers/wavestrap.pdf`

[24] W. S. Cleveland, Visualizing Data, Hobart Press, Summit, New Jersey, U.S.A., 1993.

[25] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (2010).
URL http://www.R-project.org

[26] W. Constantine, D. Percival, wmtsa: Insightful Wavelet Methods for Time Series Analysis, r package version 1.0-5 (2010).
URL http://CRAN.R-project.org/package=wmtsa

[27] O. Perpiñán, solar: Solar photovoltaic systems, R package version 0.14 (2010).
URL http://CRAN.R-project.org/package=solaR

[28] A. Zeileis, G. Grothendieck, zoo: S3 infrastructure for regular and irregular time series, Journal of Statistical Software 14 (6) (2005) 1–27.
URL http://www.jstatsoft.org/v14/i06/

[29] D. Sarkar, lattice: Lattice graphics, R package version 0.18-3 (2010).
URL http://CRAN.R-project.org/package=lattice

[30] D. Sarkar, F. Andrews, latticeExtra: Extra Graphical Utilities Based on Lattice, r package version 0.6-12/r148 (2010).
URL http://R-Forge.R-project.org/projects/latticeextra/