Characterizing Temporal Variability in Measurements of Surface Solar Radiation and its Dependence on Climate
Marc Bengulescu, Philippe Blanc, Lucien Wald

To cite this version:

HAL Id: hal-01406104
https://hal-mines-paristech.archives-ouvertes.fr/hal-01406104
Submitted on 30 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Characterizing temporal variability in measurements of surface solar radiation and its dependence on climate

Marc Bengulescu*, Philippe Blanc, Lucien Wald

MINES ParisTech, PSL Research University, Centre for Observation, Impacts, Energy, CS 10207 - 06904 Sophia Antipolis, France

Abstract

The intrinsic temporal scales of the variability of the surface solar radiation are addressed by means of the empirical mode decomposition. High quality measurements of the solar radiation impinging on a horizontal plane at ground level, from different BSRN ground stations, are analysed. By first extracting all the embedded oscillations that share a common local time-scale, followed by Hilbert spectral analysis, the characteristic scales of variability, along with the fluctuations in the intensity of the pyranometric signal, are revealed. It is shown that data from stations with different local climates share some common features, most notably a high-frequency plateau of variability whose amplitude is found to be modulated by the seasonal cycle. The study has possible implications on the modelling and the forecast of the surface solar radiation, at different local time-scales.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Peer-review under responsibility of the organizing committee of the General Assembly of the European Geosciences Union (EGU)

Keywords: solar radiation; temporal variability; climate; empirical mode decomposition; Hilbert spectrum; BSRN

1. Introduction

The surface solar irradiance (SSI) has been identified by the Global Climate Observing System program as one of the Essential Climate Variables that is of utmost importance in characterizing Earth climate, thus better knowledge of the SSI represents a key factor in decision making in both science and policy circles [1]. Furthermore, solar irradiance is the main driver behind the weather and climate systems on the planet, with a dynamic range varying from seconds to millions of years [2]. Despite this wide span of the associated time-scales, most of the studies that have raised the scientific understanding with respect to the SSI have primarily focused on either global averages and long-term trends [3], or have only addressed the short-term, or high-frequency part of its temporal variability [4].

In light of this need to understand the variability of the SSI field at different time horizons, an appropriate data processing technique for extracting the intrinsic modes of variability at vastly different temporal scales is required. Added to this, the nonlinear and nonstationary nature of the measured solar radiation data [5] further restricts the pool

* Corresponding author. Tel.: +33-493-678-947 ; fax: +33-493-678-908.
E-mail address: marc.bengulescu@mines-paristech.fr
of signal processing candidates to those fit to handle such time-series, without making any beforehand assumptions about the nature of the signal. It has been proposed that the Hilbert-Huang Transform (HHT), an adaptive, data-driven analysis method is a suitable technique for analysing such solar radiation datasets [6–8]. Here, this point of view is also adopted, therefore the inner workings and the rationale behind the choice of this particular data processing method will be only briefly reiterated in a dedicated subsection.

The focus of the present work thus lies on extracting and analysing the intrinsic modes of the temporal variability of the SSI, as identified by the HHT. By scrutinizing long-term (i.e. 10+ years) daily measurements of the solar radiation impinging on the ground, at different geographical locations, it is shown that the HHT is able to identify the low frequency variability components, i.e the multi-year trend and seasonal variation, and multiple high-frequency intrinsic variability modes. Although these short-timescale features appear to be distinct from the seasonal cycle, a correlation can nevertheless be found between their amplitude envelopes and the seasonal cycle, indicative of non-linear multiplicative cross-scale coupling [9]. Based on these characteristics, the difference in local climatic conditions of the measuring stations can be inferred. Possible implications of the findings on the modelling and forecast of the SSI are then discussed.

The study is organised as follows. In section 2 the data and the methods used will be described. Sub-section 2.1 discusses the data sources and the preprocessing, while 2.2 will address the analysis method. Section 3 will present the results obtained, with the discussion thereof being deferred to section 4. Lastly, acknowledgements and a bibliographical list conclude the study.

2. Data and methods

2.1. Data sources and pre-processing

The data analysed in this study consist of time-series of high-quality SSI ground measurements at two distinct geographical locations: one located in Carpentras, France (CAR – 44.083° N, 5.059° E), and the other in Payerne, Switzerland (PAY – 46.815° N, 6.944° E). For both stations the data span a ten year period, from January 1st, 2001 to December 31st, 2010 (figure 1). The measuring stations belong to the Baseline Surface Radiation Network (BSRN), a worldwide radiometric network providing accurate readings of the SSI at high temporal resolution [10]. The climate at the CAR station is of a humid subtropical, Mediterranean type (Köppen-Geiger: Csa), while the one at PAY experiences a marine west coast, mild type of climate (Köppen-Geiger: Cfb) [11]. Both stations are set in rural areas, with the surroundings having a cultivated surface type, and local hilly topography [12]. Lastly, since the stations are located in Europe, any reference to seasons and seasonal phenomena shall be understood as occurring in the northern hemisphere.

For the scope of the study, daily means of the global SSI received on a horizontal plane (GHI) were employed. The original time-series of the GHI, having a one-minute integration time, have been quality checked according to [13]. Daily means of GHI were then calculated only if more than 80% of the data during daylight were valid. Isolated missing values in the daily means series were then completed by applying linear interpolation to the daily clearness index.

2.2. Adaptive data analysis

Among the different techniques commonly used in geophysical signal processing, as for exampled those reviewed in [14], a suitable candidate for the study at hand was found to be the Hilbert-Huang Transform (HHT), a completely data-driven and adaptive signal analysis technique [15]. The HHT consists of two steps.

First, by means of the Empirical Mode Decomposition (EMD), oscillations that present a common local time-scale embedded within the signal are extracted. These oscillations, also called Intrinsic Mode Functions (IMFs), can be thought of as being amplitude modulation - frequency modulation (AM-FM) components of the data that have a zero average and a well behaved Hilbert transform. Owing to the adaptive nature of the EMD, the IMFs represent the data-driven, a posteriori chosen basis of the decomposition [16]. By contrast, other mathematical transforms employ basis functions that are fixed in advance, such as the trigonometric functions in the case of the Fourier transform or the wavelet functions of the eponymous transform. This adaptive approach of the decomposition not only endows the
method with the capability of dealing with data issued from the non-linear interaction of physical processes, often also found under the influence of non-stationary external forcings, but it also ensures that the extracted components carry physical meaning, and that the influence of method-inherent mathematical artefacts on the rendered picture of temporal variability is kept to a minimum [17].

Once the empirical mode decomposition is completed, the second step of the HHT is to perform Hilbert spectral analysis. Each IMF and its Hilbert transform are used to construct the complex-valued analytic signal, from which an amplitude modulation - frequency modulation (AM-FM) model of the data is obtained. Thus, under this model any IMF, say $x(t)$, can be written as $x(t) = \Re\{a(t) \cdot e^{i \int_{\tau}^{t} \omega(\tau) d\tau}\}$, with $a(t)$ instantaneous amplitude and $\omega(t)$ instantaneous frequency. From this representation the Hilbert energy spectrum of the data can be computed, basically an energy density distribution overlaid on the time-frequency space [18].

In this study, a recent iteration of the algorithm is employed, the improved complete ensemble EMD [19], as it allows for an exact decomposition of the data and is more robust with respect to noise. Furthermore, a fast EMD routine, described in [20], has been used to decrease the computation time.

3. Results

The IMFs obtained for the CAR time-series are presented in figure 2 (top panels, IMF1...8, numbers indicate extraction order from the highest frequencies to the lowest). The IMFs for PAY share similar characteristics, the only notable difference is that the CAR time-series is decomposed into 8 IMFs, while the PAY data has 9 components. Owing to the EMD being a time-domain only decomposition, the IMFs have the same temporal support as the original data and, by construction, zero average and symmetrical upper and lower envelopes. As the decomposition progresses, the local time-scale of the IMFs increases, which is to say that the period of the intrinsic oscillations is getting larger and larger within each newly extracted component. A graphical support of this is provided in the lower panel of figure 2, where the power spectral density (PSD) and an estimate of the mean period, expressed in days, of each component are plotted.

The PSD reveals that for the first five components the spectral shapes are similar in form, and the median period seems to roughly double with increasing IMF number, following an approximately dyadic sequence: $3.1 \rightarrow 7.4 \rightarrow 15.7 \rightarrow 32.0 \rightarrow 54.9$. Next, IMF6 can be readily identified by its median period of 364.9 days with the seasonal cycle. It also accounts for the most prominent visual feature in the original time-series (top panel, CAR, of figure 1), its peaks and troughs denoting summer and winter, respectively. Another feature of this sixth IMF is its sharp peak of relatively high power, as opposed to the bell curve shape and the frequency spread of the previous five components. By contrast, no distinct visual features stand out in the PSD of IMF7 and IMF8, with median Fourier periods of 655.5 and 1370.5 days, respectively; nevertheless their time-domain representations reveal that they have very low amplitudes. Lastly, also noteworthy is that the first five modes seems to exhibit amplitude excursions in accordance with IMF6, i.e. their amplitudes are greater during summer than during winter.

After having extracted all the IMFs, all that is left of the original time-series is a residual, or EMD trend, that cannot be considered an oscillation at the time span of the original data (10 years). Stated otherwise, the trend can be thought of as an offset onto which all the zero mean embedded oscillatory components are riding, or as a low-pass filtered approximation of the original data [21]. This can be clearly seen in figure 3, where the trend for CAR, plotted in blue, has a non-zero mean and is devoid of the oscillations of the IMFs depicted in in figure 2. To graphically
convey this, the yearly mean of the SSI, found in [22], is plotted in red alongside the EMD trend, in figure 3, whence it can be seen that the two curves follow each other closely, with the EMD trend taking on slightly higher values.

An alternative illustration of the spectral contents of the data, in terms of its Hilbert AM-FM decomposition, is presented in figure 4. Using a box plot, the frequency distribution within each IMF is conveyed, giving an indication of the local time-scale captured by the component. The first five IMFs have closely matching median periods, that once again seem to follow a dyadic sequence for both CAR and PAY (in parantheses): 3.2 (3.0) → 7.0 → 14.1 (14.6) → 26.8 (27.4) → 55.7 (53.4). Not only are the medians of these modes in close agreement for the two time-series, but both in terms of the interquartile range (IQR), and in terms of total range, the first five modes of PAY exhibit the same frequency variability as those for CAR. The sixth IMF for both series can clearly be associated, in terms of their median periods (364.3 for CAR and 356.2 for PAY), with the seasonal cycle given by the revolution of the Earth around the Sun. For the PAY data, the IQR and especially the range of this sixth mode are significantly greater than for CAR, overlapping part of the support of the previous IMF. This is an illustration of the mode mixing phenomenon that may arise in the EMD, i.e. the coexistence of different modes in the same IMF, principally caused by signal intermittence and noise [17]. However, this box plot representation is incomplete, as it only accounts for the frequency spread of the modes while completely ignoring amplitude, or power. It will later be shown that the spectral part of IMF6 which overlaps IMF5 has diminished power, thus this phenomenon does not hinder the validity of the analysis.

As a comparison, for the CAR high-frequency IMFs 1...5 most of the power in the Fourier PSD from figure 2 is assigned to a frequency band that closely corresponds to the Hilbert range, which do not extend beyond 100 days in figure 4. The same is also true for IMF6, whose Hilbert frequency distribution is concentrated in a very narrow range, that would correspond to the sharp peak in the PSD of the same component. Hence, while the Fourier representations of the IMFs span the whole frequency scale, the Hilbert frequency distributions have more compact support.

Coming back to the discussion of the IMFs in figure 4, it is found that after all the components with yearly and sub-yearly time-scales have been extracted the rest of the data, trend notwithstanding, is assigned into minimally overlapping (within the same dataset) modes in a band that spans the spectrum beginning at the one year mark. The
CAR data has a slightly longer spectral range, from one to six years, with a break at about three years that divides it into two modes with median periods of 469.5 days for IMF7 and 1305.1 days for IMF8. The PAY time-series features components with narrower spectral support, namely two IMFs that cover the one to two and a half years band having median frequencies of 413.6 and 707.5 days, and one compactly supported IMF around 4.5 years (∼1668 days). For both SSI time-series IMF7 has the same lower end support, but with higher median – 469.5 days vs. 413.6 days – IQR and upper end support for the CAR data. Moreover, when taken together, IMF7 and IMF8 for PAY seem to emulate IMF7 for CAR, with the distinction that the upper limit is somewhat less than three years, as is the case for the latter.

With the discussion of the embedded time-scales of the IMFs from figure 4 now complete, it must be reiterated that the picture that has been painted, although insightful, is incomplete, chiefly because information pertaining to amplitude or temporal localisation has not yet been exploited. A more thorough representation is presented in figure 5, which portrays the two SSI time-series as an energy density distribution over-imposed on a time-frequency space (left panel for CAR, right panel for PAY). Each pixel in these two so-called Hilbert spectra has three attributes through which it denotes the local power (color, log-scale), at a specific instant in time (abscissa), at a specific time-scale (ordinate, log-scale). For improved readability, the spectra have been binned and smoothed, hence because of the aliasing that may have been introduced by these procedures a mostly qualitative interpretation should be preferred. The side-plots on the right show the Hilbert marginal spectrum, a time-integrated (line-by-line sum) variant of each spectrum which is comparable, but not identical to, the Fourier spectrum of the same time-series.

Referring to the CAR Hilbert spectrum, depicted in the left panel of figure 5, some features can be identified as follows. A high-frequency plateau between 2 days and ∼100 days is notable, corresponding to the first five IMFs of the time-series. As previously seen in figure 4, the frequency support of these modes overlap, hence the appearance on the Hilbert spectrum of a contiguous plateau, instead of multiple discrete bands. The power of this feature is low, as can be inferred from the very slight “dent” in the marginal Hilbert spectrum corresponding to this region. Moreover, a periodic shift in color can be observed in the 4 days to 32 days band, indicative of an amplitude modulation phenomenon in accordance with the seasonal cycle – lighter yellow tones during winter turn blue-green during summer. Next, between 100 and 300 days a gap in the spectrum is apparent, as it can also be deduced by the lack of support in this region for any of the IMFs in figure 4. The dark blue trace, corresponding to IMF number six, is oscillating around the one year period, in the range of 300 to 450 days, which is also the support of this mode in the frequency box plot. These two last components have such low power that they fail to leave a mark on the marginal spectrum.

In the Hilbert spectrum for the PAY data, shown in the right panel of figure 5, the contiguous high-frequency plateau between 2 days and ∼100 days is also present, unsurprisingly, since first five IMFs of the time-series resemble those for the CAR data. The power in this band, however, is slightly greater than for CAR, especially considering the predominance of blue hues in the 2 – 4 days region that corresponds to the first and second IMFs. These two last components have such low power that they fail to leave a mark on the marginal spectrum.

In the Hilbert spectrum for the CAR data, shown in figure 4, the frequency support of these modes overlap, hence the appearance on the Hilbert spectrum of a contiguous plateau, instead of multiple discrete bands. The power of this feature is low, as can be inferred from the very slight “dent” in the marginal Hilbert spectrum corresponding to this region. Moreover, a periodic shift in color can be observed in the 4 days to 32 days band, indicative of an amplitude modulation phenomenon in accordance with the seasonal cycle – lighter yellow tones during winter turn blue-green during summer. Next, between 100 and 300 days a gap in the spectrum is apparent, as it can also be deduced by the lack of support in this region for any of the IMFs in figure 4. The dark blue trace, corresponding to IMF number six, is oscillating around the one year period, in the range of 300 to 450 days, which is also the support of this mode in the frequency box plot. This is also the component with the greatest power, as indicated by the large peak on the marginal spectrum. Interestingly enough, the corresponding time-scale fluctuations are centred in 365 days, and modulation towards greater frequencies takes place predominantly during 2002 through 2006. Between 2007 and 2009, however, the frequency modulations are less pronounced – corresponding to a period when solar activity is at a minimum. The last two low-frequency, orange-green traces on the spectrum correspond to IMF7 and IMF8 of the CAR data. For the seventh component mode mixing is apparent through the occasional sharing of the yearly time-scale band with IMF6, between the end of 2003 and the first half of 2006, and shortly again before 2009. These two last components have such low power that they fail to leave a mark on the marginal spectrum.

In the Hilbert spectrum for the PAY data, shown in the right panel of figure 5, the contiguous high-frequency plateau between 2 days and ∼100 days is also present, unsurprisingly, since first five IMFs of the time-series closely resemble those for the CAR data. The power in this band, however, is slightly greater than for CAR, especially considering the predominance of blue hues in the 2 – 4 days region that corresponds to the first and second IMFs. This is also apparent when looking at the marginal spectrum, where a distinct peak at this time-scale can be clearly made out. The amplitude modulation phenomenon in phase with the seasonal variations, previously identified in
the CAR data, is even more pronounced in the PAY spectrum; once again light yellow tones that occur during low insolation in winter turn blue during the high irradiance regime of summer. Here too, between 100 and 300 days a gap in the spectrum is also apparent, with the notable exception of the mode mixing phenomena associated with IMF6 that occur during 2003, 2007 and 2009. As previously shown in figure 4, the lower range of the frequency distribution of the sixth mode overlaps the high frequency plateau, which is portrayed by the three jutting spikes from the yearly band into the sub-100 days plateau on the PAY spectrum. Since these protruding filaments have such low power that they leave no imprint on the marginal spectrum, their most probable origin can be attributed to some sort of numerical artefacts. The yearly variability of the data can be, unsurprisingly, identified with the sharp peak at roughly 365 days on the marginal spectrum. In terms of the Hilbert time-frequency-energy representation, however, the seasonal cycle cannot be attributed to one sole component. This is due to mode mixing as seen in figure 4 where the range of IMF7 completely overlaps that of IMF6. It could be argued that IMF6 should represent the "true" seasonal cycle, as indicated by the median of its frequency distribution, however for 2003 and 2007 the Hilbert spectrum reveals that whenever IMF6 extends its tendrils into the high-frequency plateau and drastically reduces its power, IMF7 seems to "pick up the slack" by reaching in to the vacated one year band. During 2008, and briefly during spring 2004, it is found that the two components trade places altogether, with IMF6 assuming lower values than IMF7 on the frequency scale. The last two low-frequency components, IMF8 and IMF9, can be seen here too to have relatively low power – both only just manage to make a very slight indentation on the marginal spectrum. For IMF9, the frequency spread in the Hilbert representation is in good agreement with its narrow support from figure 4.

The main findings of this section can now be summed up as follows. Both time-series are shown to share several common features: a high-frequency plateau between 2 days and 100 days, a relatively low power region in the marginal spectrum between 100 days and 300 days, a high powered seasonal or yearly cycle peaking at 365 days, and two low-frequency components in the 1.5 years to 6 years band.

4. Discussion and conclusion

It has been observed in figure 4 that the median periods of the first five IMFs follow a dyadic repartition, that can be approximated by the series 3.1, 7, 14, 27, and 54 days. It is found that this type of frequency doubling in the IMFs of geo- and astrophysical signals is recurrent throughout the scientific literature. Specifically, when investigating three independent datasets of satellite observations of the total (extraterrestrial) solar irradiance, [23] consistently find a
Fig. 5. Hilbert spectra of the two SSI time-series – CAR (left) and PAY (right). For each station, the plot on the right depicts a time-integrated version of the corresponding spectrum. Year markers correspond to January 1st.

Similar dyadic scale progression of modes at 13.5, 27, and 54 days, statistically significant within the 95% level, that correspond to the 27-day solar rotation period and its (sub-)harmonics. Similarly, [24] also report intrinsic periodicities having an average of $25^{+2}_{-2}$ and $44^{+10}_{-2}$ days in the signals of five solar proxies: 10.7 cm radio flux intensity, helioseismic frequency shift, and sunspot area for the whole Sun, and separately for the northern and southern solar hemispheres (error bars in sub-/superscript estimated at the half-level width of the corresponding probability histogram). However intriguing, the imprint of a solar rotation signature on ground measurements of the SSI is rather improbable, as it would imply the existence of hitherto unknown physical mechanisms in Earth’s atmosphere [25]. Moreover, amplitude modulation of high-frequency components by lower frequency ones is also found in sunspots number time series [24,26], and in surface air temperatures [27], and is indicative of non-linear cross-scale interactions [9].

Similar remarks can be made for the IMFs on time-scales greater than one year. The corresponding IMFs for the CAR time-series have median periods of 469.5 and 1305.1 days, while those for PAY have medians of 413.6, 707.5 and 1668.6 days. The time-scales of these oscillations seem to agree once again with those related to the so-called solar quasi-biennial oscillations (QBO), which are variations in the Sun’s activity with periodicities between 0.6 and 4 years [28]. In [24] the aforementioned solar proxies IMFs have average periods of $395^{+46}_{-46}$, $626^{+69}_{-113}$, and $1423^{+196}_{-146}$ days respectively. The same solar QBOs in the 1 year to 4.5 years range are also identified as a fundamental time-scale of variability in solar magnetic synoptic maps [29]. In [30] a 1.68 year (613 days) spectral peak is also reported, found in retrospectively determined cloud cover from measurements of the surface solar radiation, indicating a galactic cosmic ray, rather than solar irradiance, cloud effect. Nevertheless, the cautionary statement from the previous paragraph is reiterated, as no physical mechanism in Earth atmosphere has been proposed.

Orbital mechanics and parameters of the Earth-Sun system may account for the spectral peaks detected around the one year mark in both CAR and PAY datasets. However, it has been shown that for the PAY time-series, two components are needed to explain the seasonal variability of the data. For this particular time-series, no clear physical significance can be attributed to either mode. Rather, it is the frequency representation of an aggregate of the two that is able to account for variability at this time-scale.

To sum up, the HHT analysis of decennial time-series of daily means of SSI measurement, at two distinct BSRN stations has revealed the following: the presence of a high-frequency (2-100 days) plateau composed of five IMFs that are amplitude-modulated by the yearly cycle; a distinct band having very low power with time scales of 100-300 days; a distinct spectral peak at the one year mark accounting for the seasonal variability; multiple low power “QBOs”; and a non-linear trend that, for one station at the least, is clearly associated to the yearly means of the data. These characteristic features of the temporal variability allow to differentiate between the local climates of the ground measurement stations and have possible implications on the modelling and forecasting of the surface solar radiation field.
References


