Adv. Sci. Res., 14, 35–48, 2017
www.adv-sci-res.net/14/35/2017/
doi:10.5194/asr-14-35-2017

Advances in Science & Research
Open Access Proceedings

16th EMS Annual Meeting & 11th European Conference on Applied Climatology (ECAC)

Do modelled or satellite-based estimates of surface solar irradiance accurately describe its temporal variability?

Marc Bengulescu[1], **Philippe Blanc**[1], **Alexandre Boilley**[2], **and Lucien Wald**[1]

[1]MINES ParisTech, PSL Research University, Centre for Observation, Impacts, Energy, CS 10207, 06904 Sophia Antipolis CEDEX, France
[2]Transvalor SA, 694 Avenue du Dr. Maurice Daunat, 06255 Mougins CEDEX, France

Correspondence to: Marc Bengulescu (marc.bengulescu@mines-paristech.fr)

Received: 14 December 2016 – Revised: 8 February 2017 – Accepted: 13 February 2017 – Published: 21 February 2017

Abstract. This study investigates the characteristic time-scales of variability found in long-term time-series of daily means of estimates of surface solar irradiance (SSI). The study is performed at various levels to better understand the causes of variability in the SSI. First, the variability of the solar irradiance at the top of the atmosphere is scrutinized. Then, estimates of the SSI in cloud-free conditions as provided by the McClear model are dealt with, in order to reveal the influence of the clear atmosphere (aerosols, water vapour, etc.). Lastly, the role of clouds on variability is inferred by the analysis of in-situ measurements. A description of how the atmosphere affects SSI variability is thus obtained on a time-scale basis. The analysis is also performed with estimates of the SSI provided by the satellite-derived HelioClim-3 database and by two numerical weather re-analyses: ERA-Interim and MERRA2. It is found that HelioClim-3 estimates render an accurate picture of the variability found in ground measurements, not only globally, but also with respect to individual characteristic time-scales. On the contrary, the variability found in re-analyses correlates poorly with all scales of ground measurements variability.

1 Introduction

The Sun is of the utmost importance for the planet Earth. Not only does it play a central role in our solar system, but it also represents the main power source for the Earth, being the main driver behind weather and climate (Trenberth et al., 2009). As such, the Global Climate Observing System (GCOS) has defined the surface solar irradiance (SSI) as an Essential Climate Variable (ECV), used for the characterization of the state of the global climate system and for long-term climate monitoring (GCOS2010).

The SSI is known to exhibit variations on a large dynamic range with respect to both time and geographical position. This is due to a wide array of factors. Some of these operate at long time-scales, from decades to millennia and beyond, and are related to the stellar variability of the Sun (Beer, 2000), or to changes in the orbital parameters of the Earth, e.g. obliquity of the ecliptic, eccentricity of the orbit, and axial precession (Beer et al., 2006). The revolution of the Earth around the Sun gives rise to the yearly cycle. Another well known factor is the rotation of the Earth around its own axis, that yields variations in the SSI with a period of 1 day. Atmospheric effects such as scattering and absorption, mainly due to O_2 and stratospheric O_3, clouds, and aerosols (Madronich and Flocke, 1999) may exhibit even shorter time-scales. Orographic factors, such as shadowing effects due to relief, influence the SSI on a yearly to very short time-scale. The variability of the SSI at different time-scales is, thus, a leitmotif of this study.

In order to analyse the temporal variability of the SSI, measurements of this physical quantity are needed. A primary way of obtaining such data is recording the values of the SSI at ground stations by using pyranometers or pyrheliometers. Nevertheless, even sources of high quality solar radiation measurements, such as the Baseline Surface Radiation Network (BSRN), a worldwide radiometric network providing accurate readings of the SSI at high temporal resolution (Ohmura et al., 1998), are spatially very sparse, capture only temporal variability on a very limited set of locations and, in addition, have been found to have a data gap percentage ranging from 4.4 to 13 % (Roesch et al., 2011). Furthermore, measurements from point sources can only reveal information about the temporal variability of the solar radiation at one particular site.

In practice, information about the SSI is often required at geographical locations different from any measuring station. But extending the representativity of ground station measurements to surrounding areas cannot be applied to regions where the physical and/or climatological distance between stations is large (Zelenka et al., 1992). Using the nearest-neighbour approximation, Zelenka et al. (1992) have found that the relative root-mean-square error (relative RMSE) of the estimates is proportional to the square root of the distance between sites.

Another option is to make use of satellite based methods, which are a good supplement in long term solar resource assessment. Perez et al. (1997) have shown that when the distance from a station exceeds 50 km, in the case of daily aggregates, or 34 km for hourly data, – Zelenka et al. (1999) estimate this as low as 20 km – satellite-derived values of the SSI are more accurate than estimating using the nearby measuring point. Thus, given the scarcity and spatial sparsity of long-term ground measurements of the SSI, satellite-derived estimates of surface solar radiation remain a good complement to ground station data (Lefèvre et al., 2014).

Yet another possibility for estimating the solar radiation at ground level is provided by global atmospheric re-analyses from numerical weather models. The main benefits are the wide, even global, coverage and the spanning of multi-decennial time periods. However, some authors have found a large uncertainty relative to satellite-based irradiance estimates and advise against using data from re-analyses (Lohmann et al., 2006; Boilley and Wald, 2015). Others nevertheless find SSI datasets from re-analyses to be suitable for photovoltaic applications (Richardson and Andrews, 2014). Efforts to improve the adaptation of re-analyses solar radiation datasets to a specific geographical site are also ongoing (Polo et al., 2016).

In this context, we investigate and analyse here the temporal variability in time-series of daily means of SSI for two geographical locations, at different time-scales, as found in the outputs of different models, satellite estimates, re-analyses, and ground measurements. To gain better insight into the causes of variability of the SSI, we follow the downwelling solar shortwave irradiance along its path through the atmosphere towards the surface. The modelled top of the atmosphere (TOA) solar irradiance is first analysed as a clean input signal, devoid of any atmospheric perturbations, in order to reveal the natural variability of the exo-atmospheric solar input. To account for variability owing to atmospheric effects such as scattering or absorption due to water vapour or aerosols, but excluding any influence of clouds, the output of a clear-sky (i.e. cloud-free) model of the SSI is scrutinized. The role of clouds on variability is lastly inferred by analysing pyranometric ground measurements. The fitness for use of satellite estimates and re-analyses data is then assessed, by comparison with the measured data.

The novelty of our work stems from the fact that, unlike previous studies where global statistical indicators are employed (Espinar et al., 2009), here we decompose the datasets into their constituent characteristic time-scales before doing the analysis. To this end, we employ the adaptive, data-driven Hilbert–Huang Transform (HHT) (Huang et al., 1998), which has been shown to be a good analysis method on solar radiation datasets (Duffy, 2004; Calif et al., 2013; Bengulescu et al., 2016a, b, c). In this way, the fitness for use of the modelled and estimated SSI can be assessed not only at a global, whole dataset level, but also on an per time-scale basis.

The study develops as follows. Section 2 describes the datasets and the analysis technique. Results are presented in Sect. 3, discussion thereof being deferred to Sect. 4. The conclusions and outlook are given in Sect. 5. The availability of the software code and the datasets needed to reproduce our findings is indicated in Sects. 6 and 7, respectively.

2 Data and methods

2.1 Data

The data consists of multiple time-series of daily means of solar irradiance corresponding to two geographical locations in Europe: Vienna, Austria (48.25° N; 16.35° E; elevation 203 m), and Kishinev, Moldova (47.00° N; 28.82° E; elevation 205 m). The temporal coverage of the data is 9 years, from 1 February 2004 to 31 January 2013. The number of samples in each dataset is $9 \times 365 = 3285$ days (leap days are omitted). The datasets for Vienna (VIE) are plotted in Fig. 1. A similar plot for Kishinev (KIV) is provided in the Supplement.

Six datasets are used for each location:

- modelled exo-atmospheric irradiance;
- modelled clear-sky irradiance at ground level;
- pyranometric measurements of the SSI;
- Meteosat satellite-based SSI estimates;
- radiation products from the ERA-Interim re-analysis;
- radiation products from the MERRA2 re-analysis.

Data availability and Digital Object Identifiers (DOIs) are indicated in Sect. 7.

The top of the atmosphere (TOA) solar irradiance time-series has been generated using the SG2 algorithm (Blanc and Wald, 2012). This dataset is built using the constant value of 1367 W m^{-2} for total solar irradiance, though it could have made use of satellite measurements of this quantity. Recently,

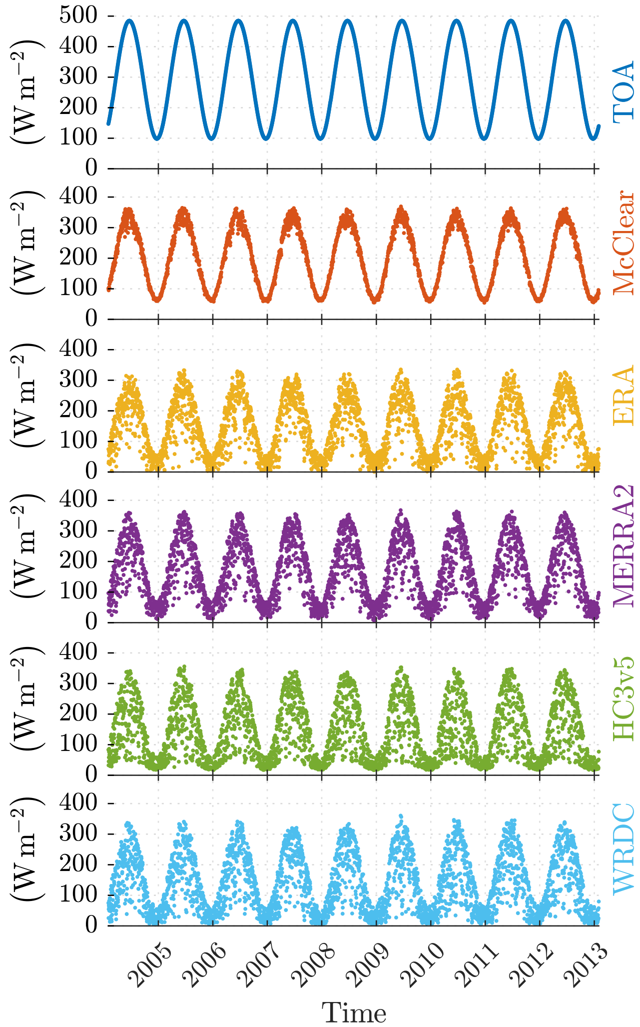

Figure 1. The six solar irradiance time-series for VIE investigated in this study, spanning 1 February 2004 to 31 January 2013. From top to bottom: TOA, McClear, ERA, MERRA2, HC3v5, and WRDC. Each point corresponds to a daily mean of irradiance. Time markers on the abscissa indicate the start of the corresponding year.

the yearly mean of the total solar irradiance has been revised to 1361 W m^{-2} (Prša et al., 2016) or 1362 W m^{-2} (Meftah et al., 2014). This discrepancy does not impact the validity of our analysis.

The dataset of downwelling surface solar irradiance, under clear-sky conditions (i.e. cloud-free), is generated using the McClear model (Lefèvre et al., 2013). The McClear model is part of the Copernicus Atmosphere Monitoring Service (CAMS) and its inputs comprise 3 h estimates of the aerosol properties and total column contents of water vapour and ozone also provided by CAMS.

Estimates of the SSI derived from Meteosat satellite imagery by the Heliosat-2 method, as described by Rigollier et al. (2004) and modified by Qu et al. (2014), are obtained from the HelioClim-3 (HC3v5) database (Blanc et al., 2011). The daily means of TOA, McClear, and HC3v5 irradiance were obtained directly from the SoDa website (http://www.soda-pro.com).

Pyranometric ground measurements of the daily SSI were obtained from the World Radiation Data Centre (WRDC) (Tsvetkov et al., 1995) for the two stations. No detailed information on data quality is provided except for a quality flag. Considering the high quality of maintenance of these stations, we consider that the data obey the good quality level set by the World Meteorological Organization (CIMO2014) which specifies a 5 % uncertainty, expressed as the percentile 95 of the deviations (P95). If a normal law is assumed for the deviations, then there is a 0.3 % chance that a deviation exceeds 1.5 times P95, i.e. 7.5 % of the SSI. Given the global means of SSI of the WRDC datasets, both greater than 135 W m^{-2}, the 7.5 % uncertainty is 10.2 W m^{-2} for VIE and 11.3 W m^{-2} for KIV.

The ERA-Interim product "Surface Solar Radiation Downwards" (ECMWF2009), from 2004 to 2014, was retrieved using the `ecmwfapi` python library on the Meteorological Archival and Retrieval System (MARS). The raw ERA-Interim data is a forecast of accumulated SSI expressed in J m^{-2}, and has a spatial resolution of 0.75° × 0.75°. Both the H+12 forecast from 00:00 UT re-analysis and from 12:00 UT re-analysis were summed and then divided by 24 to obtain a daily SSI value in W m^{-2} for the 4 nearest points around the location. These values were then bi-linearly interpolated at the exact location.

The 1 h radiation diagnostics M2T1NXRAD from MERRA2 have been extracted for the four nearest points and bi-linearly interpolated to generate the time series at the exact location. The MERRA2 data are in W m^{-2} directly and have a spatial resolution of 0.5° × 0.65° in latitude and longitude. Values were then summed over each day and divided by 24 to obtain a daily SSI value in W m^{-2}.

2.2 The Hilbert–Huang transform (HHT)

The goal of the study at hand is to first decompose the scrutinized time-series into uncorrelated sub-constituents that have distinct characteristic time-scales. Analysis then ensues at each distinct scale of intrinsic variability. These time-scales, or characteristic periods, are nothing more than the inverse of the frequency of the various processes from which the data stems. As such, analysis techniques that depict the changes with respect to time of the spectral content of a time-series are to be favoured, since they enable both the identification of periodicities and the following of the dynamic evolution of the processes generating the data. For a review of such regularly employed methods in geophysical signal processing see Tary et al. (2014).

The non-linear and non-stationary characteristics of the SSI (Zeng et al., 2013) are also worth consideration. Handling such data issued from the non-linear interaction of physical processes, often also found under the influence of non-stationary external forcings calls for an adaptive data analysis approach (Wu et al., 2011). The ideal analysis technique must make no a priori assumptions regarding the character of the data, i.e. neither linearity, nor stationarity should be presumed, since the nature of the processes that have generated the data is not usually known before the analysis is carried out. Adaptivity to the analysed data is also a desirable feature, i.e. letting the data itself decompose onto a set of basis functions determined by its local characteristic time scales, instead of a projection onto a predefined set of patterns. This 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 (Wu et al., 2011).

As such, this study employs the Hilbert–Huang Transform, an adaptive, data-driven analysis technique. The HHT is ideally suited for non-linear and non-stationary data and it adaptively decomposes any time-series into basis functions derived from the local properties of the data (Huang et al., 1998). The method is used here to extract, depict and analyse the characteristic time-scales of variability of solar irradiance time-series. The data analysis method operates in the time domain and makes no beforehand assumptions regarding the analysed dataset (stationarity or linearity). The method is also adaptive, letting the data decompose itself onto a finite number of locally derived data-driven basis functions (Wu et al., 2011), in contrast with the Fourier or wavelet transforms that impose a predefined set of functions for the decomposition, such as trigonometric functions or wavelets (Huang and Wu, 2008). Further details on how the HHT compares to other spectral methods, the Fourier or wavelet transforms included, can be found in Huang et al. (1998) and Wu et al. (2011). A case study comparing the HHT and the wavelet transform, as applied to surface solar radiation data, is offered in Bengulescu et al. (2016a).

The HHT consists in two steps, the empirical mode decomposition (EMD), followed by Hilbert spectral analysis (HSA), both detailed hereafter.

2.2.1 Empirical mode decomposition (EMD)

The EMD is algorithmic in nature, and iteratively decomposes data into a series of oscillations; within a series, oscillations have a common local time-scale, called Intrinsic Mode Function (IMF). An IMF is a function that satisfies two criteria: (1) its number of zero crossings and number of extrema differ at most by one; (2) at any point, the mean value of its upper and lower envelopes is zero. The theoretical signal model for IMFs is an amplitude modulation–frequency modulation (AM–FM) one. Given the adaptive nature of the EMD, the IMFs represent the basis functions onto which the data is projected during decomposition. Any two IMFs are locally orthogonal for all practical purposes, however, given the empirical nature of the method no theoretical guarantee can be provided. In practice, it is found that the relative difference between the variance of the input signal and the sum of variances of the IMFs (i.e. the spectral leakage) is typically less than 1 %; only for extremely short data ranges does the leakage increase to 5 %, comparable to that of a collection of pure trigonometric functions having the same data length (Huang et al., 1998). By design, IMFs have a well behaved Hilbert transform (Huang et al., 1998). The EMD can be sketched as follows:

1. let $r(t)$ hold the data, initialize IMF counter $k = 1$;
2. let $h(t) \leftarrow r(t)$
3. a. find the minima and maxima of $h(t)$;
   b. interpolate minima to find lower envelope: $L(t)$;
   c. interpolate maxima to find upper envelope: $U(t)$;
   d. find mean of envelopes: $m(t) \leftarrow \frac{L(t)+U(t)}{2}$;
   e. substract the mean: $h(t) \leftarrow h(t) - m(t)$;
4. if $h(t)$ is not an IMF, repeat step 3;
5. store IMF: $c_k(t) \leftarrow h(t)$
6. update the residual: $r(t) \leftarrow r(t) - c_k(t)$
7. if $r(t)$ is not monotonic, increment k and go to step 2;
8. return IMFs $c_k(t)$, $k \in \{1, \ldots, N\}$ and residual $r(t)$.

Step 3 is called the sifting loop and it controls the filter character of the EMD. An infinite number of sifting iterations would asymptotically approach the result of the Fourier decomposition (i.e. constant amplitude envelopes) (Wang et al., 2010). Flandrin et al. (2004) have shown the wavelet-like dyadic filter bank character of the EMD and Wu and Huang (2010) have found that this dyadic property is enforced by keeping the number of sifting iterations small, around 10, which also assures maximum component separation and minimum leakage. The IMFs can also be shown to satisfy the envelope–carrier relationship, thus guaranteeing the existence of a unique true intrinsic amplitude function and of a unique phase function (Huang et al., 2013).

After all the IMFs are extracted, what is left of the data is called a trend or residue, which can no longer be considered as an oscillation at the span of the data. Bengulescu et al. (2016c) have shown that, for time-series of daily means of solar irradiance, the trend approximates the yearly mean. More generally, the trend can be interpreted as low-pass filtered version of the data (Moghtaderi et al., 2013), therefore it is excluded from this analysis. To illustrate the EMD process, the IMFs obtained from the decomposition of the WRDC time-series for VIE are plotted in Fig. 2. The modes IMF1…IMF10 and the residual (Res.) are plotted as YZ slices along the x axis, with time running on the y axis, and amplitude on the z axis. The zero-centred oscillatory nature of the IMFs can be clearly seen, as well as the local time-scale increase with mode number. The IMF1 has a

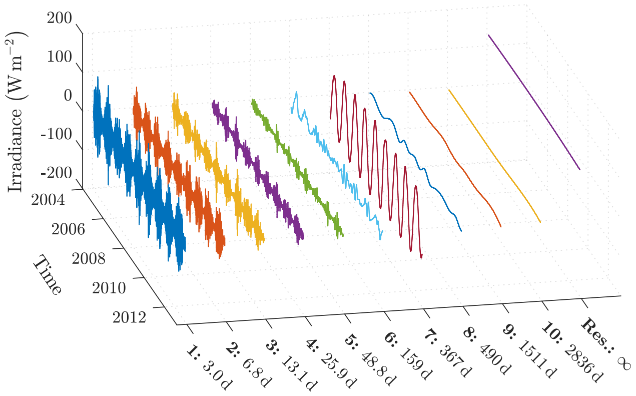

Figure 2. The IMFs obtained by decomposing the WRDC time-series for VIE, plotted as YZ slices, with the x axis denoting their number (IMF1…IMF10) and mean time-scales in days. For reference, the residual (Res.) is also included. The amplitude, or irradiance, of each IMF is plotted on the z axis, and time runs on the y axis. Time markers denote 1 January of the corresponding year.

mean time-scale of 3.0 days and exhibits a large variability in time. As the IMF rank increases, the time-scale increases and the variability decreases. The exception is for IMF7 at 367 days which exhibits the greatest variability, as discussed later. Edge effects in the EMD appear because of oscillations of the interpolating splines and are usually contained within a half-period of a component at data boundaries (Wu et al., 2011).

This study uses a modified version of the original EMD algorithm, the Improved Complete Ensemble Empirical Mode Decomposition, introduced by Colominas et al. (2014). This enables the exact decomposition of the data, i.e. the sum of all IMFs, including the trend, reconstructs the original time series, and is more robust with respect to noise. To decrease computation time, the fast EMD routine proposed by Wang et al. (2014) is employed. See Sect. 6 for code availability.

2.2.2 Hilbert spectral analyis (HSA)

Once the EMD has decomposed the data into IMFs, the last step of the HHT consists in the Hilbert spectral analysis. For each IMF $c_k(t)$ its Hilbert transform is computed as given by Eq. (1), where k designates the kth IMF, and P indicates the Cauchy principal value.

$$\sigma_k(t) = \mathcal{H}(c_k(t)) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{c_k(\tau)}{t-\tau} \mathrm{d}\tau \tag{1}$$

The pair can then be used to construct the complex-valued analytic signal proposed by Gabor (1946), described by an amplitude modulation–frequency modulation (AM–FM) model, as in Eq. (2).

$$z_k(t) = c_k(t) + i \cdot \sigma_k(t) = a_k(t) \cdot e^{i \cdot \theta_k(t)} \tag{2}$$

In the AM–FM model, the instantaneous amplitude is given by Eq. (3).

$$a_k(t) = \sqrt{c_k^2(t) + \sigma_k^2(t)} \tag{3}$$

The instantaneous phase can be derived from Eq. (4).

$$\theta_k(t) = \tan^{-1}\left(\frac{\sigma_k(t)}{c_k(t)}\right) \tag{4}$$

The instantaneous frequency, i.e. the inverse of the local time-scale, is then just the first time derivative of the instantaneous phase, as in Eq. (5).

$$\omega_k(t) = \frac{1}{2\pi} \frac{\mathrm{d}\theta_k(t)}{\mathrm{d}t} \tag{5}$$

The role of HSA is to decompose each IMF into two time-varying components, namely instantaneous amplitude and instantaneous frequency, in order to determine, in a time-dependent manner, how much power (amplitude squared) occurs at which time-scales, as in Eq. (6). This representation, called the Hilbert energy spectrum, plots the data as an energy density distribution overlaid on the time-frequency space (Huang et al., 2011).

$$S(\omega, t) = \sum_{k=1}^{N} a_k^2(t) \cdot e^{i \int \omega_k(\tau) \mathrm{d}\tau} \tag{6}$$

The time-integrated version of Eq. (6), the Hilbert marginal spectrum $S_\mathrm{M}(\omega)$, is similar, but not identical to, the traditional Fourier spectrum, and is given in Eq. (7).

$$S_\mathrm{M}(\omega) = \int_0^T S(\omega, t) \mathrm{d}t \tag{7}$$

3 Results

The 12 datasets (6 datasets per station) have been decomposed by the EMD into 10 IMFs and a residual, as shown in Fig. 2. Similar plots are available for the rest of the datasets in the Supplement. To summarize the results, the mean characteristic scales of variability for all the IMFs of the VIE datasets have been compiled into Table 1, while the corresponding mean amplitudes are given in Table 2. Similar summaries for the KIV datasets are provided in Appendix A, as Table A1 for time-scales and Table A2 for amplitude, respectively.

Table 1. Mean IMF time-scales in days for the VIE datasets.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
TOA	6.5	9.3	17.0	30.8	344	366	366	1431	3134	3242
McClear	3.0	6.9	13.2	25.3	55.3	366	417	803	2971	3432
WRDC	3.0	6.8	13.1	25.9	48.8	159	367	490	1511	2836
HC3v5	3.0	6.7	13.2	26.2	49.8	156	368	475	914	1645
ERA	3.1	6.8	13.1	25.8	50.8	180	371	576	1042	2838
MERRA2	3.0	6.7	12.8	26.0	52.1	217	368	568	1536	3351

Table 2. Mean IMF amplitudes in $W m^{-2}$ for the VIE datasets.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
TOA	0.1	0.1	0.1	0.1	3.1	186	5.4	1.2	2.0	0.4
McClear	6.4	4.6	3.7	3.0	3.2	128	10	1.6	1.7	0.4
WRDC	44	24	19	16	12	16	96	8.3	4.2	2.8
HC3v5	47	26	20	17	13	14	91	12	6.5	2.1
ERA	33	20	16	13	10	19	89	6.9	5.0	2.2
MERRA2	38	21	16	13	10	30	94	4.7	4.5	3.1

3.1 Intrinsic time-scales of variability

From Tables 1 and 2 it can be seen that the most significant time-scale of variability present in the TOA time-series is around the 1 year mark, as evidenced by IMF6. The mean period of this component is 366 days and its mean amplitude of $186 W m^{-2}$ is two orders of magnitude greater than the other modes. From this, it can be inferred that this yearly mode, a result of the revolution of the Earth around the Sun, is the only significant scale of variability of the TOA. This result is unsurprising, since it can also be inferred from Fig. 1 top panel, where the TOA series does not exhibit any other variability apart of this mode, i.e. it is almost a perfect sine wave with a period of one year. Adjacent to this sixth mode, IMF5 and IMF7 also display similar time-scales of 344 and 366 days, however their mean amplitudes are 3.1 and $5.4 W m^{-2}$ respectively, thus their origin is attributed to spectral leakage from the main mode. The first four IMFs have negligible amplitudes ($0.1 W m^{-2}$) and are probably the manifestation of residual noise. IMF8 has a mean time-scale of 1431 days and its low amplitude of $1.2 W m^{-2}$ is an indication that it might be a numerical artefact. IMF9 and IMF10 are also assumed to be non-physical, because their large periods of over 3100 days are indication that these components are entirely immersed in the time coverage where edge effect are non-negligible. As a reminder, edge effects are important at the half-period of a component at data boundaries, which for these two latter IMFs yields 1550 days forward from 1 February 2004 and 1550 backwards from 31 January 2013, spanning almost the entire data range.

For the McClear time-series, as well as for the rest of the VIE datasets, IMF1...IMF5 display remarkably similar features, such as monotonically decreasing amplitudes and time-scales that exhibit period doubling, roughly following the dyadic scale: 3 days $\rightarrow$ 6.8 days $\rightarrow$ 13.1 days $\rightarrow$ 26 days $\rightarrow$ 51 days. The amplitudes for McClear are 3 to 7 times less than in other time-series, and less than $6.5 W m^{-2}$. The break in the monotonic decrease of amplitude with scale for McClear – $3.2 W m^{-2}$ for IMF5 vs. $3 W m^{-2}$ for IMF4 – is considered an artefact, since this monotonicity holds for the other datasets and also for KIV (see Table A2). The presence of these first five dyadic scales in all the datasets, with the exception of TOA, allude to their possible origin as being cloud-free atmospheric affects. This can also be observed in Fig. 1, where, as opposed to the TOA, the raw McClear time-series is seen to exhibit slight high-frequency variability, and which tends to increase during the summer months. For McClear, there is little to no variability in the 2–3 months to 1 year band (see also Fig. 3, discussed later on). As for TOA, IMF6 of the McClear time-series is clearly associated with the yearly variability, by its mean period and amplitude of 366 days and $128 W m^{-2}$, respectively. Here too, this is the most energetic spectral component of the dataset. IMF7 of McClear has a median time-scale of 417 days and a median amplitude of $10 W m^{-2}$. IMF8...IMF10 have very low amplitudes of less than $1.8 W m^{-2}$, hence their origin cannot be unambiguously determined. This is especially the case with IMF9 and IMF10, which reside almost entirely in the edge effect region because of their large periods of about 3000 days and greater.

The WRDC dataset unsurprisingly shares many features with HC3v5, ERA and MERRA2 datasets, since the latter three are intended to be accurate estimates of the former. The rest of the results will be presented in a lumped form for these four time-series. For IMF1...IMF5, HC3v5 agrees bet-

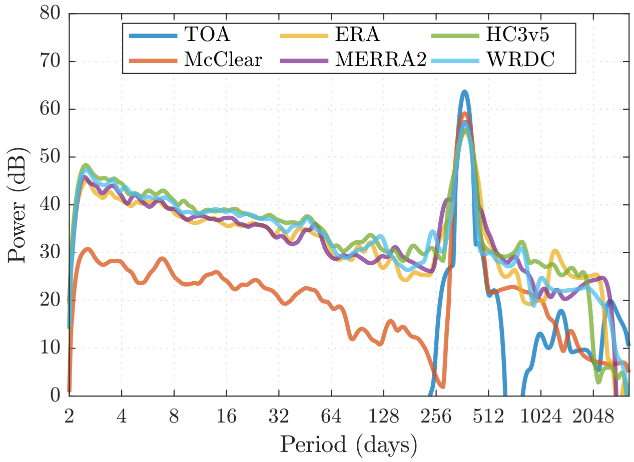

Figure 3. The Hilbert marginal spectra for the VIE datasets: TOA, McClear,WRDC, HC3v5, ERA, MERRA2. The abscissa indicates the time-scale on a binary logarithm, and the ordinate denotes power in dB.

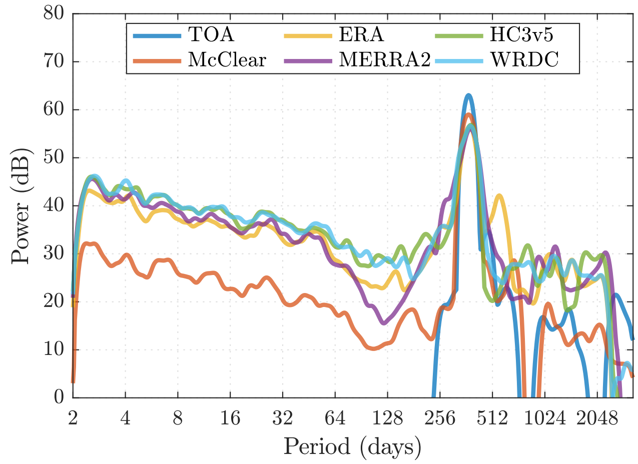

Figure 4. The Hilbert marginal spectra for the KIV datasets: TOA, McClear,WRDC, HC3v5, ERA, MERRA2. The abscissa indicates the time-scale on a binary logarithm, and the ordinate denotes power in dB.

ter with WRDC than the re-analyses in terms of mean amplitudes and takes on only slightly higher values (1 to 3 $\mathrm{W\,m^{-2}}$ or less than 5 % on average, Table 2). This can also be seen in Fig. 1, where the HC3v5 data exhibits a slight overestimation of the low SSI values for WRDC during winter. For the first five modes, both ERA and MERRA2 have mean amplitudes that are on average 17 % less than those of WRDC. This is also apparent in Fig. 1, where fewer samples of less that 150 $\mathrm{W\,m^{-2}}$ are occurring between spring and autumn for the re-analyses time-series; this is mostly visible for the year 2007. For IMF6, HC3v5 closely follows WRDC, both in terms of time-scale: 156 days vs. 159 days, and amplitude: 14 $\mathrm{W\,m^{-2}}$ vs. 16 $\mathrm{W\,m^{-2}}$. For the same IMF6, both ERA and MERRA2 exhibit significantly greater time-scales: 180 and 217 days, respectively, and amplitudes: 19 and 30 $\mathrm{W\,m^{-2}}$. The first six modes are responsible for the high-frequency variability, which is also manifest in Fig. 1, i.e. the estimates and ground measurements of the SSI are significantly more variable than the McClear time-series. For all these datasets the yearly variability cycle is captured by IMF7, with generally good agreement across datasets in terms of both time-scale (367 to 371 days) and amplitude (89 to 96 $\mathrm{W\,m^{-2}}$). The greatest modes, IMF8...IMF10, have amplitudes less than the uncertainty threshold, with the exception of IMF8 for HC3v5.

Generally similar results are also obtained for KIV, as summarized in Tables A1 and A2. Apart minor differences in numerical values, the only notable exception in the KIV datasets are the significantly greater time-scales for IMF6 of 280 and 298 days respectively for ERA and MERRA2, as opposed to 199 and 176 days for WRDC and HC3v5, respectively.

3.2 Results of Hilbert spectral analysis

The previous summary of the results, although informative, is static in the sense that only two features are used to characterize, in an approximative manner, each time-evolving IMF: the long-term average amplitude and time period. To make use of the full potential of the HHT, which can follow both the temporal and the spectral evolution of the data, Hilbert spectra were also computed for all the datasets (not shown) and are provided in the Supplement. From these Hilbert spectra the marginal, time-integrated, versions were computed and are presented in Fig. 3 for VIE and in Fig. 4 for KIV.

The TOA spectrum for VIE in Fig. 3 confirms the previous findings. For this dataset the only significant mode of variability is found at the one year mark, and has a power of about 63 dB. A slight peak of 17 dB is also present around 1400 days, corresponding to IMF8. The end region of the spectrum, from 2500 days onwards also contains some power, but the respective IMFs have been shown to be heavily affected by edge effects, thus the origin of this feature is ambiguous at best.

The McClear dataset is seen to introduce variability in the high-frequency regime, whose power decreases almost monotonically from 30 dB at 2 days, to about 2 dB at roughly 300 days. Most of this variability occurs during summer, as observed in Fig. 1 (see also the full Hilbert spectrum of McClear in the Supplement). The yearly variability component stands out again, this time with just less than 60 dB in power. From here, power decreases to a minor spectral shoulder of 22 dB at 800 days (IMF8), after which it fades out towards larger frequencies.

As previously shown, the high frequency variability (IMF1...IMF5) of HC3v5 matches more closely that of WRDC, while the re-analyses have slightly (2–5 dB) less power. The power of these features is 15 dB greater in the estimates and ground measurements of SSI than that found in the clear-sky regime. From 170 days to 256 days, WRDC and HC3v5 overpower the re-analyses. After 256 days, the power in the re-analyses overcomes that of WRDC and HC3v5, until approximately one year. This can also be seen from Tables 1 and 2, where IMF6 for the re-analyses is seen to be greater than the other two time-series, both in terms of amplitude and of time-scale. Again, the yearly variability component is the largest spectral characteristic, with WRDC peaking at 57 dB. Here, the other datasets closely agree with the ground measurements as per the VIE results summary tables. After the 500 days mark the interpretation of the different spectral features is ambiguous, both because of their mean amplitudes failing to rise above the uncertainty level, and also because of the progressively large impact of the edge effects, especially towards the end of the spectrum.

The spectra for the KIV datasets in Fig. 4 are very similar to their VIE counterparts. For KIV too, TOA exhibits only a sharp yearly variability component (IMF6), while McClear introduces an almost monotonically decreasing band of high-frequency variability. Unlike for VIE, the McClear does not drop abruptly up to one year, but exhibits a rebound of 17 dB around 150 days. This is interpreted as an artefact, induced by energetic oscillations in and near the left edge effect boundary in the full Hilbert spectrum (see Supplement). From 2 days to 2 years, the HC3v5 spectrum follows the WRDC one more closely that the re-analyses. The latter two exhibit large downward excursions of 5 to 10 dB, from 70 days to roughly 150 days. For the yearly variability component, there is better agreement than for VIE between the estimates and the ground measured SSI data, all four of them peaking simultaneously at 57 dB. The large peak of 42 dB found in ERA at 530 day should be ignored, since it is caused by an energetic oscillation in the edge effect region (see Supplement). As for the VIE datasets, after the 500 days mark the interpretation spectral features becomes ambiguous.

3.3 Time-scale comparison of SSI estimates and ground measurements

Still another possibility of investigating the data is to make use of the adaptive, data-driven, time-domain filter character of the EMD. Looking at pairs of IMFs in the time domain, it is possible to construct 2-D histograms of the satellite and re-analyses estimates of SSI compared to the concomitant ground measurements. This gives a good overview of the similarity, at each characteristic time-scale of variability, between satellite estimates of the SSI or re-analyses radiation products and the WRDC measurements, which serve as ground truth.

Figure 5 illustrates the 2-D histogram for the second IMF of the HC3v5 and WRDC datasets for the VIE station. The colour of each pixel denotes the relative frequency of the WRDC–HC3v5 irradiance pairs, as encoded on the colour-

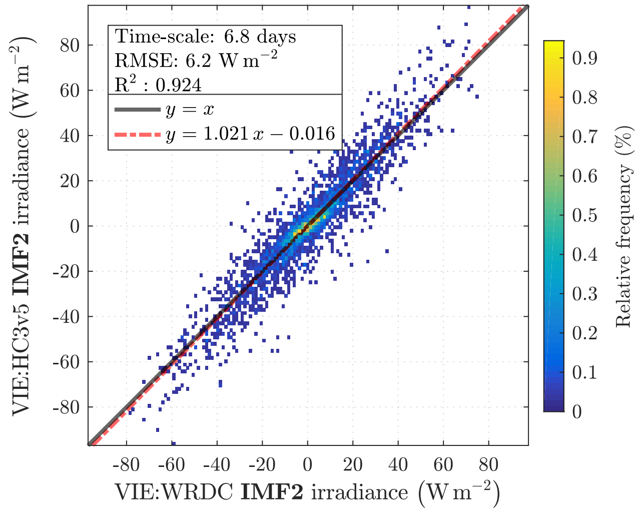

Figure 5. The 2-D histogram for IMF2 of HC3v5 and WRDC for VIE. Each pixel encodes relative frequency according to the colour-bar on the right. The solid black line denotes the identity line and the dash-dotted red line represents the best fit line. The linear regression equation is indicated in the legend. The time-scale, root-mean-square error and coefficient of determination are indicated in the panel above the legend.

bar on the right. In Fig. 5, the pattern of dots has a positive slope which indicates a positive correlation between the two variables. A robust best-fit linear regression has been performed, with the resulting line shown in dash-dotted red. Plotted in solid black is the identity line; were the two datasets identical the scatters would fall exactly onto it. The equations for the two lines are indicated in the legend. The line describing the best-fit can be seen to deviate very little from the identity line, with a slope of 1.021 indicating that for IMF2 at VIE the satellite-derived SSI slightly overestimates the SSI measured at ground level, and an intercept of -0.016, which is the expected mean irradiance value of HC3v5 when then WRDC irradiance is zero. Indeed, the linear regression model manages to explain 92.4 % of the total variability, as indicated by the coefficient of determination ($R^2 = 0.924$). The points do not fall exactly onto the best-fit line and exhibit a small scattering, with a root-mean-square error (RMSE) of 6.2 W m^{-2}. The characteristic time-scale of variability is also indicated on the plot, as 6.8 days.

Figure 6 presents a similar graph, but for ERA and WRDC datasets. Here the best-fit line deviates significantly from the identity line, with a slope of 0.647 and an intercept of 0.119. This is also reflected in the coefficient of determination of $R^2 = 0.636$. The scatter is also significantly larger, with a RMSE of 10.3 W m^{-2}.

Similar plots to those in Figs. 5 and 6 have been computed for IMF1…IMF5, IMF7 and globally, for the whole time-series, for the ground measurements and for the satellite and re-analysis estimates. Graphs are available in the Supple-

Table 3. Statistical indicators for correlations at different time-scales between SSI estimates and ground measurements for VIE.

		IMF1	IMF2	IMF3	IMF4	IMF5	IMF7	Glob.
HC3v5	R^2	0.937	0.924	0.945	0.938	0.934	0.992	0.979
	RMSE*	9.1	6.2	4.2	3.6	2.8	5.7	14.1
ERA	R^2	0.639	0.636	0.675	0.755	0.703	0.987	0.921
	RMSE*	15.5	10.3	7.8	5.5	4.6	7.0	24.8
MERRA2	R^2	0.684	0.686	0.697	0.732	0.762	0.985	0.928
	RMSE*	16.1	9.7	7.5	5.6	4.3	8.4	26.4

* RMSE has units of $\mathrm{W\,m^{-2}}$.

Table 4. Statistical indicators for correlations at different time-scales between SSI estimates and ground measurements for KIV.

		IMF1	IMF2	IMF3	IMF4	IMF5	IMF7	Glob.
HC3v5	R^2	0.926	0.924	0.929	0.921	0.885	0.992	0.984
	RMSE*	8.5	5.7	4.5	4.0	3.1	6.3	12.6
ERA	R^2	0.611	0.677	0.692	0.753	0.670	0.984	0.931
	RMSE*	13.9	8.7	6.5	5.2	4.5	8.4	24.5
MERRA2	R^2	0.621	0.643	0.671	0.800	0.590	0.982	0.934
	RMSE*	16.4	10.1	7.4	5.1	4.8	8.5	26.2

* RMSE has units of $\mathrm{W\,m^{-2}}$.

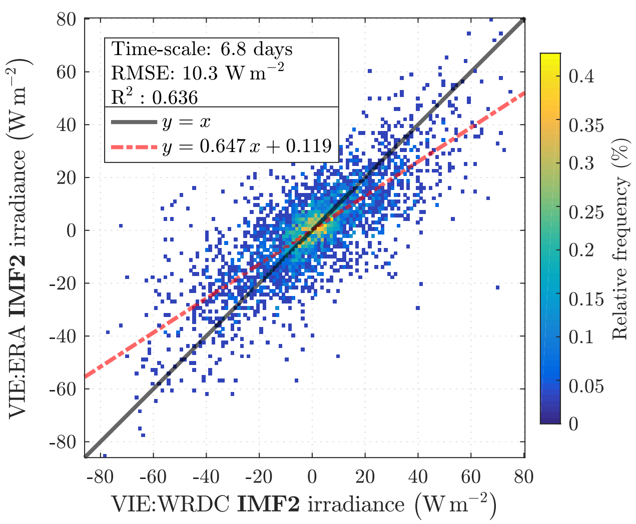

Figure 6. The 2-D histogram for IMF2 of ERA and WRDC for VIE. Each pixel encodes relative frequency according to the colourbar on the right. The solid black line denotes the identity line and the dash-dotted red line represents the best fit line. The linear regression equation is indicated in the legend. The time-scale, root-mean-square error and coefficient of determination are indicated in the panel above the legend.

ment. Results, in terms of root-mean-square error and coefficient of determination, are indicated in Tables 3 and 4 for VIE and KIV, respectively. IMF6 has been excluded, because Tables 1 and A1 show that for the re-analyses the time-scale of this mode deviates significantly from the ground measurements, thus the comparison is meaningless. IMF8…IMF10 have also been excluded because the mean amplitudes are below the uncertainty level (see Tables 2 and A2).

Table 3 shows that for VIE, on a per time-scale basis as well as globally, the closest estimate of ground measurements of the SSI is the HC3v5 dataset, both in terms of explained variance, and in terms of scatter. The lowest coefficient of determination for HC3v5 is 0.924 for the weekly variability (IMF2), meaning that Fig. 5 represents the worst case scenario for this particular dataset. The largest percentage of variance explained, 99.2 %, is attained for the yearly variability (IMF7). Globally, the HC3v5 accounts for 97.9 % of the observed variability (RMSE = 14.1 $\mathrm{W\,m^{-2}}$), outclassing ERA with 92.1 % (RMSE = 24.8 $\mathrm{W\,m^{-2}}$) and MERRA2 with 92.8 % (RMSE = 26.4 $\mathrm{W\,m^{-2}}$). For IMF1…IMF5, MERRA2 outperforms ERA in terms of R^2 except for the monthly variability (IMF4), as also reflected in the range of this coefficient of [0.684; 0.762] for MERRA2 as opposed to a range of [0.636; 0.755] for ERA. The yearly variability of the ground measurements is better expressed by ERA than by MERRA2, both in terms of coefficient of determination and RMSE. Generally, all the datasets exhibit monotonically decreasing RMSE for the first five modes and very good agreement for the yearly variability ($R^2 > 0.985$).

Similar statements can be made about the results for KIV, presented in Table 4. Here too, the HC3v5 dataset outperforms ERA and MERRA2 both in terms of coefficient of determination and RMSE, across all time-scales and also at the whole time-series level. The minimum R^2 for HC3v5 is 0.885 for IMF5, while the minimum values of R^2 for ERA end MERRA occur for IMF1 and are 0.611 and 0.621, respectively. As for VIE, all the datasets exhibit monotonically decreasing RMSE for the first five modes and very good agreement for the yearly variability ($R^2 > 0.982$).

4 Discussion

The results in the previous section have highlighted some features of the data that will be expanded upon here.

It has been inferred from the mean time-scales and mean amplitudes of the decomposed data (Tables 1 and 2) as as well as from the Hilbert marginal spectra (Figs. 3 and 4) that the yearly mode of variability is the most prominent feature of all the datasets. This result is unsurprising since both stations are situated at mid-latitude and it can be explained in terms of orbital geometry through the yearly cycle of seasons; it can also be inferred by visually inspecting the raw data from Fig. 1 which shows large variability with periodicity of one year. For WRDC, HC3v5, ERA and MERRA2, there is good agreement with respect to this mode (IMF7) both in terms of amplitude and of time-scale (see also Tables 3 and 4).

Apart from this yearly component, the TOA exhibits no other form of significant variability, also in good agreement with its trace from Fig. 1, which registers as an undisturbed sinusoidal waveform.

High-frequency variability, from 2 days up to 2–3 months, is manifest in McClear through its first five IMFs. This feature is also present in the rest of the datasets with greater power when compared with McClear. Hence, this feature can be attributed to clear-sky (no cloud) atmospheric effects (scattering and absorption by ozone, water vapour, aerosols, etc.). Looking at the McClear graph in Fig. 1, it becomes evident that this high-frequency variability manifests itself more strongly during the summer than during the winter months. In other words, the yearly cycle modulates the power of this high-frequency feature through a non-linear cross-scale amplitude-phase coupling. This feature is also apparent in the HC3v5, WRDC, ERA and MERRA2 datasets (see Supplement) and is in agreement with the findings of Bengulescu et al. (2016b) who underlined its stochastic nature. The time-scales for the individual modes composing this high-frequency feature agree well across these latter four datasets. In terms of amplitudes, however, for VIE HC3v5 slightly overestimates the WRDC measurements while ERA and MERRA2 underestimate more severely, while for KIV all the SSI estimates yield lower values than the ground data, although less so for HC3v5. This is also readily apparent in Tables 3 and 4, where for the first five IMFs, HC3v5 outperforms ERA and MERRA2 by a large margin. This result is a major contribution of the study, since Inman et al. (2013) have proposed that the accuracy of SSI forecasts crucially depends on the ability to forecast the stochastic component. As such, practitioners interested in modelling and forecasting the daily SSI are better off using satellite estimates that radiation products from re-analyses, at least for the two sites studied herein. This finding can be explained by the fact that re-analyses assimilate state variables such as temperature, moisture and wind, while the SSI is diagnostic. Stated differently, in re-analyses, radiation and cloud properties are derived from a model and, as such, they include the uncertainty of this model. Re-analyses often predict clear sky conditions while the actual conditions are cloudy (Boilley and Wald, 2015). HC3v5 is based on Meteosat imagery (Sect. 2.1) and as such directly takes clouds into account.

Another significant result of this study is the fact that, for both VIE and KIV, the McClear datasets do not have a variability component in between this high-frequency feature and the yearly cycle. In other words, IMF6 for the McClear data represents the yearly cycle, unlike the ground measurements or the SSI estimates, where IMF6 is an intermediate component before the yearly component represented by IMF7. This has first been discussed as a "variability gap" by Bengulescu et al. (2016a), when analysing a decennial dataset of daily means of SSI measured by BSRN ground station at Carpentras, France, that experiences clear-sky conditions for most of the year. Subsequently, Bengulescu et al. (2016b) have shown that this "variability gap" is also manifest for a similar dataset of ground measurements taken at Boulder, Colorado, USA, a location that also experiences a high number of days with clear skies. Hence, this study confirms the fact that, indeed, a clear-sky atmosphere does not introduce any spectral features in between 2–3 months and 1 year. Since the ground data for both VIE and KIV feature such spectral components, it can be concluded that these two locations do not experience so many cloud-free days and/or that they experience a lot of broken clouds conditions. Here too, ERA and MERRA2 are outperformed by HC3v5, since the time-scale of this IMF6 in the re-analyses datasets is greatly different form the true time-scale found in ground measurements, and which is accurately reflected by the satellite estimates.

Larger scales of variability (IMF8…IMF10) have been discarded from this analysis because of their failing to stand above the uncertainty threshold.

5 Conclusions and outlook

In this study we have investigated the characteristic time-scales of variability found in long-term time-series of daily means of SSI. We have also studied the fitness for use of satellite estimates of the SSI and radiation products re-analyses as alternatives to pyranometric ground measurements. The novelty of our work is the use of the adaptive, data-driven Hilbert–Huang Transform (HHT) to decompose the datasets into their distinct characteristic time-scales of variability before undergoing analysis.

We have shown that the TOA only presents variability at the one year time-scale. The clear-sky atmosphere introduces stochastic high frequency variability, from 2 days to 2–3 months, which exhibits non-linear cross-scale phase-amplitude coupling with the yearly cycle. This feature is also present, and amplified, in ground measurements, satellite estimates and re-analysis products. The fact that the cloud-free atmosphere does not introduce variability from 2–3 months to one year, i.e. the "variability gap" alluded to in previous studies, has been confirmed. It has also been shown that, HC3v5 outperforms ERA and MERRA2 by a large margin in terms of estimating the measured SSI, not only at a global, whole dataset level, but also on an per time-scale basis, and especially with respect to the stochastic variability component. This has implications on the forecast and modelling of the SSI, where satellite estimates should be preferred instead of re-analysis products. Our study, hence, refines the existing methodology to assess the fitness for use of surrogate SSI products, through an improved in-depth comparison of their local time-scales of variability.

A limitation of our study needs to be pointed out. Before carrying out the analysis, we have used the EMD on each time-series and have only compared modes with similar time-scales. That is, we have used the mono-variate version of the EMD, where mode alignment (identical time-scales for the IMFs across datasets) is not enforced. Nevertheless the non-alignment of modes is not to be considered a weakness of our approach. Because identical time-series will be decomposed into identical modes, by not enforcing similar time-scales across the modes of different datasets, changes in the time-scales of the modes (e.g. IMF6 of HC3v5 matches WRDC unlike ERA or MERRA2) also provide supplementary clues as to the fitness for use of the surrogate SSI datasets in lieu of ground measurements. Mode alignment can be enforced by more advanced, multi-variate versions of the EMD. Two such techniques are the noise-assisted multivariate empirical mode decomposition (NA-MEMD) introduced by Rehman et al. (2013) or the adaptive-projection intrinsically transformed multivariate empirical mode decomposition (APIT-MEMD) proposed by Hemakom et al. (2016). The latter method is particularly of interest since it is also able to deal with power imbalances and inter-channel correlations found in multichannel data. With NA-MEMD or APIT-MEMD, all the datasets would be treated in a unitary manner as a single multi-variate signal, thus mode alignment would be enforced. This would also enable the use of of more advanced descriptors, such as multi-scale measures suitable for multi-variate datasets and inter-component measures, e.g. intrinsic correlation, intrinsic sample entropy or intrinsic phase synchrony (Looney et al., 2015). However, the exercise is significantly more technical and is proposed as a future study.

Lastly we recognize the restrained geographical character of the study and, as a future exercise, we propose its extension to many more geographical locations and possibly including several different satellite estimates and re-analyses radiation products, in order to determine whether the findings reported herein also hold for different regions and for different SSI surrogates.

6 Code availability

The software used for this study, comprising general EMD and HSA routines is publicly available online, as follows.

The fast EMD routine used in this study is provided by Wang et al. (2014) and can be downloaded at: http://rcada.ncu.edu.tw/FEEMD.rar.

Methods pertaining to Hilbert spectral analysis are part of a general HHT toolkit provided by Wu and Huang (2009) and can be downloaded at: http://rcada.ncu.edu.tw/Matlabruncode.zip.

The code for the ICEEMD(AN) algorithm (Colominas et al., 2014) is provided by María Eugenia Torres on her personal webpage, and can be downloaded at: http://bioingenieria.edu.ar/grupos/ldnlys/metorres/metorres_files/ceemdan_v2014.m

7 Data availability

The data can be accessed as follows:

The ERA-Interim data set (ECMWF2009) can be accessed at: https://doi.org/10.5065/D6CR5RD9.

The MERRA2 radiation diagnostics M2T1NXRAD timeseries (GMAO2015) is available at: doi:10.5067/Q9QMY5PBNV1T.

TOA and McClear data from Copernicus Atmosphere Monitoring Service (Copernicus2015, Hoyer-Klick et al., 2015) can be retrieved at: http://www.soda-pro.com/web-services/radiation/mcclear.

The WRDC global radiation daily sums for Europe (WRDC2014) can be accessed at: http://wrdc.mgo.rssi.ru/wrdc_en.htm.

The HelioClim-3v5 dataset was downloaded from the SoDa Service web site (http://www.soda-pro.com) managed by the company Transvalor. Data are available to anyone for free for years 2004–2006 as a GEOSS Data-CORE (GEOSS Data Collection of Open Resources for Everyone) and for-pay for the most recent years with charge depending on requests and requester.

Appendix A: Mean IMF time-scales and amplitudes for KIV

Tables A1 and A2 present the mean time-scales and respectively the mean amplitudes of the IMFs for the KIV datasets.

Table A1. Mean IMF time-scales in days for the KIV datasets.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
TOA	6.5	9.3	16.0	31.8	344	366	389	796	3157	3407
McClear	2.9	6.9	12.6	25.1	54.9	366	387	1054	2981	3813
WRDC	3.1	6.9	13.3	26.9	49.0	199	369	959	1552	1760
HC3v5	3.1	6.9	13.3	27.4	49.7	176	370	726	1639	2680
ERA	3.1	6.9	13.3	27.2	48.3	280	370	1015	1540	3367
MERRA2	3.0	6.9	13.4	26.8	48.3	298	374	746	1658	3533

Table A2. Mean IMF amplitudes in $W\,m^{-2}$ for the KIV datasets.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
TOA	0.1	0.1	0.1	0.1	3.3	182	5.6	1.4	2.6	0.5
McClear	7.7	5.1	3.8	3.5	3.1	130	6.6	1.6	0.9	0.5
WRDC	42	25	20	18	12	22	103	6.6	5.4	0.6
HC3v5	40	24	19	17	11	19	103	7.4	6.6	1.2
ERA	29	18	14	13	9.3	31	91	4.9	4.5	2.3
MERRA2	36	21	16	14	9.3	41	89	5.5	5.9	0.6

The Supplement related to this article is available online at doi:10.5194/asr-11-35-2017-supplement.

Author contributions. All authors contributed equally to this work.

Competing interests. The authors declare no competing interests. HC3v5 solar radiation products are commercialized by Transvalor SA through its online SoDa Service at http://www.soda-pro.com, except for years 2004–2006 which are available to anyone for free as a GEOSS Data-CORE.

Acknowledgements. The authors thank the World Radiation Data Centre for maintaining the radiation archives and hosting the website for downloading data. They thank the ground station operators of the WMO network for their valuable measurements. The authors are indebted to the company Transvalor SA which is taking care of the SoDa Service for the common good, therefore permitting an efficient access to the HelioClim databases.

Edited by: S.-E. Gryning
Reviewed by: two anonymous referees

References

Beer, J.: Long-term indirect indices of solar variability, Space Sci. Rev., 94, 53–66, doi:10/d8pzkn, 2000.

Beer, J., Vonmoos, M., and Muscheler, R.: Solar variability over the past several millennia, Space Sci. Rev., 125, 67–79, doi:10/djw2kj, 2006.

Bengulescu, M., Blanc, P., and Wald, L.: On the temporal variability of the surface solar radiation by means of spectral representations, Adv. Sci. Res., 13, 121–127, doi:10/bp6g, 2016a.

Bengulescu, M., Blanc, P., and Wald, L.: On the intrinsic timescales of temporal variability in measurements of the surface solar radiation, Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-38, in review, 2016b.

Bengulescu, M., Blanc, P., and Wald, L.: Characterizing temporal variability in measurements of surface solar radiation and its dependence on climate, Energy Procedia, 97, 164–171, doi:10/bt46, European Geosciences Union General Assembly 2016, EGU Division Energy, Resources & the Environment (ERE), https://www.journals.elsevier.com/energy-procedia (last access: February 2017), 2016c.

Blanc, P. and Wald, L.: The SG2 algorithm for a fast and accurate computation of the position of the Sun for multi-decadal time period, Solar Energy, 86, 3072–3083, doi:10/2h9, 2012.

Blanc, P., Gschwind, B., Lefèvre, M., and Wald, L.: The HelioClim project: Surface solar irradiance data for climate applications, Remote Sensing, 3, 343–361, doi:10/chpk5b, 2011.

Boilley, A. and Wald, L.: Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface, Renew. Energy, 75, 135–143, doi:10/f259gv, 2015.

Calif, R., Schmitt, F. G., Huang, Y., and Soubdhan, T.: Intermittency study of high frequency global solar radiation sequences under a tropical climate, Solar Energy, 98, 349–365, doi:10/7xg, 2013.

CIMO – Commission for Instruments and Methods of Observation: Guide to Meteorological Instruments and Methods of Observation, Tech. Rep. WMO-No. 8, 2014 Edn., World Meteorological Organization, Geneva, Switzerland, http://tinyurl.com/CIMO2014v8 (last access: 14 December 2016), 2014.

Colominas, M. A., Schlotthauer, G., and Torres, M. E.: Improved complete ensemble EMD: A suitable tool for biomedical signal processing, Biomed. Sig. Proces. Cont., 14, 19–29, doi:10/7tp, 2014.

Duffy, D. G.: The application of Hilbert-Huang transforms to meteorological datasets, J. Atmos. Ocean. Tech., 21, 599–611, doi:10/b9td58, 2004.

Espinar, B., Ramírez, L., Drews, A., Beyer, H. G., Zarzalejo, L. F., Polo, J., and Martín, L.: Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations, Solar Energy, 83, 118–125, doi:10/ftmgqt, 2009.

Flandrin, P., Rilling, G., and Gonçalves, P.: Empirical mode decomposition as a filter bank, IEEE Sig. Proc. Lett., 11, 112–114, doi:10/c9g7q3, 2004.

Gabor, D.: Theory of communication. Part 1: The analysis of information, J. Inst. Elect. Eng.-Pt. III, 93, 429–441, doi:10/2rz, 1946.

Hemakom, A., Goverdovsky, V., Looney, D., and Mandic, D. P.: Adaptive-projection intrinsically transformed multivariate empirical mode decomposition in cooperative brain–computer interface applications, Philos. T. Roy. Soc. A, 374, 20150199, doi:10/bvjm, 2016.

Hoyer-Klick, C., Lefèvre, M., Schroedter-Homscheidt, M., and Wald, L.: User's guide to the MACC-RAD services on solar energy radiation resources, doi:10.13140/RG.2.1.5016.7521, 2015.

Huang, N. E. and Wu, Z.: A review on Hilbert-Huang transform: Method and its applications to geophysical studies, Rev. Geophys., 46, 1–23, doi:10/fr2hk2, 2008.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, P. Roy. Soc. Lond. A, 454, 903–995, doi:10/dr5ngd, 1998.

Huang, N. E., Chen, X., Lo, M.-T., and Wu, Z.: On Hilbert spectral representation: a true time-frequency representation for nonlinear and nonstationary data, Adv. Adapt. Data Anal., 3, 63–93, doi:10/b7m8xt, 2011.

Huang, N. E., Young, V., LO, M., WANG, Y. H., Peng, C.-K., Chen, X., Wang, G., Deng, J., and Wu, Z.: The uniqueness of the instantaneous frequency based on intrinsic mode function, Adv. Adapt. Data Anal., 5, 1350011, doi:10/bx8p, 2013.

Inman, R. H., Pedro, H. T., and Coimbra, C. F.: Solar forecasting methods for renewable energy integration, Prog. Energy Combust. Sci., 39, 535–576, doi:10/brhc, 2013.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter-Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J. W., and Morcrette, J.-J.: McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions, Atmos. Meas. Tech., 6, 2403–2418, doi:10.5194/amt-6-2403-2013, 2013.

Lefèvre, M., Blanc, P., Espinar, B., Gschwind, B., Ménard, L., Ranchin, T., Wald, L., Saboret, L., Thomas, C., and Wey, E.: The HelioClim-1 database of daily solar radiation at Earth surface: an example of the benefits of GEOSS Data-CORE, IEEE J. Select. Top. Appl. Earth Obs. Rem. Sens., 7, 1745–1753, doi:10/f259gx, 2014.

Lohmann, S., Schillings, C., Mayer, B., and Meyer, R.: Long-term variability of solar direct and global radiation derived from ISCCP data and comparison with reanalysis data, Solar Energy, 80, 1390–1401, doi:10/fjn6mc, 2006.

Looney, D., Hemakom, A., and Mandic, D. P.: Intrinsic multi-scale analysis: a multi-variate empirical mode decomposition framework, P. Roy. Soc. A, 471, 20140709, doi:10/bvjn, 2015.

Madronich, S. and Flocke, S.: The Role of Solar Radiation in Atmospheric Chemistry, in: Environmental Photochemistry, vol. 2/2L of The Handbook of Environmental Chemistry, edited by: Boule, P., Springer, Berlin, Heidelberg, 1–26, doi:10/f259g2, 1999.

Meftah, M., Dewitte, S., Irbah, A., Chevalier, A., Conscience, C., Crommelynck, D., Janssen, E., and Mekaoui, S.: SOVAP/Picard, a spaceborne radiometer to measure the total solar irradiance, Solar Physics, 289, 1885–1899, doi:10/bvqs, 2014.

Moghtaderi, A., Flandrin, P., and Borgnat, P.: Trend filtering via empirical mode decompositions, Comput. Stat. Data Anal., 58, 114–126, doi:10/c4nqzz, 2013.

Ohmura, A., Gilgen, H., Hegner, H., Müller, G., Wild, M., Dutton, E. G., Forgan, B., Fröhlich, C., Philipona, R., Heimo, A., König-Langlo, G., McArthur, B., Pinker, R., Whitlock, C. H., and Dehne, K.: Baseline surface radiation network (BSRN/WCRP): new precision radiometry for climate research, B. Am. Meteorol. Soc., 79, 2115–2136, doi:10/bvpw97, 1998.

Perez, R., Seals, R., and Zelenka, A.: Comparing satellite remote sensing and ground network measurements for the production of site/time specific irradiance data, Solar Energy, 60, 89–96, doi:10/fn3t4w, 1997.

Polo, J., Wilbert, S., Ruiz-Arias, J. A., Meyer, R., Gueymard, C., Súri, M., Martín, L., Mieslinger, T., Blanc, P., Grant, I., Boland, J., Ineichen, P., Remund, J., Escobar, R., Troccoli, A., Sengupta, M., Nielsen, K. P., Renne, D., Geuder, N., and Cebecauer, T.: Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets, Solar Energy, 132, 25–37, doi:10/bs5k, 2016.

Prša, A., Harmanec, P., Torres, G., Mamajek, E., Asplund, M., Capitaine, N., Christensen-Dalsgaard, J., Depagne, É., Haberreiter, M., Hekker, S., Hilton, J., Kopp, G., Kostov, V., Kurtz, D. W., Laskar, J., Mason, B. D., Milone, E. F., Montgomery, M., Richards, M., Schmutz, W., Schou, J., and Stewart, S. G.: Nominal Values for Selected Solar and Planetary Quantities: IAU 2015 Resolution B3, Astron. J., 152, 41, doi:10/bs5d, 2016.

Qu, Z., Gschwind, B., Lefevre, M., and Wald, L.: Improving HelioClim-3 estimates of surface solar irradiance using the McClear clear-sky model and recent advances in atmosphere composition, Atmos. Meas. Tech., 7, 3927–3933, doi:10.5194/amt-7-3927-2014, 2014.

Rehman, N., Park, C., Huang, N. E., and Mandic, D. P.: EMD via MEMD: multivariate noise-aided computation of standard EMD, Adv. Adapt. Data Anal., 5, 1350007, doi:10/3t6, 2013.

Richardson, D. B. and Andrews, R. W.: Validation of the MERRA dataset for solar PV applications, in: 40th IEEE Photovoltaic Specialists Conference (PVSC-40), 8–13 June 2014, at the Colorado Convention Center in mile-high Denver, Colorado, USA, 0809–0814, doi:10/bs5s, 2014.

Rigollier, C., Lefèvre, M., and Wald, L.: The method Heliosat-2 for deriving shortwave solar radiation from satellite images, Solar Energy, 77, 159–169, doi:10/bfrn55, 2004.

Roesch, A., Wild, M., Ohmura, A., Dutton, E. G., Long, C. N., and Zhang, T.: Assessment of BSRN radiation records for the computation of monthly means, Atmos. Meas. Tech., 4, 339–354, doi:10.5194/amt-4-339-2011, 2011.

Tary, J. B., Herrera, R. H., Han, J., and Baan, M.: Spectral estimation – What is new? What is next?, Rev. Geophys., 52, 723–749, doi:10/f259g3, 2014.

Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's global energy budget, B. Am. Meteorol. Soc., 90, 311–323, doi:10/d3xccg, 2009.

Tsvetkov, A., Wilcox, S., Renne, D., and Pulscak, M.: International solar resource data at the World Radiation Data Center, in: Proceedings of Solar '95: The 1995 American Solar Energy Society Annual Conference, 15–20 July 1995, Minneapolis, Minnesota, edited by: Campbell-Howe, R. and Wilkins-Crowder, B., American Solar Energy Society, Boulder, Colorado, USA, 1995.

Wang, G., Chen, X.-Y., Qiao, F.-L., Wu, Z., and Huang, N. E.: On intrinsic mode function, Adv. Adapt. Data Anal., 2, 277–293, doi:10/b466fq, 2010.

Wang, Y.-H., Yeh, C.-H., Young, H.-W. V., Hu, K., and Lo, M.-T.: On the computational complexity of the empirical mode decomposition algorithm, Physica A, 400, 159–167, doi:10/3nz, 2014.

Wu, Z. and Huang, N. E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal., 01, 1–41, doi:10/fdnzfq, 2009.

Wu, Z. and Huang, N. E.: On the filtering properties of the empirical mode decomposition, Adv. Adapt. Data Anal., 2, 397–414, doi:10/bscvxv, 2010.

Wu, Z., Huang, N. E., and Chen, X.: Some considerations on physical analysis of data, Adv. Adapt. Data Anal., 3, 95–113, doi:10/cftghv, 2011.

Zelenka, A., Czeplak, G., d'Agostino, V., Josefson, W., Maxwell, E., and Perez, R.: Techniques for supplementing solar radiation network data, in: Report No. IEA-SHCP-9D-1, vol. 1, International Energy Agency, http://tinyurl.com/kvc6hck (last access: 14 December 2016), 1992.

Zelenka, A., Perez, R., Seals, R., and Renné, D.: Effective accuracy of satellite-derived hourly irradiances, Theor. Appl. Climatol., 62, 199–207, doi:10/ct3pqh, 1999.

Zeng, Z., Yang, H., Zhao, R., and Meng, J.: Nonlinear characteristics of observed solar radiation data, Solar Energy, 87, 204–218, doi:10/bcr8, 2013.