Meteorologische Zeitschrift, PrePub DOI 10.1127/metz/2017/0836

Performance Evaluation of Radiation Sensors for the Solar Energy Sector

Laurent Vuilleumier[1*], Christian Félix[1], Frank Vignola[2], Philippe Blanc[3], Jordi Badosa[4], Andreas Kazantzidis[5] and Bertrand Calpini[1]

[1]Federal Office of Meteorology and Climatology, MeteoSwiss, Payerne, Switzerland
[2]Department of Physics, University of Oregon, Eugene, Oregon, USA
[3]MINES ParisTech, PSL Research University, O.I.E. Centre Observation, Impacts, Energie, Sophia Antipolis, France
[4]Laboratoire de Météorologie Dynamique, Ecole Polytechnique, Palaiseau, France
[5]Laboratory of Atmospheric Physics, Physics Department, University of Patras, Greece

(Manuscript received October 26, 2016; in revised form May 2, 2017; accepted May 29, 2017)

Abstract

Rotating Shadowband Irradiometers (RSI) and SPN1 Sunshine Pyranometers allow determining the diffuse and direct components of solar radiation without sun trackers; they can be deployed in networks for continuous field operation with modest maintenance. Their performances are evaluated here by analyzing their errors with respect to well characterized references. The analysis is based on 1-minute data recorded over a 15-month period at the Payerne BSRN station in Switzerland. The analysis was applied both to the whole dataset and data subsets reflecting particular conditions to allow a better understanding of how instrument performance depends on such conditions.

The overall performance for measuring global horizontal irradiance (GHI) is satisfactory with deviations compatible with an expanded uncertainty of ±25 Wm^{-2} (±10 %). For diffuse horizontal irradiance (DfHI), RSIs exhibited errors on the order of ±20 Wm^{-2} (±13 %) with some of them being affected by small systematic negative biases on the order of −5 Wm^{-2} (median). SPN1s underestimate DfHI by about −10 Wm^{-2} (median) with a relatively large range of the expanded error distribution between −45 Wm^{-2} and 20 Wm^{-2} (−35 % to 13 %). For direct normal irradiance (DNI), the extended error range for RSIs is on the order of ±40 Wm^{-2} (±5–6 %) with some instruments presenting no bias while others are affected by median biases up to −15 Wm^{-2}. SPN1s exhibit a relatively large median bias of 40 Wm^{-2}, and an extended range of the error distribution between −45 Wm^{-2} and 125 Wm^{-2} (−6 % to 19 %). Typical errors on the integrated yearly energy per unit surface area are on the order of a few percent or less (< 5 %) for RSI with negligible errors on DNI for some RSI instruments. SPN1 integrated errors are negligible for GHI, but on the order of −8 % for DfHI, and between 9 % and 11 % for DNI.

For RSIs, GHI and DfHI errors showed similar amplitude and dependence on solar elevation, while DNI errors were significantly smaller in relative terms than GHI or DfHI errors. This suggests that RSIs are optimized for providing good estimates of DNI, at the expense of – and resulting in – a correlation between GHI and DfHI errors. RSI uncertainty for DNI is about twice the uncertainty of a good quality pyrheliometer under favorable conditions. SPN1 instruments exhibit the opposite behavior with GHI and DfHI errors of opposite signs, resulting in large DNI errors. While the SPN1 performances for measuring GHI are similar to those of RSI, corrections are required to obtain satisfactory performances for DNI.

Keywords: shortwave radiation, accuracy, uncertainty, pyranometer, rotating shadowband irradiometer, RSI, SPN1

1 Introduction

Surface solar irradiance is a parameter of key importance for understanding and monitoring the global climate system and was identified as an Essential Climate Variable within the Global Climate Observing System (GCOS). Of course, it is also the key parameter for solar energy, and its high variability is an issue for solar energy production. This variability should be monitored, assessed and if possible forecasted. For this purpose, estimates of solar irradiance are required, and the only straightforward way to measure irradiance at the surface is by ground-based observation (World Meteorological Organization, 2010). These estimates are required for solar resource assessment and monitoring, for validation (Gueymard, 2014) or calibration (Vernay et al., 2013) of estimates derived from satellite data or Numerical Weather Predictions (NWP) models, and sometimes for very-short term forecasting (nowcasting – Chaabene and Ben Ammar, 2008; Yang et al., 2013).

*Corresponding author: Laurent Vuilleumier, Federal Office of Meteorology and Climatology, MeteoSwiss, ch. de l'Aérologie 1, 1530 Payerne, Switzerland, e-mail: laurent.vuilleumier@meteoswiss.ch

Table 1: List of tested instruments.

	Acronym used and name of instrument as indicated by provider	Provider	# instr.	Principle
RSI	IRR: Rotating Shadowband Radiometer	Irradiance Inc.	2	Rotating shadowband + LI-COR
	CSPA: Rotating Shadowband Irradiometer	CSP Services GmbH	2	Rotating shadowband + 2 LI-CORs (prototype)
	CSPB: Rotating Shadowband Pyranometer	CSP Services GmbH	2	Rotating shadowband + LI-COR
Δ-T	SPN1: Sunshine Pyranometer	Delta-T Devices Ltd.	3	Seven thermopile sensors + elaborate fix shading radiation shield

Knowledge of the accuracy of ground-based measurements is therefore crucial. Accuracy targets have been published for *reference* networks such as the Baseline Surface Radiation Network (BSRN – McArthur, 2005), which follows strict measurement guidelines ensuring high-quality data of known accuracy. But such stations are impractical for networks dedicated to solar energy assessments because of their cost and high level of maintenance required. Separate determination of the solar direct and diffuse radiation components is nonetheless desired for assessing the solar energy input onto collection devices because of their varied configurations, orientation and tilt. While this is achieved using pyrheliometers and pyranometers mounted on sun trackers at reference stations, alternative instruments have been developed by which diffuse and direct components of solar radiation may be inferred without sun trackers. These instruments are compact, lightweight with low power consumption and do not need mechanical adjustments of the shading device. They operate in a robust and cost effective way, and can be deployed in networks for continuous field operation with limited maintenance. Typically, RSI systems (including data logger) are designed to be powered with a 10 W solar panel, while sun trackers alone need more than 20 W to be operated (e.g., SOLYS 2 from Kipp & Zonen). They measure the global horizontal irradiance (GHI) and its diffuse component (DfHI) either in rapid succession (RSI) or simultaneously (SPN1). The direct irradiance is inferred from the difference between GHI and DfHI. These instruments are of two basic designs, the Rotating Shadowband Irradiometers (RSI, also sometimes called Rotating Shadowband Pyranometer – RSP) from different manufacturers, and the SPN1 radiometer from Delta-T Devices Ltd.

The accuracy of irradiance measurements has been assessed in different studies. They include uncertainty assessments (Myers et al., 1989; Myers, 2005; Reda, 2011; Vuilleumier et al., 2014), and studies reporting results of inter-comparison campaigns (Michalsky et al., 2003; 2011; Dutton and Long, 2012; Wilcox and Myers, 2008; Habte et al., 2016). Only one study conducted at the US National Renewable Energy Laboratory (NREL – Wilcox and Myers, 2008; Habte et al., 2016) evaluated the accuracy of both RSI and SPN1. Thus, the performance of RSI or SPN1 instruments has rarely been evaluated by comparison to reference instruments for standard operational conditions. The Payerne BSRN station was proposed as a test bed for conducting such an evaluation over a 15 month period (16.06.2012 to 15.09.2013, 457 days in total). This BSRN station is part of the Payerne meteorological observatory in the western part of the Swiss Plateau between the Jura Mountains and the Alps (46.812° N, 6.942° E, 491 m above sea level).

The results of the performance evaluation comparing RSI and SPN1 radiometers to high accuracy radiation sensors (references) from the Payerne BSRN station are presented here. The reference instruments are traceable to the World Radiometric Reference, and their accuracy has been assessed by Vuilleumier et al. (2014).

2 Instrument descriptions

2.1 Tested instruments

Three different RSI models, as well as the SPN1 radiometer were evaluated. Two instruments of each RSI model and three SPN1 instruments were tested (Table 1). These instruments were installed on a test field at the Payerne BSRN station and operated as recommended by the providers. If using corrections is part of the instrument standard operation, we used them; but in no case were such corrections determined with knowledge of the reference data.

A majority of RSI models have a rotating shadowband (Fig. 1) that makes a rapid rotation over the sensor for determining DfHI; typically the sensor is only shaded during a fraction of a second. This requires a sensor with a very fast time response such as the LI-COR photodiode sensor (LI-COR, Inc., Lincoln, Nebraska, USA). The RSIs used in this study all included LI-COR photodiode sensors. Neglecting corrections described in the next paragraphs, their basic principle is using the minimum irradiance measured during the rotation of the shadowband as an estimate of DfHI, determining GHI from the average reading taken when the shadowband is far from shading the diffuser and inferring the direct normal irradiance (DNI) from $(\mathrm{GHI} - \mathrm{DfHI})/\sin(\xi_s)$, where ξ_s is the solar elevation angle. The main factors affecting the RSI performance are the non-uniform spectral response of the LI-COR sensor, the non-ideal (non-Lambertian) directional response of the sensor's light

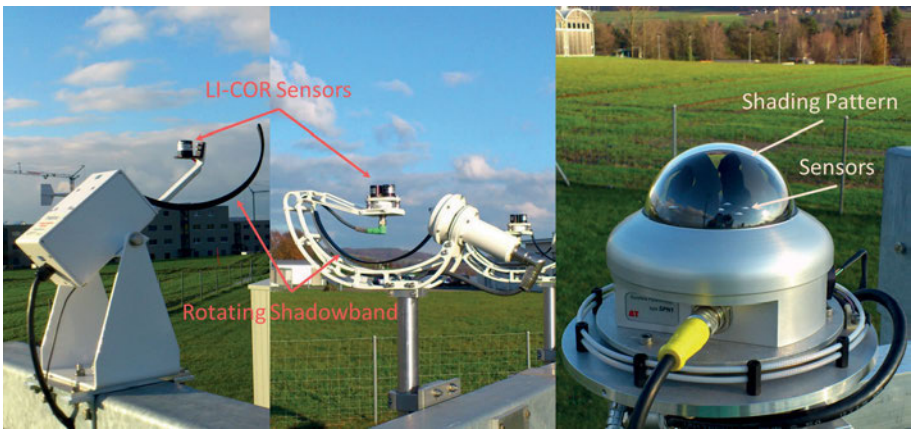

Figure 1: Rotating Shadowband Irradiometers (left: CSPA; center: CSPB) and SPN1 Shading Pattern Pyranometer (right).

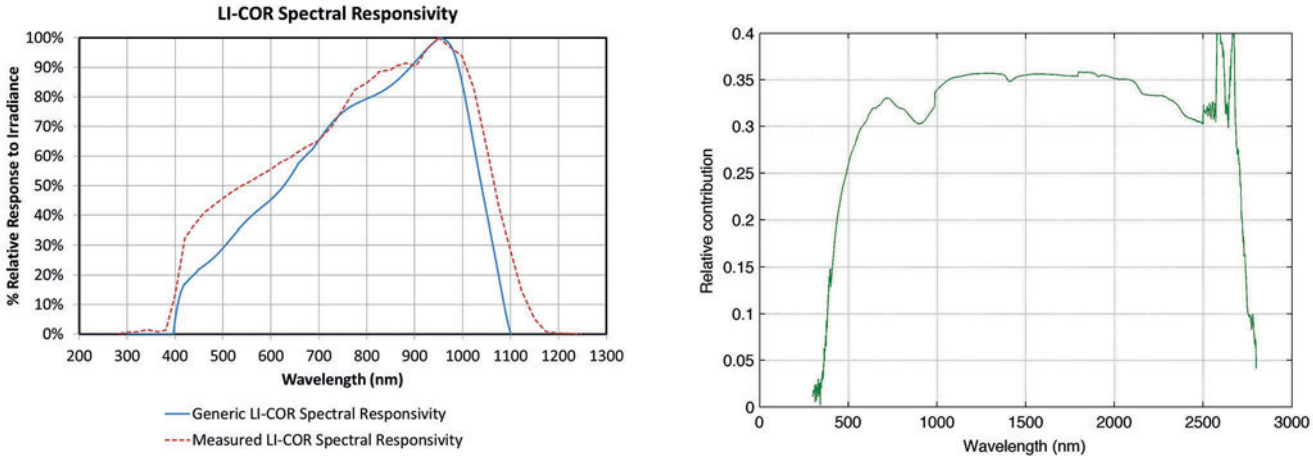

Figure 2: Left: Typical spectral responsivity of a LI-COR pyranometer. Solid blue line: spectral responsivity published by LI-COR. Dashed red line: spectral responsivity of one LI-COR pyranometer measured at NREL. Right: Typical spectral responsivity of a SPN1 sunshine pyranometer.

diffusing aperture window, and the temperature dependence of the sensor.

LI-COR sensors are mainly sensitive to wavelengths between ~ 400 nm and ~ 1150 nm with a maximum sensitivity between 950 nm and 1000 nm (Fig. 2). Outside these limits, the sensitivity is very low. This spectral range is a fraction of the range of the standard thermopile sensor, typically from 300 nm to 4000 nm, and this non-uniform spectral response results in a broadband responsivity that depends on the spectral distribution of the incident irradiance. The global, direct and diffuse radiation spectra are different, and they depend on sun elevation, ξ_s, the atmospheric composition and cloud cover (Michalsky et al., 1987; Vignola, 1999; Vignola et al., 2016).

The directional response of the LI-COR is also not Lambertian, especially at large incidence angles corresponding to low solar elevation (Michalsky et al., 1987; 1995; Vignola, 2006), which introduces a sensitivity dependent on the solar elevation angle when direct radiation is present. Since the effects caused by the spectral dependence of the pyranometer can also be modeled as a function of the solar elevation angle, modeled corrections for the deviation from the Lambertian response have often included some of the deviation caused by the spectral dependence of the instrument. Finally, the response of photodiodes in general, and of the LI-COR sensor in particular, is sensitive to the temperature (Michalsky et al., 1987; 1991; Vignola, 1999; 2006). If not regulated, the temperature of the sensor depends on many meteorological factors and particularly the solar irradiation itself.

Corrections for the above-mentioned dependences have been devised and improved over the years (Michalsky et al., 1987; 1991; Vignola, 1999; 2006; Augustyn et al., 2002; 2004; Geuder et al., 2008, 2010; Vignola et al., 2016). These corrections are often implemented in the firmware of RSIs and significantly improve their performance. Nonetheless, it is impossible, in practice, for such corrections to fully account for the local and complex dependences and smaller but non-negligible uncertainties and dependences remain after corrections. Such uncertainty sources often have a similar influence and may all induce effects that vary with

solar elevation angle. They are thus relatively difficult to distinguish.

The SPN1 Sunshine Pyranometer of Delta-T Devices Ltd. uses a radiation shield with an elaborate pattern and seven thermopiles (six on a hexagonal grid and one at the center) to measure the global solar radiation and its diffuse component (Fig. 1). This instrument does not have moving parts and is heated to keep the glass dome free of dew and help melting solid water or evaporating liquid water. The instrument's spectral response extends from 400 nm to 2700 nm and is relatively constant between 500 nm and 2700 nm (Fig. 2). It has a response time less than 200ms. The specific design of the shading pattern ensures that it masks half of the hemisphere for each thermopile and that, in any situation, at least one sensor is fully exposed to the direct solar beam (reading the maximum output $U_{\max}$) and another is completely shaded (reading the minimum output $U_{\min}$). Given the fact that the shading-pattern blocks half of the upper hemisphere and under the assumption of isotropic diffuse sky radiance, all sensors receive half of the diffuse irradiance (Badosa et al., 2014). DfHI is thus assumed to be $2U_{\min}$, GHI $= U_{\max} + U_{\min}$, and DNI can be inferred with DNI $= (U_{\max} - U_{\min})/\sin(\xi_s)$. More detailed technical information, including extra corrections applied to the readings, can be found in the User Manual (Wood, 2012). For this study, GHI and DfHI measurements were taken from the instrument outputs, and DNI values were calculated from (GHI – DfHI)/ $\sin(\xi_s)$.

A comprehensive study of the uncertainty sources affecting SPN1 pyranometers was conducted by Badosa et al. (2014). That study concluded that while GHI measurements by SPN1 and reference pyranometers compare relatively well, DfHI measurements are systematically underestimated, and the resulting DNIs are consequently overestimated. The main uncertainty sources come first from the geometry of the instrument (sensor placements with respect to the dome and the shading pattern as described in the next paragraphs), and second from the SPN1 spectral response.

More precisely, the geometry of the SPN1 is a source of errors because the working definition of DNI includes the radiance up to 2.5° from the sun, based on the half-opening angle recommended by the WMO CIMO guide (World Meteorological Organization, 2010). Conversely, for consistency, DfHI is defined as the irradiance from the upper hemisphere excluding the radiance in the DNI. For the SPN1, the region around the sun that is shaded for the sensor recording $U_{\min}$ is larger than a cone with a 2.5° opening half-angle. Badosa et al. (2014) computed SPN1 "first touch angles", which imply a notion similar to the opening angle. They found that depending on the sensor and the sun position in the sky, this "effective" opening half-angle is always more than 5°, with a maximum value of 25° for low solar elevation. This often results in a bright region around the sun being excluded from DfHI, which contradicts the assumption of isotropic diffuse radiance behind DfHI $= 2U_{\min}$. DfHI will be underestimated by the SPN1 in such cases. This error depends on the position of the sun in the sky and on the circumsolar radiance.

Another geometric uncertainty source affects the direct beam component. Because six of the sensors are not located at the geometrical axis of the glass dome, the direct beam is bent by the glass dome due to a lensing effect. This effect is predictable; in addition unpredictable sensor mismatches (e.g., differences in the calibration of the different sensors or imprecision in the leveling of the sensors) can also lead to uncertainties and abrupt changes when switching from one sensor to the other for $U_{\max}$ as the subset of sensors exposed to beam radiation change with time. These uncertainties can reach up to $\pm 15\,\mathrm{Wm}^{-2}$ when DNI is high (Badosa et al., 2014).

Fig. 2 shows the SPN1 spectral response, which has a higher sensitivity in the near infrared but falls off at shorter wavelengths (blue part of the visible spectrum). The sensitivity of the SPN1 is consequently different for clear-sky DfHI and GHI measurements because of the predominantly blue weighting of the diffuse spectrum in such cases. In addition, the difference in clear-sky sensitivity for GHI and DfHI will change with the solar elevation angle, because of changes in direct and diffuse irradiance spectra.

2.2 Reference instruments

The reference data for GHI, DfHI and DNI are the data collected by the Payerne BSRN station. The direct and diffuse components are monitored separately in addition to the global irradiance. All measurements are performed using secondary standards or first-class instrument. GHI and DfHI are measured by three, respectively two instruments (redundancy), while DNI is measured by a pyrheliometer constantly compared to a PMO6 absolute cavity radiometer (see Table 2). The PMO6 absolute open cavity radiometer is a secondary standard providing traceability to the World Radiometric Reference (WRR); this allows permanently checking the stability of the reference instruments as described by Vuilleumier et al. (2014). Table 2 lists the reference instruments used during the study period. The direct and diffuse component measurements are carried out using computer-controlled sun-trackers and collimation or shading devices. Vuilleumier et al. (2014) investigated the accuracy of the reference measurements taken during this performance evaluation. They found that GHI and DfHI expanded uncertainties are less than 1.8 %, while DNI uncertainty is on the order of 1.5 %. They describe in detail the measurements taken as reference values for this study.

The uncertainties determined by Vuilleumier et al. (2014) for reference instruments are a fraction – approximately by a factor five – of the uncertainties of the tested instruments, and the differences between reference and tested instrument measurements should therefore be strongly dominated by tested instrument uncertainties. However, reference instrument uncertainties linked to systematic errors could have a non-negligible

Table 2: List of reference instruments.

	Instrument	Serial #	Calibration	
			Period	Sensitivity
Direct	Kipp & Zonen CHP1	110740	Nov-2011 Kipp & Zonen	$7.88\,\mu V/(Wm^{-2}) \pm 0.09\,\mu V/(Wm^{-2})$
	PMOD/WRC PMO6	891002	Sep/Oct-2011 PMOD/WRC	$135.39\,(m^2 \cdot \Omega)^{-1} \pm 0.06\ (m^2 \cdot \Omega)^{-1}$
Diffuse	Kipp & Zonen CMP22	80001	Sep-2011 PMOD/WRC	$8.74\,\mu V/(Wm^{-2}) \pm 0.06\,\mu V/(Wm^{-2})$
	Kipp & Zonen CM21	61653	Sep-2011 PMOD/WRC	$11.83\,\mu V/(Wm^{-2}) \pm 0.16\,\mu V/(Wm^{-2})$
Global	Kipp & Zonen CMP22	80002	2008 Kip & Zonen	$9.40\,\mu V/(Wm^{-2}) \pm 0.10\,\mu V/(Wm^{-2})$
	Kipp & Zonen CM21	51436	Nov-2010 MeteoSwiss	$10.52\,\mu V/(Wm^{-2}) \pm 0.12\,\mu V/(Wm^{-2})$
	Kipp & Zonen CM21	41306	Nov-2010 MeteoSwiss	$10.73\,\mu V/(Wm^{-2}) \pm 0.12\,\mu V/(Wm^{-2})$

effect on our results. Specifically, a systematic error in the reference measurements may induce biases in the differences between reference and tested instrument measurements and affect the median and means of the difference distributions in a non-negligible way. The possible sources of reference instrument uncertainties are thoroughly investigated by Vuilleumier et al. (2014), and the following may lead to systematic uncertainties: calibration errors, thermal offset (for DfHI and GHI) and directional errors (for GHI).

In the framework of this performance evaluation, the fact that the reference pyrheliometer was continuously compared and eventually adjusted to an absolute cavity radiometer, while the pyranometers measuring GHI and DfHI were linked to the pyrheliometer measurements with a method similar to the continuous Sun and shade method for pyranometer calibration allowed substantially reducing the risk of a systematic calibration error. The expanded uncertainty on the reference instrument sensitivity factors was estimated to be 1 % for DNI and GHI, and 1.4 % for DfHI. Thermal offsets are known to affect thermopile-based pyranometers (Dutton et al., 2001; Ji and Tsay, 2000). These are almost always negative and may induce a systematic error. However, the correction method proposed by Dutton et al., (2001) was applied to the reference data, which strongly reduces the risk of systematic errors linked to thermal offsets (Dutton et al., 2001; Michalsky et al., 2003). The corresponding expanded uncertainty for the reference instruments is estimated at $2\,Wm^{-2}$ for DfHI and $4\,Wm^{-2}$ for GHI (Vuilleumier et al., 2014). Directional errors are linked to the non-Lambertian response of pyranometers. They mainly affect GHI because the diffuse radiance has a much more isotropic distribution and DfHI is less affected by deviations from a Lambertian response in pyranometers. Directional errors should not result in systematic errors in general, but when the analysis is restricted to specific solar elevations, systematic errors are possible. Vuilleumier et al. (2014) investigated such a problem by repeating their study of reference GHI uncertainty for four different solar elevation bands, and they found results agreeing within 1 %.

3 Method

The performance of the tested instruments is evaluated by studying the differences between their measurements and the reference ones. It is the nature of measurements to have a distribution around the unknown true value described as the measurement uncertainty. Because the uncertainty of the reference data is known to be small, the difference between test instrument and reference data will be defined as "error". If the average of the distribution is different from the true value, it is said to have a bias. This bias may be dependent on experimental conditions such as temperature, solar spectral distribution, or solar elevation. It is possible to develop algorithms to account for systematic errors and this has been done with the instruments being tested. The use of such algorithms can be problematic as various systematic errors are often difficult to separate from each other and from the random measurement uncertainties. Therefore if the adjustments were developed at one location under a given set of circumstances, the adjustment algorithm needs to be tested at other locations under different sets of circumstances to ensure that they appropriately reduce or eliminate the systematic error. The tests at Payerne are designed to evaluate the performance of the RSI and SPN1 instruments with their adjustment algorithms under conditions experienced at Payerne. The evaluation in this study used 1 min data. For reference instruments, these data are 1-min averages of 1 Hz sampling. For the tested instruments, measurements are 1-min averages except

for the DfHI from RSIs for which different sampling strategies are used (for a description of instrument sampling strategies see Appendix A1 in the Supplementary Material).

Standard performance indicators such as bias, root mean square error (RMSE) and mean absolute error (MAE) derived from the differences between tested and reference instruments are used (see Section 3.1). We also consider quantiles (median, quartiles, etc.) of the errors represented by boxplot diagrams for a better understanding of the error distributions. In addition, we divided the dataset in subsets using criteria (see Section 3.2) differentiating conditions that could influence the performances of the instrument as suggested by the studies mentioned in Section 2. These criteria use the solar elevation (low vs. high solar elevation) and the short-term variability of the DNI (low vs. high variability), the latter being influenced by the clouds. Studying the subsets independently facilitates the understanding of the dependences of the instrument performance on these conditions.

However, the analysis described above provides only limited information on the time scales of the variability. A further step is to analyze the "time frequency signature" of the tested instruments compared to the reference ones, which may differ notably because of their different response time. We investigated this using power spectral density analysis (see Sections 3.3 and 4.5). This also enables assessment of the influence of using longer time scale aggregation on the agreement between tested and reference instruments.

3.1 Estimators for the statistical analysis

In general, the statistical estimators are calculated using the difference (Δ) between tested and reference data: $\Delta = I_{\text{test}} - I_{\text{ref}}$, where I_{test} is the irradiance (GHI, DNI or DfHI) measured by the instruments under test at a given time, while I_{ref} is the corresponding irradiance measured by the reference instrument. Although the true value of GHI, DNI or DfHI is unknown, the uncertainty of the reference instruments is known to be small (Vuilleumier et al., 2014), and we will thereafter consider the reference data as the true values and consider Δ as the error. We define the average bias as $\frac{1}{n}\sum_{i=1}^{n}\Delta_i$, the Mean Absolute Error (MAE) as $\frac{1}{n}\sum_{i=1}^{n}|\Delta_i|$ and the Root Mean Square Error (RMSE) as $\sqrt{\frac{1}{n}\sum_{i=1}^{n}\Delta_i^2}$. In addition, the α-quantile (p_α) is the data value for which α % of the (ordered) errors are smaller than the value p_α, the median, lower and upper quartiles being p_{50}, p_{25} and p_{75} respectively. The Interquartile Range (IQR) is then defined as $p_{75} - p_{25}$. Following a standard boxplot analysis convention, we define as outlier any observation deviating outside the domain defined by $p_{25} - 1.5\text{IQR}$ to $p_{75} + 1.5\text{IQR}$ corresponding to approximately $\pm 2.7\sigma$ and 99.3 % coverage for a normal distribution.

The quantities defined above were computed for the whole dataset and for different subsets. The changes in the error distributions with time were first analyzed by disaggregating the dataset into weekly subsets, computing the statistical indicators for the subsets and verifying how the indicators evolved during the inter-comparison. In a further step, the statistical indicators of the difference distributions were studied using the boxplot analysis, which was applied on the whole dataset and subsets reflecting particular conditions (high or low solar elevation, high or low DNI variability). This was performed in order to verify whether the performance of a given instrument changed when changing the conditions; this was completed by a detailed analysis of the error distribution dependences on solar elevation and cloudiness. For GHI and DfHI, the time evolution, the boxplot analyses and the study of the dependences on solar elevation angle and cloudiness are conducted for data measured when the sun elevation is sufficiently high to avoid horizon effects (see Section 3.2 for criteria). In addition for DNI, very low irradiance data ($< 5\,\text{Wm}^{-2}$) are excluded, thus excluding cases when thick clouds obscure the sun and times of very low sun elevation.

3.2 Data selection

The analysis was performed on data measured when ξ_s was higher than 10° to avoid perturbations by the horizon and shadowing by two trees at sunrise during some periods of the year. The reference data flagged as suspicious by short term quality control (QC, semi-automated performed within two days) and longer term quality analysis (QA, performed over months) were also excluded from the analysis. The QC and QA procedures are described by Vuilleumier et al. (2014). Reference data validated by both QC and QA procedures represent 96–99 % of data collected for $\xi_s > 10°$, depending on the parameter (DNI, DfHI or GHI). In addition, a simplified version of the reference data automated QC procedure was also performed on the data from tested instruments for early detection of instrument failures so that erroneous data were not included in the performance analysis. The simplified QC for tested instruments singled out data for verification, which resulted in the exclusion of some data from the tested instruments, mainly snow events on 14.12.2012, from 19.01.2013 to 22.01.2013 and from 11.02.2013 to 14.02.2013, because of snow accumulation on RSIs and data recorded after 01.06.2013 for CSPA RSIs. In the latter case, erroneous data were identified in the beginning of June 2013, which was finally linked to humidity infiltrating the sensor head (CSPA instruments were prototypes; the follow-up model was redesigned to avoid such issues).

The data subsets reflecting particular conditions were selected according to the sun elevation and the DNI variability. Data collected at solar elevation $10° \le \xi_s < 30°$ were categorized as low solar elevation events, while high solar elevation events included data collected for $30° \le \xi_s < 67°$ (the maximum solar elevation at Payerne is approximately 67°). The DNI variability was assessed with the change between successive measurements of

Table 3: Subdivision of the reference dataset depending on the DNI variability and solar elevation.

	Low DNI variability	High DNI variability
Low solar elevation ($10° \leq \xi_s < 30°$)	8 %	38 % (difficult cases)
High solar elevation ($30° \leq \xi_s < 90°$)	14 % (favorable cases)	40 %

DNI by the reference pyrheliometer (K&Z CHP1, see Table 2). Low DNI variability events were selected as instants when DNI was greater than $30\,\mathrm{Wm}^{-2}$ and for which the difference between two successive 1 min measurements was lower than $10\,\mathrm{Wm}^{-2}$. The latter condition on the rate of change was requested to be true for at least 15 minutes for events to be characterized as low DNI variability. These conditions on DNI can be expressed as:

$$\mathrm{DNI}(t_i) > 30\,\mathrm{Wm}^{-2}$$
$$\max_{j\in[i,i+15\,\mathrm{min}]} \left|\mathrm{DNI}(t_{j+1}) - \mathrm{DNI}(t_j)\right| < 10\,\mathrm{Wm}^{-2}\ \forall\ t_j$$

This typically selects data measured when clouds do not obscure the sun during at least 15 minutes. It is not the same as clear-sky conditions where no clouds are present in the sky. The two sets of criteria divide the dataset into four subsets including from 8 % to 40 % of the total dataset as indicated in Table 3. The analysis and discussion is based mainly on the two most contrasting subsets: 1) high solar elevation / low DNI variability (stable sunny conditions with the sun high over the horizon), thereafter called "favorable" conditions, and 2) low solar elevation / high DNI variability (cloudy conditions and sun relatively low over the horizon), thereafter called "difficult" conditions.

3.3 Power spectral density analysis

As mentioned, the reference instruments and tested instruments have different time responses and this can change the nature of the error distributions that may have different "frequency signatures". A first approach to assess the performance of the different sensors in term of temporal frequency response is to distinguish the performances in situation of slowly or rapidly variable irradiance (see Section 3.2 and Table 3). This can be completed by a finer analysis of the frequency signatures using the power spectral density (PSD).

The PSD of a second-order stationary signal x, noted $S[x](f)$, is defined as the Fourier transform of its autocorrelation $\Gamma[x](f)$. Similarly, the cross-PSD of two second-order stationary signals x and y, noted $S[x,y](f)$, is defined as the Fourier transform of their cross-correlation $\Gamma[x,y](f)$. Here, the PSD and cross-PSD have been estimated using Welch's method for noise reduction (Welch, 1967).

In order to remove dominant non-stationary components related to the daily and annual cycles of solar radiation, the irradiance time series have been normalized by the corresponding European Solar Radiation Atlas clear-sky model predictions (Rigollier et al., 2000). This normalization transforms the irradiance time series into corresponding clearness index time series reflecting only the deviation of the measurements from a clear-sky model. This corresponds mainly to the influence of clouds and deviations of local aerosol from the aerosol climatology included in the model.

To compute the different PSD and cross-PSD, 347 days – out of the 457 available – with at least 8 hours for which the solar elevation angle is greater than 10° have been selected. Days with more than 15 min of invalid data within these 8 hours have also been removed. The number of remaining days ranges between 178 and 285, depending on the considered test instrument. The time series of clear-sky index for the remaining days have been then completed by linear interpolation, if necessary. The different PSD and cross-PSD have been computed with the Welch's method for each of these days and then averaged. If K_c^t and K_c^r are the clearness index time series from the test and the reference sensors respectively, the frequency-based indicators that we analyzed are the ratio of the clearness index variance ratio $R(f)$, the signal-to-noise ratio $SNR(f)$ and the correlation coefficient $CC(f)$:

$$R(f) = \frac{S\left[K_c^t\right](f)}{S\left[K_c^r\right](f)}$$
$$SNR(f) = \frac{S\left[K_c^r\right](f)}{S\left[K_c^t - K_c^r\right](f)}$$
$$CC(f) = \frac{\left|S\left[K_c^r, K_c^t\right](f)\right|}{\sqrt{S\left[K_c^t\right](f) S\left[K_c^r\right](f)}}$$

$R(f)$ represents the ratio of variance of the test time series to the reference one, for a given temporal frequency f. It should ideally be equal to one. $SNR(f)$ is the ratio of the variance of the reference time series to the variance of the error for a given temporal frequency. It should be as large as possible, and ideally infinite. $CC(f)$ is a measure of the temporal coherence between the tested and the reference time series with respect to frequency and should ideally be equal to one.

4 Results and discussion

The uncertainty of the tested instruments is assessed by comparing their measurements with those of references whose uncertainty was previously evaluated by Vuilleumier et al. (2014). To be clear, the uncertainty of the references also contributes to the errors presented for the tested instruments. But the uncertainties of the references are at least a factor 5 smaller than the errors reported here, for instance the 10 % GHI or DfHI expanded uncertainty reported below for the tested instruments.

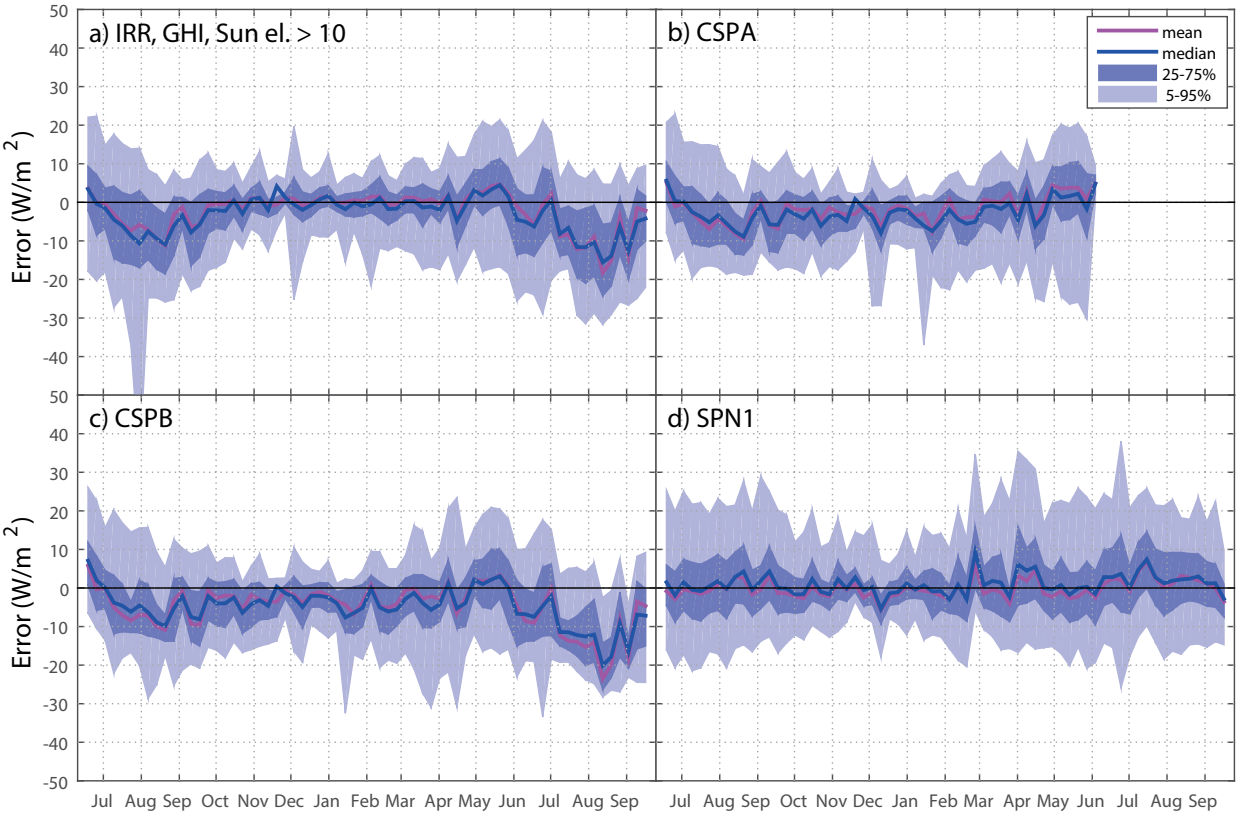

Figure 3: Mean, median, interquartile range and 5–95 percentile range of weekly error distribution between tested and reference instrument for GHI.

The mean, median, IQR and 5–95 percentile range of the error distributions are first presented as temporal series at a weekly resolution. This temporal resolution allows capturing seasonal effects, while ensuring that each subset includes enough data for statistical significance. After verifying that the error distributions do not show unexpected seasonal deviations, the statistical indicators are studied for the whole dataset, and for the subsets described in Section 3.2. The discussion is mainly based on results given in irradiance units, but it is also useful to present results in relative values. To avoid repetition, the main results are given here but corresponding figures for relative error distributions are given in Appendix A2 in the Supplementary Material.

4.1 Time series

Fig. 3 shows the GHI error distribution time series grouped by instrument model type. The error distributions stay most of the time within $\pm 20\,\mathrm{Wm}^{-2}$, with the 90 % range sometimes outside these limits in summer when GHI is stronger. For both summers (2012 and 2013), the RSIs exhibit negative bias on the order of 10–20 Wm^{-2} for the median or the mean, while the 90 % range extends as low as $-30\,\mathrm{Wm}^{-2}$. The relative error distributions (Appendix A2.1) stay within ±10 % (90 % range), except for some winter spikes due to low GHI values, especially for weeks with high cloudiness.

Corresponding to Fig. 3, Fig. 4 shows the error distribution time series for DfHI. The error distribution spreads for RSIs are smaller than for GHI, although the DfHI signal is usually smaller than GHI, and in relative value the DfHI spreads are larger (Appendix A2.1). For instruments CSPA and CSPB, a small systematic negative bias on the order of $-5\,\mathrm{Wm}^{-2}$, to $-10\,\mathrm{Wm}^{-2}$ is present. SPN1 instruments underestimate DfHI by an average $-5\,\mathrm{Wm}^{-2}$ to $-20\,\mathrm{Wm}^{-2}$ with a relatively large spread of the error distributions, the 90 % range sometime reaching as low as $-60\,\mathrm{Wm}^{-2}$.

Fig. 5 shows the same distributions for DNI. Because the spread of the error distributions for all types of instruments are significantly larger than for GHI or DfHI, Fig. 5 *y*-axis scale is different than for the preceding figures. For DNI, the IRR instruments present an underestimation by $10\,\mathrm{Wm}^{-2}$ to $30\,\mathrm{Wm}^{-2}$ in average, while the CSPA and CSPB instruments do not show bias in general. For these instruments, the distribution spread (90 % range) is of the order of $\pm 50\,\mathrm{Wm}^{-2}$. The SPN1 instruments exhibit a DNI overestimation by 25–80 Wm^{-2} in average, with the 90 % range reaching up to $150\,\mathrm{Wm}^{-2}$. The DNI is not a parameter provided by the SPN1 instrument, and we simply derived it from GHI and DfHI without corrections. Thus, since DfHI is underestimated by SPN1, DNI is overestimated.

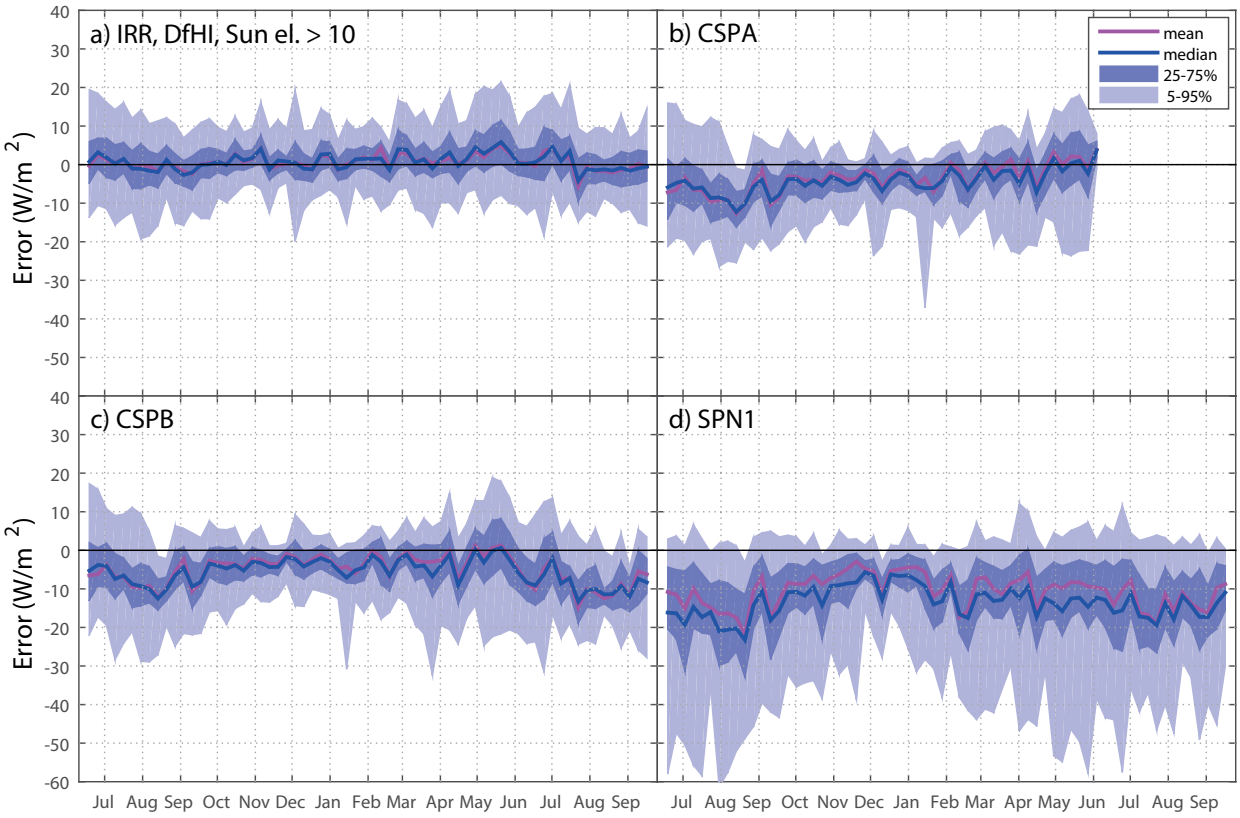

Figure 4: Same statistical indicators as given in Fig. 3, but for DfHI.

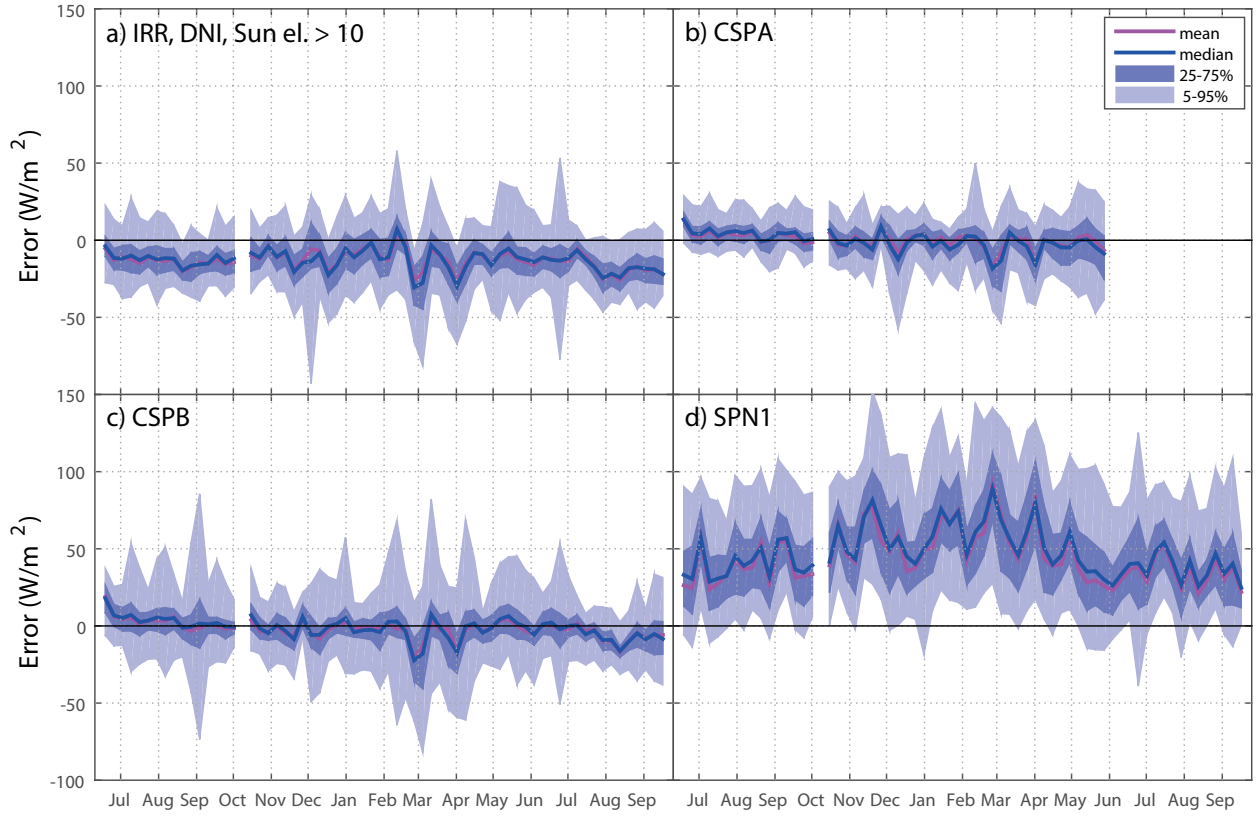

Figure 5: Same statistical indicators as given in Fig. 3, but for DNI.

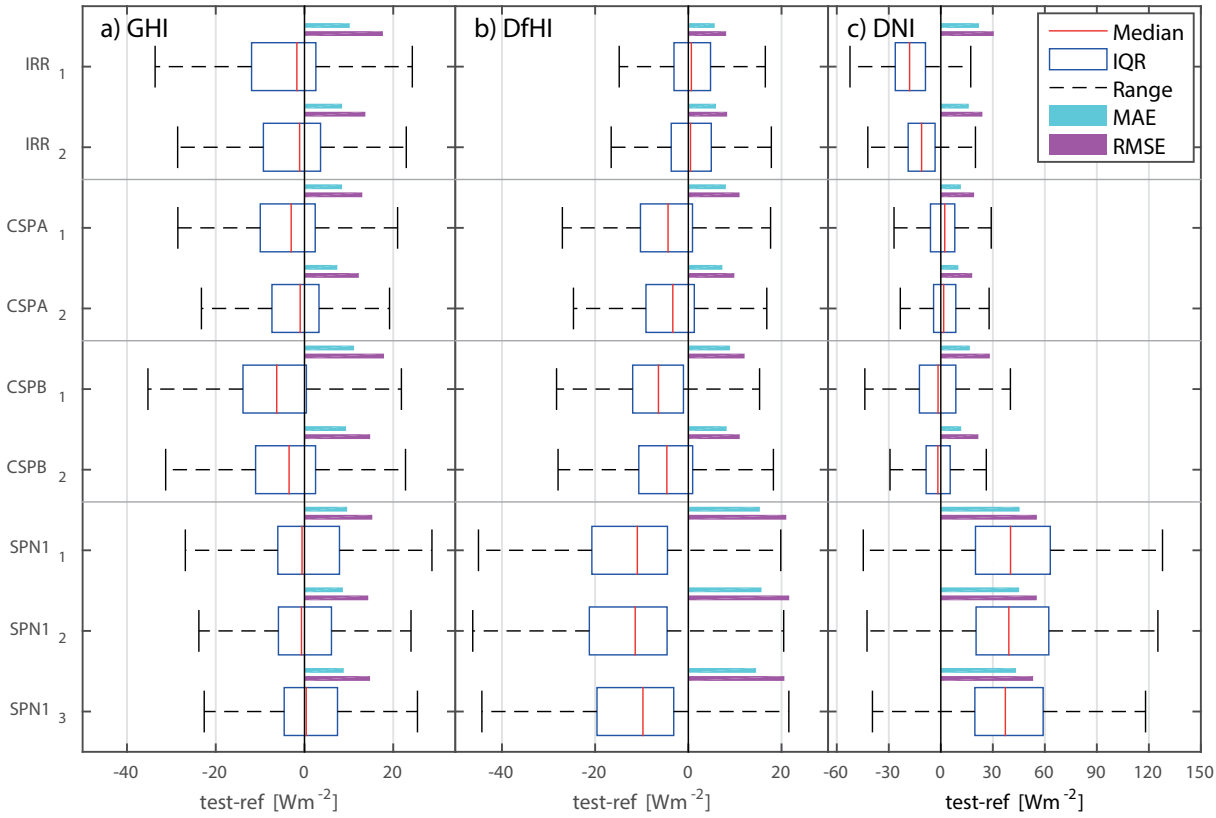

Figure 6: Boxplot diagram of the overall error distributions between measurements made by the tested and reference instruments. The boxplots show the median, IQR and the extended error range excluding outliers. In addition the RMSE and MAE are indicated. The statistical indicators are shown for a) GHI, b) DfHI and c) DNI, and are based on 1-min averages.

4.2 General performances

Fig. 6 shows statistical indicators for the error distributions presented as boxplot diagrams for all 1-min events fulfilling the selection described in Section 3.2. Each boxplot shows the median (red line), the IQR range (blue box), and the extended range of the distribution excluding outliers (whiskers, see Section 3.1). The latter range includes in all cases more than 95 % of the distributions. Beside each boxplots, the magenta and the cyan bars show the RMSE and MAE (Section 3.1).

With respect to GHI, RSI instruments and SPN1 instruments exhibit similar performance with error distributions ranging between $-35\,\mathrm{Wm^{-2}}$ and $25\,\mathrm{Wm^{-2}}$, which would correspond to an average expanded uncertainty of about $\pm 25\,\mathrm{Wm^{-2}}$ ($\pm 10\,\%$, Appendix A2.2). The DfHI uncertainties for the two types of instruments are different. For RSI, the errors are smaller than for GHI ($\sim \pm 20\,\mathrm{Wm^{-2}}$), but DfHI is usually also significantly smaller than GHI resulting in DfHI relative errors of about $\pm 13\,\%$. For some RSI, a small negative bias (about $5\,\mathrm{Wm^{-2}}$) is found. SPN1 instruments underestimate DfHI by about $-10\,\mathrm{Wm^{-2}}$ for the median of error (bias), with an expanded error distribution between $-45\,\mathrm{Wm^{-2}}$ and $20\,\mathrm{Wm^{-2}}$ (uncertainty range).

For DNI, the performance of RSI and SPN1 instruments also differ significantly. The extended range of the error distribution for RSI is of the order of $\pm 40\,\mathrm{Wm^{-2}}$ (± 5–$6\,\%$) with some instruments presenting no bias while others are affected by negative ones (the median can be down to $-15\,\mathrm{Wm^{-2}}$). SPN1 instruments exhibit a relatively large median bias on the order of $40\,\mathrm{Wm^{-2}}$ and an extended range of the error distribution between $-45\,\mathrm{Wm^{-2}}$ and $125\,\mathrm{Wm^{-2}}$ ($-6\,\%$ to $19\,\%$).

4.3 Performance for specific conditions

Fig. 7 shows the same error statistics as Fig. 6, but for favorable (sunny and high solar elevation, panels a, b, c) and difficult cases (cloudy and low solar elevation, panels d, e, f). When restricting the analysis to the subset of favorable cases, the RSIs exhibit GHI negative median biases (-22 to $-13\,\mathrm{Wm^{-2}}$); while in difficult cases the biases are much smaller or negligible. For these instruments, the behavior of DfHI biases is very similar. The DNI medians of errors are small: negligible for some instruments and of the order of $-20\,\mathrm{Wm^{-2}}$ for others. It should be noted that in favorable cases the instruments exhibiting the lowest DfHI biases are those showing the highest DNI biases. This suggests a bias compensation mechanism in the derivation of DNI. The more similar the GHI and DfHI biases are, the better such a bias compensation would be and the smaller the

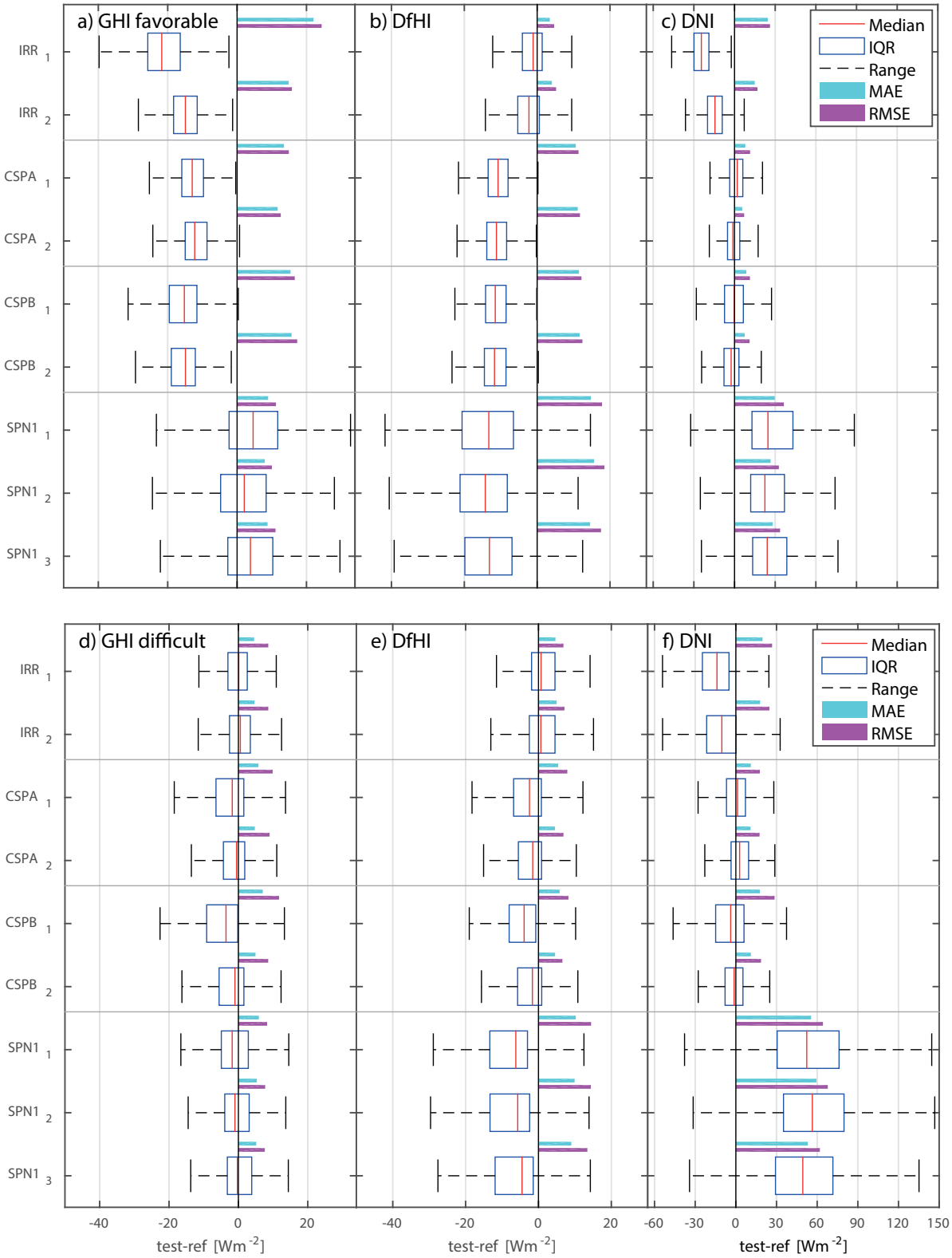

Figure 7: Boxplot diagrams of the error distributions between measurements made by the tested and reference instruments presented as boxplots for favorable cases (panel a, b and c for GHI, DfHI and DNI respectively) and for difficult cases (panel d, e and f for GHI, DfHI and DNI respectively) and for difficult cases (panel d, e and f for GHI, DfHI and DNI respectively). The statistical indicators shown are the same as for Fig. 6.

final DNI biases. For difficult cases, both the GHI and DfHI biases are small, and again very similar, leading to small DNI biases. The error ranges are of the same order for GHI and DfHI in both favorable and difficult cases (widths between 25 and 40 Wm^{-2} for extended ranges), resulting in significantly larger error ranges in relative values for DfHI for favorable cases. The DNI error extended ranges have widths between 45 and −55 Wm^{-2} for favorable cases and between 55 and 85 Wm^{-2} for difficult cases, leading to significantly smaller relative DNI errors for favorable cases (about 3–5 % vs. 8–15 %, Appendix A2.3).

The "bias behavior" of the SPN1 instruments is almost opposite to that of RSI and leads to an error amplification instead of compensation for DNI. They exhibit a small positive median of GHI error (< 5 Wm^{-2}, < 1 %), but a significant negative median of the DfHI error (about −15 Wm^{-2}, −15 %) for favorable cases, while they have negligible median of GHI error and small negative median of DfHI error (about −5 Wm^{-2}, −5 %) for difficult cases. For favorable cases, this results in a bias addition for the computed DNI, followed by an error amplification due to the presence of $\sin(\xi_s)$ value in the denominator. For difficult cases, the small DfHI biases lead to large DNI biases due to a strong $\sin(\xi_s)$ amplification for low solar elevation cases. The resulting median of DNI error are about 20–25 Wm^{-2} (3 %) and 50 Wm^{-2} (16 %), for favorable and difficult cases respectively, while the IQR and extended ranges are 30 Wm^{-2} and 120 Wm^{-2} for favorable cases, and 45 Wm^{-2} and 180 Wm^{-2} for difficult cases. Given the homogeneity of the SPN1 DNI error distributions, and their similar dependence on $\sin(\xi_s)$ (see Section 4.4), it could be appropriate to explore a $\sin(\xi_s)$ dependent correction directly for DNI. However, one still needs to explore how general these corrections would be since they may strongly depend on the location.

4.4 Influence of solar elevation and cloud cover

The results discussed above clearly show that the instrument performances are not uniform when the conditions change, the most likely causes being geometrical and spectral dependences. Fig. 8 shows the variations in error distribution median bias and IQR as a function of $\sin(\xi_s)$, which is approximately inversely proportional to the air mass, i.e. the amount of atmosphere traversed by the direct solar beam. Values at the left and right side of the plots correspond to low and high solar elevation angles, respectively. The upper part of the figure (panels a–i) are for GHI, while the lower one (panels j–r) are for DfHI. GHI and DfHI panels are divided in a first row (panels a–f and j–o) corresponding to results for RSI and a second row (panels g–i and p–r) for SPN1. The separation in two data subsets characterized with low DNI variability (blue curves – sunny situations, see Section 3.2) and high DNI variability (red, curves – cloudy situations) is kept while the data are distributed into 30 bins of variable width so that each bin holds the same number of data (the bin width varies from 1° to 3.5° for cloudy situations and from 1.6° to 2.8° for sunny situations). The median bias and IQR are subsequently computed for each bin.

Interpretation of Fig. 8 is not obvious for RSI instruments. Panels a-f consistently show a strong dependence of the GHI error with respect to $\sin(\xi_s)$ for sunny conditions with small or no error at low $\sin(\xi_s)$ and increasingly negative errors at higher $\sin(\xi_s)$. But in relative value (Appendix A2.4), the $\sin(\xi_s)$ dependence is less obvious and differs according to the model type. While it is present, but less strong for IRR instruments, it is not present for CSPA and CSPB instruments. This suggests that this error dependence may be simply due to the GHI signal increase with $\sin(\xi_s)$ in sunny conditions. The diffuse irradiance shows a similar pattern (see panel j–o), with the difference that for sensors of type IRR the error goes from positive values to negative values when $\sin(\xi_s)$ increases. In this case, the relative error also shows a pattern of increasingly negative error with increasing $\sin(\xi_s)$, and the error dependence on $\sin(\xi_s)$ is not only an effect of the irradiance strength for DfHI. The DfHI error dependence could thus be linked to a spectral effect: Because the sensor responsivity is not spectrally uniform and the related corrections may be not perfectly adapted to the particular conditions at the time and location of the measurement, a solar elevation error dependence can result from the changes in the diffuse irradiance spectrum related to solar elevation in sunny conditions.

This is further confirmed by Fig. 9 that presents the $\sin(\xi_s)$ dependence of the DfHI error distribution medians in a manner similar to Fig. 8, but for different cloud cover classes (CC) from 0 to 8 octas. The CC amount is determined every 10 minutes by the Automatic Partial Cloud Amount Detection Algorithm (Dürr and Philipona, 2004), which is based on long-wave measurements by the Swiss Meteorological Network. Fig. 9 shows that only overcast situations (CC = 7–8 octas) present dependences markedly different from the others, especially for CSPA and CSPB instruments. These situations are characterized by the absence of direct sun radiation and by a diffuse radiation with a spectrum relatively similar to this of GHI, i.e., *not* predominantly blue. For DfHI, the difference between these overcast situations and the others should thus be linked to the diffuse radiation spectrum.

There is a similarity between the dependence with respect to $\sin(\xi_s)$ of the error distribution for GHI and DfHI in favorable cases. But as discussed above, the GHI error dependence is likely linked to the increase of GHI with increasing $\sin(\xi_s)$, while for DfHI the most plausible cause is the non-uniform spectral responsivity of the RSI photodiode sensor. In this sense, this similarity may be a coincidence. However, multiple corrections are implemented in the firmware of RSI to account for the non-uniform response of the LICOR sensor as well as other error sources (Section 2.1). These corrections

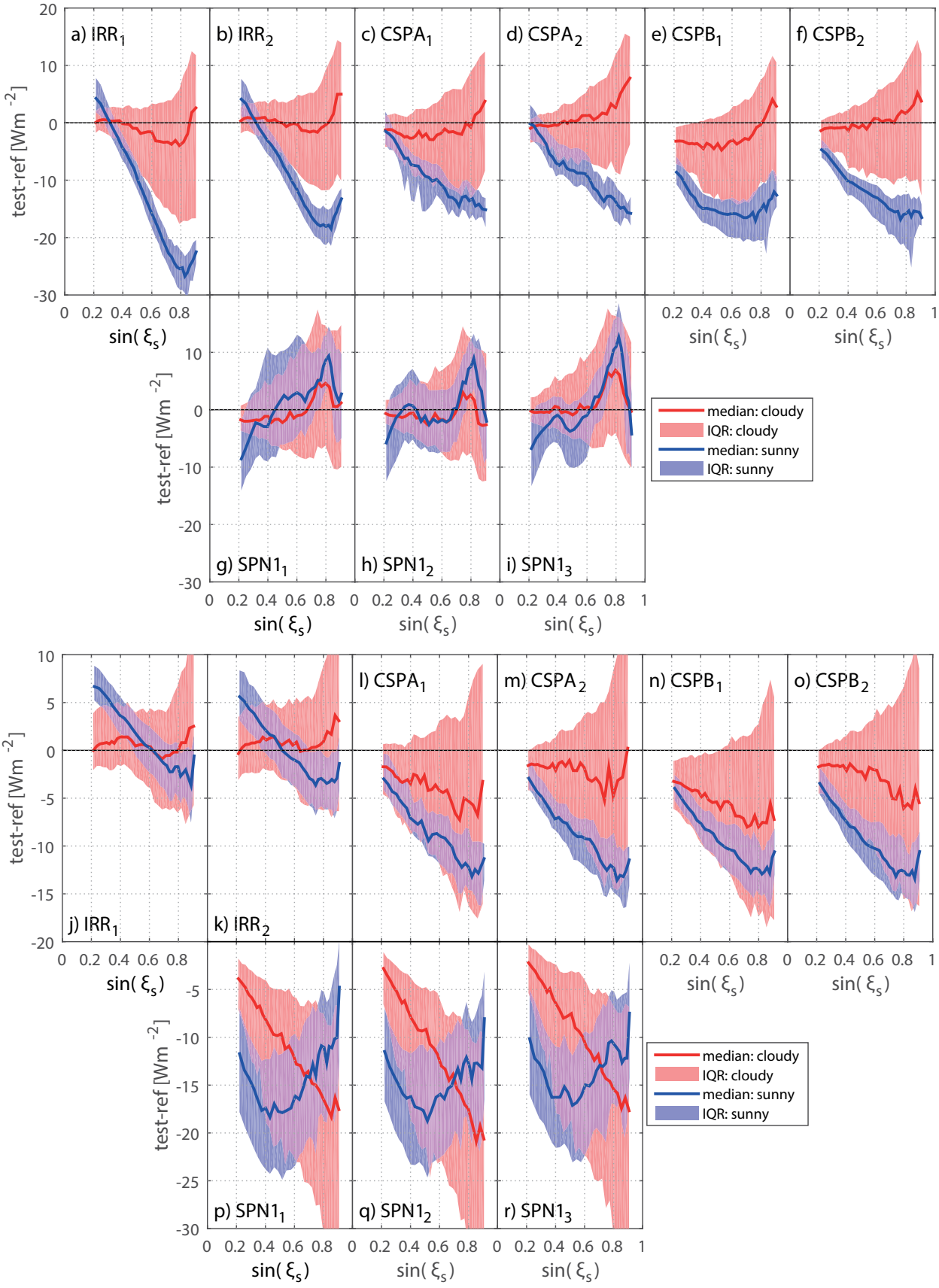

Figure 8: Error with respect to the sine of the solar elevation angle of the median and IQR of the error distributions between measurements made by the tested and reference instruments for GHI (panels a–i) and DfHI (panels j–r) for sunny (blue – low DNI variability) and cloudy conditions (red – high DNI variability).

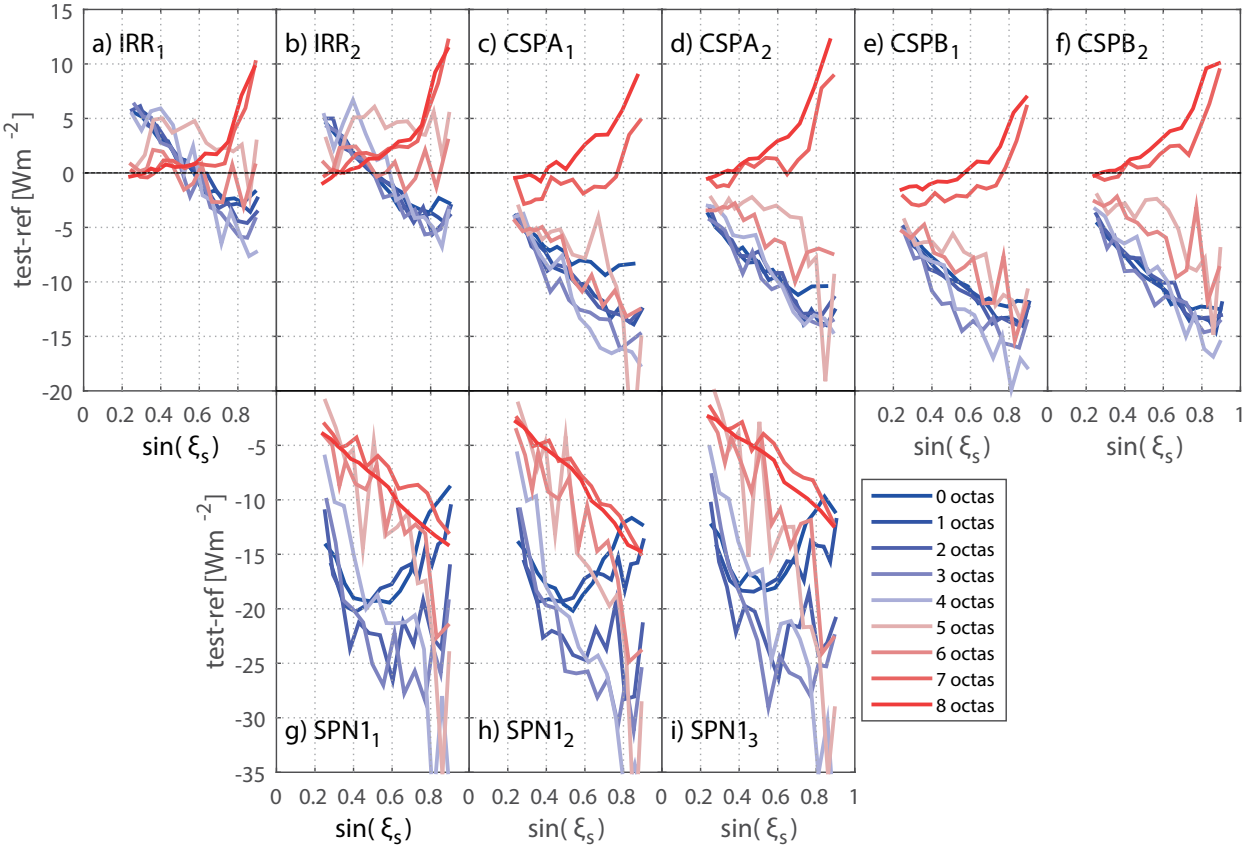

Figure 9: Error with respect to the sine of solar elevation angle and the cloud cover (in octas) of the median of the error distributions between measurements made by the tested and reference instruments, for DfHI.

seem to be optimized for reducing the error in the estimation of DNI. This could be the reason for the similarity between the GHI and DfHI error dependences.

The DNI error distributions for RSI (see Fig. 10, panels a–f) show a significant variability in their $\sin(\xi_s)$ dependence. Similarly, the analysis of DNI error distribution $\sin(\xi_s)$ dependences for different CC (not shown) does not allow detecting any consistent pattern for RSI. The medians are affected by a very significant variability that masks eventual differences between the curves for different CC, and the dependences differ from one instrument to the other. This absence of consistent error dependence pattern may be due to the fact that the RSI computation of DNI is already well optimized. The potential for further gains by a more complex $\sin(\xi_s)$ dependent correction is probably small.

Fig. 8 shows that SPN1 GHI error distributions exhibit limited dependences with respect to the solar elevation, while for DfHI, the error distributions exhibit dependence with $\sin(\xi_s)$ in both cloudy and sunny conditions. The DfHI error distributions are almost always negative – i.e., underestimation – including the upper limit of the IQR (zero is not within the IQR). Fig. 9 confirms that DfHI error medians for SPN1 are negative not only for all $\sin(\xi_s)$, but also for all CC. For low $\sin(\xi_s)$, the errors are small in irradiance, which is expected because the sun is low. In general, the errors are smaller for CC higher than 7 octas. This is also expected, because one source of DfHI underestimation for SPN1 is a geometrical effect related to a strongly non-isotropic radiance distribution in the vicinity of the sun and the fact that the fraction of the sky in the vicinity of the sun which is shaded for the DfHI measurement varies for different positions of the sun in the sky (Badosa et al., 2014). In overcast situations, the radiance distribution is relatively isotropic and such an effect is reduced. For low CC (0 or 1 octa), the underestimation is at its maximum for $\sin(\xi_s)$ between 0.4 and 0.6, while at higher $\sin(\xi_s)$, this underestimation is lower. In situations close to clear-sky, two effects can explain this behavior: first, for high elevation angles, the part of the sky around the sun that is hidden for the DfHI measurement by the SPN1 is more similar to this hidden by a regular shading device for a reference system, which reduces the error (Badosa et al., 2014); second, haze is frequent at Payerne when the weather is good, and it usually generates a stronger aureole around the sun at low than at high solar elevation. At high elevation, with a weaker aureole around the sun and a less anisotropic distribution of radiance around the sun, the error is reduced.

For GHI, the medians of the SPN1 error for various CC (not shown) are in a range between $-10\,\mathrm{Wm}^{-2}$ and $15\,\mathrm{Wm}^{-2}$ with the largest deviations at high $\sin(\xi_s)$. At high $\sin(\xi_s)$ the medians of GHI error are positive

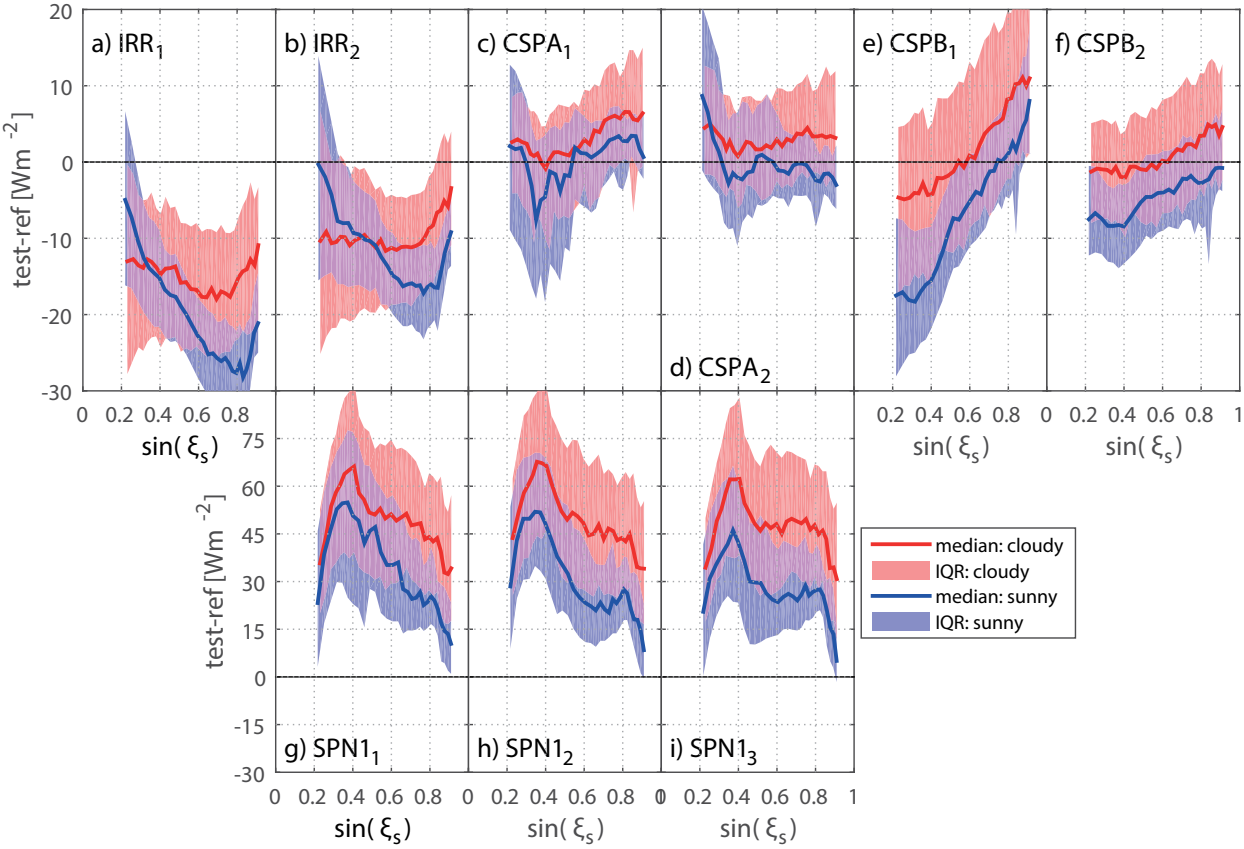

Figure 10: Error with respect the sine of the solar elevation angle of the median and IQR of the error distributions between measurements made by the tested and reference instruments for DNI for sunny (blue – low DNI variability) and cloudy conditions (red – high DNI variability).

for most CC, except for CC = 8 octas (median error of $-10\,\mathrm{Wm}^{-2}$) and CC = 7 octas (negligible median error). This detailed analysis confirms the findings of Badosa et al. (2014) pointing out geometrical effects as a major source of error in the determination of DfHI and GHI by SPN1, and related corrections would be the most effective in improving its performance.

The DNI $\sin(\xi_s)$ dependence for SPN1 among the three instruments is very similar and bears little resemblance to the dependence of the RSI instruments (Fig. 10), and the shape of the dependence is also similar for sunny and cloudy conditions: the errors are lower at low and high $\sin(\xi_s)$, and they exhibit a maximum for $\sin(\xi_s)$ between 0.3 and 0.4. This error maximum at intermediate $\sin(\xi_s)$ was also observed for DfHI in sunny conditions, and these are related, because of the way DNI is inferred from the difference between GHI and DfHI: Since the DfHI $\sin(\xi_s)$ dependence is stronger than the GHI $\sin(\xi_s)$ dependence, the DfHI influence is dominant in generating the DNI dependence.

As mentioned in Section 3.1, the events when DNI is undistinguishable from zero (nighttime or thick clouds in the sunbeam path) are not analyzed. Appendix A3 of the Supplementary Material investigates how well daylight events with no direct irradiance are detected by the tested instruments (thick clouds resulting in zero DNI events – ZDE) and its main results are given here. The probability of detection (POD) of ZDE is usually relatively high for RSI (between 79 % and 97 %, depending on the instruments). For SPN1 instruments, this POD is low (about 14 %). This is due to the overestimation of DNI by the SPN1 resulting in a positive bias, which makes it likely to measure a significantly positive DNI even in case of ZDE. A limit for the detection of ZDE was derived from the cumulative density function of the tested instrument DNI measurements taken during ZDE, using the 98th percentile. These limits for the detection of ZDE are 11 Wm^{-2} and 22 Wm^{-2} for the two IRR instruments, 8–11 Wm^{-2} for the CSPA instruments, and 4–6 Wm^{-2} for the CSPB instruments. For the SPN1 instruments, they are $\sim 40\,\mathrm{Wm}^{-2}$. This means that even if the tested instrument measures a non-zero DNI value (up to the limits described above), there is a non-negligible probability that the DNI is effectively zero.

4.5 Results of the PS-based analysis

The power spectral density (PS) analysis is used to explore timescales different than the 1-min data analyzed above. This can also be done by reapplying the analysis performed on 1-min data on data aggregated at longer timescales, typically 10-min or 1-hour averages. However, in this case a choice is made on the aggregation

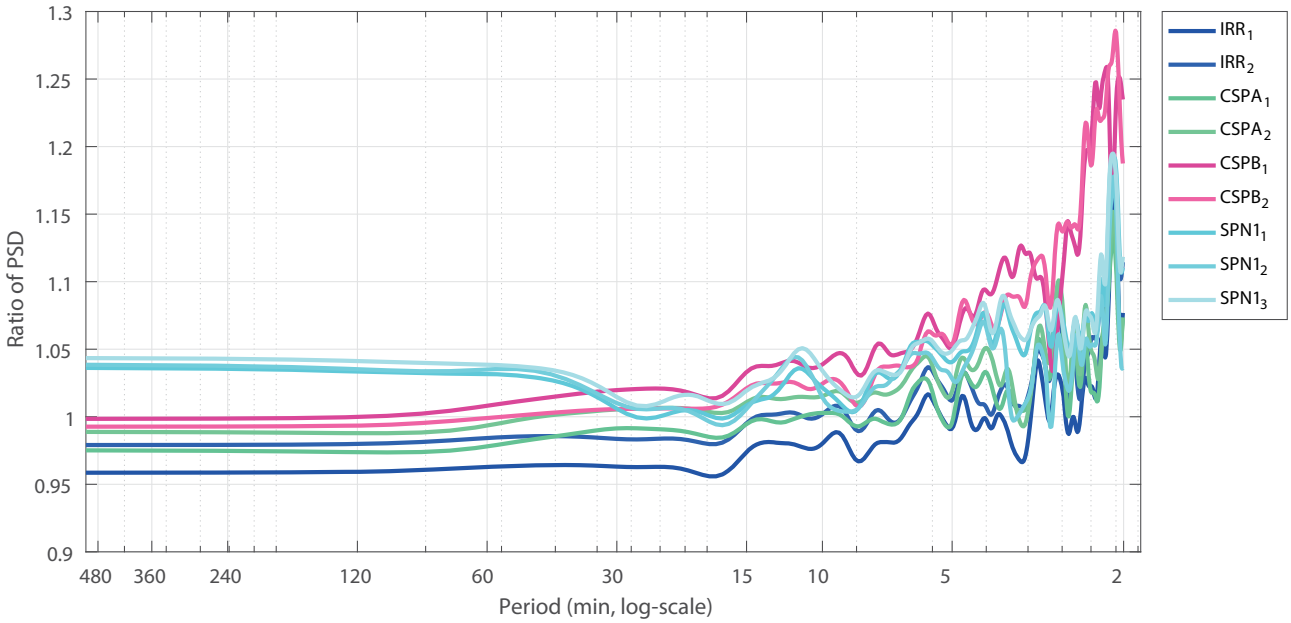

Figure 11: DNI ratio $R(f)$ of PSD for each tested instrument with respect to reference.

timescale. PS analysis allows studying given indicators (Section 3.3) as function of frequency – or equivalently timescale – and allows identifying timescales of interest where the PS analysis indicators exhibit changes or stop to exhibit changes. For instance, as shown below, the PS analysis indicators do not change anymore at timescale longer than 1–2 hours, and for these long timescales there is a consistent difference between RSI and SPN1 instruments.

But the PS analysis results are not directly comparable to those reported in Sections 4.2 and 4.3. To allow direct comparisons, results equivalent to those given in Appendices A2.2 and A2.3 are given in Appendix A4.2 for 1-hour averages computed from the 1-min measurements. The differences between results using 1-min data or 1-hour averages are as expected, with reduction in the error distribution spread but no significant change in the biases. In addition, an estimation of the error on the integrated yearly energy per unit surface area is given in Appendix A4.1. This quantity is of importance when assessing how robust solar power plant yearly energy output simulations are. Typically, errors on the order of a few percent or less (in all cases less than 5 %) are found for RSI with the best relative results obtained for DNI with negligible errors for some instruments. SPN1 errors are negligible for GHI, but on the order of −8 % for DfHI, and between 9 % and 11 % for DNI. These results confirm the pattern that the more similar the GHI and DfHI errors are, the smaller the DNI error.

The PS analysis presents a rather consistent picture. This analysis was limited to DNI, because it is the most rapidly varying parameter studied here, and consequently the most susceptible to presenting results exhibiting specificities related to given timescales. Figs. 11 to 13 present for each tested instrument and by comparison to the reference, the ratio $R(f)$, the frequency-based signal-to-noise ratio $SNR(f)$, and the frequency-based correlation coefficient $CC(f)$ as defined in Section 3.3. For the sake of readability, temporal frequencies are translated into their corresponding cycle periods, expressed in min. These figures show that the performance is similar across tested instruments when analyzing frequencies corresponding to timescales of about 1–2 hours or for longer ones. At such low frequencies, the SPN1s present a small excess of variance, while the RSIs have a variance equivalent to the reference instrument or slightly lower, depending on the instrument (Fig. 11). The *SNR* is on the order of 100 for SPN1s, while it is on the order of 650 to 950 for RSIs (Fig. 12). Measurements by RSIs and the reference present a high degree of coherence at timescales longer than 1 hour, while it is slightly lower for SPN1s (Fig. 13).

Such differences between RSIs and SPN1s at low frequency can be explained by the geometric uncertainty sources affecting the SPN1. Part of these errors is due to the fact that the sensor measuring the maximum signal changes with time (see Section 2.1). These changes occur at time scales of hours or fraction of hours, leading to a low frequency error. This is most likely the cause of the lower *SNR* for SPN1s, the excess of variance and the lack of coherence at such frequencies.

At high frequency (of the order of min^{-1}), all instruments show an excess of variance. Similarly, the *SNR* diminishes to values of the order of 2 to 7, and the correlation coefficient decreases to about 0.9. This may be partly due to the differences in time response of the tested sensor compared to the pyrheliometer. However, such an hypothesis requires further testing, especially because the time scales of the sensor responses – about 5 sec for the pyrheliometer and much shorter for the tested instruments – are shorter than the minimum aggregation time used in this study.

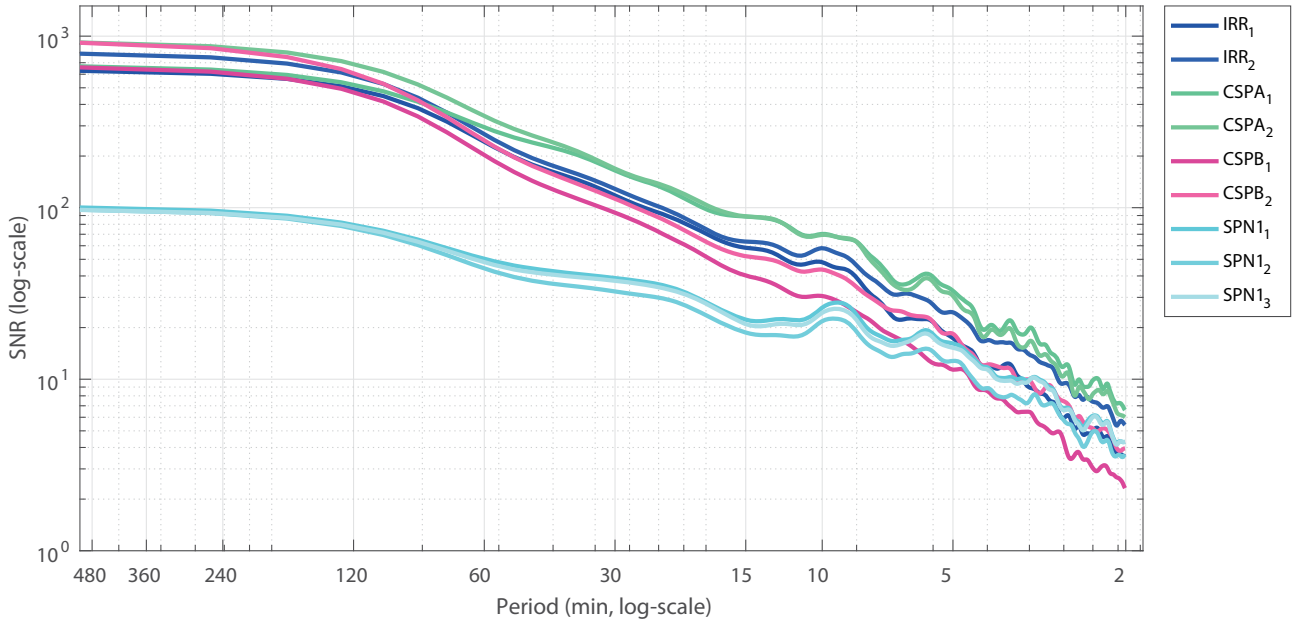

Figure 12: DNI frequency-based signal-to-noise ratio $SNR(f)$ for each instrument.

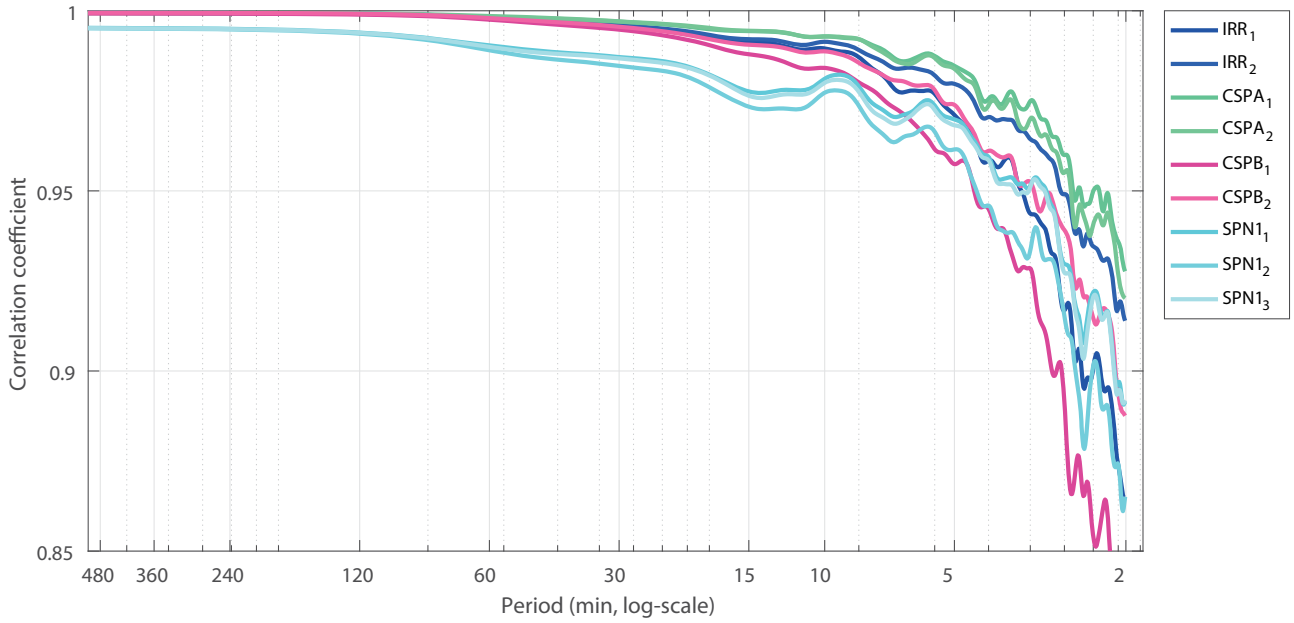

Figure 13: DNI frequency-based correlation coefficient $CC(f)$ for each instrument.

5 Conclusion

This study provides an evaluation of the performances of instruments such as RSI or SPN1 that are often used in networks or at remote isolated sites for assessing surface irradiance for the solar energy sector. From this detailed analysis the accuracy of measurements from such instruments may be assessed. The errors of the tested instruments are evaluated with respect to well characterized references and uses statistical indicators such as median, MAE, RMSE, inter-quartile range, and an extended range of the error distributions. The analysis was also applied on data subsets reflecting particular conditions (high or low solar elevation, high or low DNI variability) for better understanding the dependences of the instrument performance on such conditions. This gives some idea of how the uncertainty could change for different locations and atmospheric conditions. Caution must however be exercised in generalizing the results, because at some locations the conditions may be too far from those experienced at Payerne. For instance, the integrated water vapor column is almost always significant over Payerne, so it may not be possible to generalize the results of this study to locations with quite dry atmosphere such as arid regions or high elevation sites. The same comment stands for the solar elevation which is limited to 67° at Payerne.

It is difficult to directly compare our results to those obtained by Badosa et al. (2014) or Habte et al. (2016) because in those studies, the results of the tested instruments were normalized by factors derived from the comparison with the reference data. On the contrary, we did not allow any correction of the data from the tested instrument that would use knowledge of the reference instrument data. We still provide here some comparison with the result of these studies, but taking into account the applied correction factors. Badosa et al. report that when GHI and DHI are recalibrated (using the reference data), the mean absolute error is around 3 % or better for GHI and 2–3 times larger for DfHI and DNI. The recalibration method of Badosa et al. is based on dividing the data by the regression slope obtained by a linear fit forced through the origin of the tested data as function of the reference data. First, both the GHI and DfHI data are divided by the slope of the GHI regression and then a second normalization step is applied on the DfHI data alone using the slope of the DfHI tested data vs. reference DfHI data. The two-step process for DfHI data is assumed to first take into account general calibration error for the SPN1 sensor, and then specific issues linked to the DfHI measurement. According to Table 4 from Badosa et al. (2014), the first slope (derived from GHI, but applied to all data) range from 0.96 (Palaiseau) to 1.02 (Golden), while the second slope (DfHI) range from 0.90 (Addu Atoll) to 0.96 (Golden). The DfHI correction slopes show that the SPN1 DfHI data are underestimated by a factor between −10 % to −4 %, which correspond to the result of our study, while the mean absolute error found by Badosa et al. after recalibration are somewhat compatible with the spread found in our study.

Habte et al. (2016) report Mean Bias Difference (MBD) in percent for GHI and DNI including three RSI instruments provided by Irradiance Inc. and one provided by Solar Millennium AG that are similar to those tested in our study. One SPN1 radiometer is also included in their study. Similarly to Badosa et al., they recalibrated the tested instrument data by dividing these by a normalization factor. For each tested instrument and for both GHI and DNI, the normalization factor is the ratio of the sum of all tested instrument data collected between solar elevation angles 44° to 46° divided by the sum of all corresponding reference data. For RSI-measured GHI, mean MBD between −3 % and +3 % are reported depending on the conditions (all-sky, clear-sky, mostly cloudy) with spread of typically ±4 % around these values. The RSI GHI normalization factors are between 0.98 and 1.00 (see Table 1 of Habte et al., 2016). The mean MBD for SPN1 GHI is reported to be between 1.5 % and 5 % (depending on conditions) with a typical spread of ±5 % around these values, but a normalization factor of 1.03 is used. Without this normalization factor, the MBD would be even higher. For RSI-measured DNI, mean MBD between −2 % and 0 % are reported for all-sky conditions with spread of typically ±8 % around these values. For clear-sky conditions, similar mean MBD are found, but with smaller spread (±3 %), while for mostly cloudy conditions, more negative mean MBD are reported by Habte et al. (−8 %) with larger spread (±16 %). The RSI DNI normalization factors are between 0.98 and 1.00 (see Table 2 of Habte et al., 2016). For the SPN1-measured DNI, the mean MBD is between +3 % (±6 %) for clear-sky conditions and +21 % (±16 %) for mostly cloudy conditions. Again, a normalization factor of 1.05 is used, and without it the reported MBD would be even higher. Such large biases for SPN1-measured DNI are even higher than those that we found (see below).

In our study, the overall performance of the tested instruments for measuring GHI is not as good as that of the reference instruments described by Vuilleumier et al. (2014). But it is nonetheless satisfactory as none of the instruments exhibit error distributions significantly exceeding about ±30 Wm^{-2}, which would be the agreement expected for instruments with an expanded uncertainty of ±25 Wm^{-2} (±10 %). For DfHI, RSIs exhibited errors on the order of ±20 Wm^{-2} (±13 %) with some of them being affected by small systematic negative biases on the order of −5 Wm^{-2} (median). SPN1s instruments underestimate DfHI by about −10 Wm^{-2} (median of the error distribution) with a relatively large range of the expanded error distribution between −45 Wm^{-2} and 20 Wm^{-2} (−35 % to 13 %). For DNI, the extended range of the error distribution for RSI is on the order of ±40 Wm^{-2} (±5–6 %) with some instruments presenting no bias while others are affected by median biases up to −15 Wm^{-2}. SPN1 instruments exhibit a relatively large median bias of 40 Wm^{-2}, and an extended range of the error distribution between −45 Wm^{-2} and 125 Wm^{-2} (−6 % to 19 %). As mentioned, only events with DNI > 5 Wm^{-2} were analyzed. Typical error on the integrated yearly energy per unit surface area are of the order of a few percent or less (< 5 %) for RSI with negligible errors on DNI for some instruments. SPN1 integrated errors are negligible for GHI, but about −8 % for DfHI, and 9 % to 11 % for DNI.

The GHI and DfHI errors of RSIs were found to be very similar, especially for favorable conditions. They also showed similar dependences on solar elevation. DNI errors significantly smaller in relative value than GHI or DfHI errors were found, suggesting a cancelation of error when inferring DNI from GHI and DfHI. Such an error cancelation process requires the GHI and DfHI errors of RSIs to be correlated, which is corroborated by their similar solar elevation dependence.

Despite these similarities, the cause of RSI GHI and DfHI errors seem to be different. The GHI error solar elevation dependence seems to be linked to the increase of the GHI signal as solar elevation increases, while other –or additional – causes are necessary to explain the DfHI solar elevation dependence. Possible causes include spectral effects related to the non-uniform spectral response of the RSI sensor, non-ideal directional response and eventually effects related to the shading by a

rotating shadowband instead of the shading ball or disk used in sun tracker-based systems. These causes could have been investigated more thoroughly using for instance spectral diffuse data, data on the circumsolar radiance, etc. Unfortunately, even though the dataset used in this analysis is quite comprehensive, such data were lacking. Spectral data were only available for DNI. Vignola et al. (2016) specifically analyzed the influence of the non-uniform spectral response of the LICOR sensor on the detection of shortwave radiation for DNI using these data. Ideally, if spectral data had been available for DfHI and DNI, such an analysis could have been extended to DfHI.

The RSIs appear to be optimized for providing good estimates of DNI and the related accuracy is very reasonable; the uncertainty is about twice the uncertainty of a "good quality pyrheliometer" (World Meteorological Organization, 2010) in favorable conditions, even though DNI is not measured directly but inferred. Reducing this uncertainty further would be challenging. It would require a detailed understanding of the DfHI and GHI errors to reduce them in a first step, followed by an optimization of the corrections applied when inferring the DNI itself. For GHI, the performance of RSI instruments would place them at the limit between "moderate" and "good quality pyranometers" (World Meteorological Organization, 2010).

The SPN1 instruments, in contrast to RSI instruments, exhibit errors that tend to be of opposite signs for GHI and DfHI, especially in favorable cases: slightly positive GHI median errors and significantly negative DfHI median errors. When inferring DNI, this results in an addition of errors that is further amplified by the presence of $\sin(\xi_s)$ in the denominator. In difficult cases, the DfHI errors are smaller, but because the solar elevation is low the error amplification due to $\sin(\xi_s)$ is stronger. The resulting errors when simply inferring DNI from the difference between GHI and DfHI are large. While SPN1 performance for measuring GHI is similar to that of RSI, corrections are required to obtain satisfactory performance for DNI.

Corrections based on the detailed understanding of the sources of the error would be the best option, but such corrections have, as yet, proven to be elusive. Badosa et al. (2014) proposed some empirical corrections. The simplest one, for mid-latitude sites, would be to correct SPN1 DfHI and DNI with a multiplying factor of about 1.05 and 0.95, respectively. But given the homogeneity of the dependence of the DNI error distribution on solar elevation, it might be more appropriate to explore a $\sin(\xi_s)$ dependent correction, possibly also depending on the cloud cover. In any case, to the extent that the largest source of error for these instruments is the large and variable field of view, there is potential for improving DfHI accuracy, and thus for DNI if derived from the difference between GHI and DfHI.

This analysis was performed using 1-minute data. Aggregating the data on longer time steps usually improves measurement performance by reducing the error through averaging. The power spectral density analysis applied to DNI data show that such improvements are most significant when aggregating on timescales up to about 1–2 hours. Aggregating beyond these timescales provides little further gain in performance.

This study has analyzed the performance of GHI, DfHI and DNI measured by each type of instrument, scrupulously maintained during the whole study, under otherwise standard operating conditions. For instruments that would be less well maintained (remote operation with little on-site maintenance), the main problem is anticipated to be soiling. Michalsky et al. (1988) and Geuder and Quaschning (2006) investigated the effect of soiling on RSI instruments, and found that such instruments are less affected by soiling than other types of radiometers. There is no study known to the authors investigating the effect of soiling on SPN1 instruments. This study has further explored how the instrument uncertainty depends on the operating conditions. To assess the transferability of results it would be useful to repeat the measurements and analyses under different climatic conditions, e.g., dryer or more humid environments (arid regions or high elevation sites vs. tropics) to assess the transferability of the results obtained here. It would be especially useful in such studies to also deploy instruments allowing spectral measurements of both DNI and DfHI on a wide wavelength spectrum. Similarly, an instrument able to measure the energy available in the sun aureole might better illustrate and explain geometrical or directional error sources.

6 Acknowledgements

The Payerne instrument performance evaluation was conducted within the framework of the COST Action ES1002 Weather Intelligence for Renewable Energies (WIRE), and is a contribution to the project Direct Normal Irradiance Nowcasting methods for optimized operation of concentrating solar technologies (DNICast). DNICast received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 608623. N. Sommer, G. Paglia, M. Hauser and H. Monnard contributed to this work in the framework of internships at MeteoSwiss. M. Schwandt from Suntrace GmbH, Germany, P. Marcacci from RSE – Ricerca Sistema Energetico S.p.A., Italy and M.-C. Kotti from the Laboratory of Atmospheric Physics, University of Patras, Greece participated in the inter-comparison within Short-Term Scientific Mission with the support of the COST Action ES1002 Weather Intelligence for Renewable Energies (WIRE). The companies CSP Services GmbH, Delta-T Devices Ltd and Irradiance, Inc. are thanked for providing the instruments tested during this inter-comparison, support in the installation and operation of the instruments, and logistic support. Finally, we are grateful to two anonymous reviewers who carefully checked our work and made suggestions leading to substantial improvements.

References

Augustyn, J., T. Geer, T. Stoffel, E. Kern, R. Little, F. Vignola, R. Kessler, 2002: Improving the accuracy of low cost measurement of direct normal solar irradiance. – Proc. 2002 American Solar Energy Society Conference, 329–334, American Solar Energy Society, Boulder, Colorado, USA.

Augustyn, J., T. Geer, T. Stoffel, F. Vignola, R. Kessler, E. Kern, R. Little, B. Boyson, 2004: Update of algorithm to correct direct normal irradiance measurements made with a rotating shadow band pyranometer. – Proc. 2004 American Solar Energy Society Conference, 295–302. American Solar Energy Society, Boulder, Colorado, USA.

Badosa, J., J. Wood, P. Blanc, C.N. Long, L. Vuilleumier, D. Demengel, M. Haeffelin, 2014: Solar irradiances measured using SPN1 radiometers: uncertainties and clues for development. – Atmos. Mea.Techn. **7**, 4267–4283, DOI: 10.5194/amt-7-4267-2014.

Chaabene, M., M. Ben Ammar, 2008: Neuro-fuzzy dynamic model with Kalman filter to forecast irradiance and temperature for solar energy systems. – Renew. Energy **33**, 1435–1443, DOI:10.1016/j.renene.2007.10.004.

Dürr, B., R. Philipona, 2004: Automatic cloud amount detection by surface longwave downward radiation measurements. – J. Geophys. Res. **109**, D05201, DOI:10.1029/2003JD004182.

Dutton, E.G., C.N. Long, with contributions by M. Wild, A. Ohmura, J. Groebner, A. Roesch, 2012: Chapter 5: Long-term in-situ surface flux data products. – GEWEX Radiative Flux Assessment Vol. **1**, 135–158, WCRP Report 19/2012, World Climate Research Programme, Geneva, Switzerland. http://www.wcrp-climate.org/documents/GEWEX%20RFA-Volume%201-report.pdf (accessed at 14.10.2016).

Dutton, E.G., J.J. Michalsky, T. Stoffel, B.W. Forgan, J. Hickey, D.W. Nelson, T.L. Alberta, I. Reda, 2001: Measurement of broadband diffuse solar irradiance using current commercial instrumentation with a correction for thermal offset errors. – J. Atmos. Oceanic Technol. **18**, 297–314, DOI: 10.1175/1520-0426(2001)018<0297:MOBDSI>2.0.CO;2.

Geuder, N., V. Quaschning, 2006: Soiling of irradiation sensors and methods for soiling correction. – Solar Energy **80**, 1402–1409, DOI:10.1016/j.solener.2006.06.001.

Geuder, N., B. Pulvermüller, O. Vorbrugg, 2008: Corrections for rotating shadowband pyranometers for solar resource assessment. – Proc. SPIE Optics + Photonics **7046**, 70460F–70460F-12, San Diego, USA, DOI:10.1117/12.797472.

Geuder, N., R. Affolter, O. Goebel, B. Dahleh, M. Al Khawaja, S. Wilbert, B. Pape, B. Pulvermueller, 2010: Validation of direct beam irradiance measurements from rotating shadowband pyranometers in a different climate. – 16th SolarPACES Conference, 21.–24. September 2010, Perpignan, France.

Gueymard, C.A., 2014: A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. – Renew. Sust. Energ. Rev. **39**, 1024–1034, DOI: 10.1016/j.rser.2014.07.117.

Habte, A., M. Sengupta, A. Andreas, S. Wilcox, T. Stoffel, 2016: Intercomparison of 51 radiometers for determining global horizontal irradiance and direct normal irradiance measurements. – Solar Energy **133**, 372–393, DOI:10.1016/j.solener.2016.03.065.

Harrison, L., J.J. Michalsky, J. Berndt, 1994: Automated multifilter rotating shadow-band radiometer: an instrument for optical depth and radiation measurements. – Appl. Optics **33**, 5118–5125, DOI:10.1364/AO.33.005118.

Ji, Q., S.-C. Tsay: 2000. On the dome effect of Eppley pyrgeometers and pyranometers. – Geophys. Res. Lett. **27**, 971–974, DOI:10.1029/1999GL011093.

McArthur, L.J.B., 2005: Baseline Surface Radiation Network (BSRN) Operations Manual Version 2.1. – WMO/TD-No. 1274, World Meteorol. Organ., Geneva, Switzerland. http://bsrn.awi.de/fileadmin/user_upload/bsrn.awi.de/Publications/McArthur.pdf (accessed at 14.10.2016).

Michalsky, J.J., J.L. Berndt, G.J. Schuster, 1986: A microprocessor-based rotating shadowband radiometer. – Solar Energy **36**, 465–470, DOI:10.1016/0038-092X(86)90095-2.

Michalsky, J.J., L. Harrison, B.A. LeBaron, 1987: Empirical radiometric correction of a silicon photodiode rotating shadowband pyranometer. – Solar Energy**39**, 87–96, DOI:10.1016/S0038-092X(87)80036-1.

Michalsky, J.J., R. Perez, R. Stewart, B.A. LeBaron, L. Harrison, 1988: Design and development of a rotating shadowband radiometer solar radiation/daylight network. – Solar Energy **41**, 577–581, DOI:10.1016/0038-092X(88)90060-6.

Michalsky, J.J., R. Perez, L. Harrison, B.A. LeBaron, 1991: Spectral and temperature correction of silicon photovoltaic solar radiation detectors. – Solar Energy **47**, 299–305, DOI: 10.1016/0038-092X(91)90121-C.

Michalsky, J.J., L.C. Harrison, W.E. Berkheiser III, 1995: Cosine response characteristics of some radiometric and photometric sensors. – Solar Energy **54**, 397–402, DOI:10.1016/0038-092X(95)00017-L.

Michalsky, J.J., R. Dolce, E.G. Dutton, M. Haeffelin, G. Major, J.A. Schlemmer, D.W. Slater, J.R. Hickey, W.Q. Jeffries, A. Los, D. Mathias, L.J.B. McArthur, R. Philipona, I. Reda, T. Stoffel, 2003: Results from the first ARM diffuse horizontal shortwave irradiance comparison. – J. Geophys. Res. **108**, 4108, DOI:10.1029/2002JD002825.

Michalsky, J., E.G. Dutton, D. Nelson, J. Wendell, S. Wilcox, A. Andreas, P. Gotseff, D. Myers, I. Reda, T. Stoffel, K. Behrens, T. Carlund, W. Finsterle, D. Halliwell, 2011: An Extensive Comparison of Commercial Pyrheliometers under a Wide Range of Routine Observing Conditions. – J. Atmos. Ocean. Technol. **28**, 752–766, DOI: 10.1175/2010JTECHA1518.1.

Myers, D.R., 2005: Solar radiation modeling and measurement for renewable energy applications: data and model quality. – Energy **30**, 1517–1531, DOI:10.1016/j.energy.2004.04.034.

Myers, D.R., K. Emery, T. Stoffel, 1989: Uncertainty estimates for global solar irradiance measurements used to evaluate PV device performance. – Solar Cells **27**, 455–464, DOI: 10.1016/0379-6787(89)90055-0.

Reda, I., 2011: Method to Calculate Uncertainties in Measuring Shortwave Solar Irradiance Using Thermopile and Semiconductor Solar Radiometers. – NREL techn. report NREL/TP-3B10-52194, National Renewable Energy Laboratory, Golden, Colorado, USA, DOI:10.2172/1021250.

Rigollier, C., O. Bauer, L. Wald, 2000: On the clear sky model of the ESRA – European Solar Radiation Atlas – with respect to the heliosat method. – Solar Energy **68**, 33–48, DOI: 10.1016/S0038-092X(99)00055-9.

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

Vernay, C., P. Blanc, S. Pitaval, 2013: Characterizing measurements campaigns for an innovative calibration approach of the global horizontal irradiation estimated by

HelioClim-3. – Renew. Energy **57**, 339–347, DOI:10.1016/j.renene.2013.01.049.

Vignola, F., 1999: Solar Cell Based Pyranometers: Evaluation of the Diffuse Response. – Proc. 1999 American Solar Energy Society Conference, 255–260, American Solar Energy Society, Boulder, Colorado, USA.

Vignola, F., 2006: Removing Systematic Errors from Rotating Shadowband Pyranometer Data. – Proc. 2006 American Solar Energy Society Conference, American Solar Energy Society, Boulder, Colorado, USA.

Vignola, F., Z. Derocher, J. Peterson, L. Vuilleumier, C. Félix, J. Gröbner, N. Kouremeti, 2016: Effects of changing spectral radiation distribution on the performance of photodiode pyranometers. – Solar Energy **129**, 224–235, DOI:10.1016/j.solener.2016.01.047.

Vuilleumier, L., M. Hauser, C. Félix, F. Vignola, P. Blanc, A. Kazantzidis, B. Calpini, 2014: Accuracy of ground surface broadband shortwave radiation monitoring. – J. Geophys. Res. **119**, 13'838–13'860, DOI:10.1002/2014JD022335.

Welch, P.D., 1967: The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. – IEEE Transactions on Audio and Electroacoustics **15**, 70–73.

Wilcox, S.M., D.R. Myers, 2008: Evaluation of Radiometers in Full-Time Use at the National Renewable Energy Laboratory Solar Radiation Research Laboratory. – NREL techn. report NREL/TP-550-44627, National Renewable Energy Laboratory, Golden, Colorado, USA, DOI:10.2172/946331.

Wood, J., 2012: SPN1 Sunshine Pyranometer User Manual v3.0. – Delta-T Devices Ltd, Cambridge, UK. http://www.delta-t.co.uk/product-support-material.asp (accessed at 18.10.2016).

World Meteorological Organization, 2010: Guide to Meteorological Instruments and Methods of Observation. WMO-No. **8**, World Meteorol. Organ., Geneva, Switzerland. https://www.wmo.int/pages/prog/gcos/documents/gruanmanuals/CIMO/CIMO_Guide-7th_Edition-2008.pdf (accessed at 18.10.2016).

Yang, D., C. Gu, Z. Dong, P. Jirutitijaroen, N. Chen, W.M. Walsh, 2013: Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging. – Renew. Energy **60**, 235–245, DOI:10.1016/j.renene.2013.05.030.

The pdf version (Adobe Java Script must be enabled) of this paper includes an electronic supplement:
Table of content – Electronic Supplementary Material (ESM)

Appendix A1, A2, A3, A4