Solar Energy 155 (2017) 1451–1461

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener

Improvements in world-wide intercomparison of PV module calibration

E. Salis [a,*], D. Pavanello [a], M. Field [a], U. Kräling [b], F. Neuberger [b], K. Kiefer [b], C. Osterwald [c], S. Rummel [c], D. Levi [c], Y. Hishikawa [d], K. Yamagoe [d], H. Ohshima [d], M. Yoshita [d], H. Müllejans [a]

[a] European Commission, Joint Research Centre, European Solar Test Installation, v. E. Fermi 2749, 21027 Ispra (VA), Italy
[b] Fraunhofer Institute for Solar Energy Systems ISE, Heidenhofstr. 2, 79110 Freiburg, Germany
[c] National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, USA
[d] National Institute of Advanced Industrial Science and Technology (AIST), Research Center for Photovoltaics, Central 2, 1-1-1, Umezono, Tsukuba 305-8568, Japan

ARTICLE INFO

Article history:
Received 10 April 2017
Received in revised form 30 June 2017
Accepted 28 July 2017

Keywords:
Photovoltaic
Intercomparison
Calibration
Measurement uncertainty

ABSTRACT

The calibration of the electrical performance of seven photovoltaic (PV) modules was compared between four reference laboratories on three continents. The devices included two samples in standard and two in high-efficiency crystalline silicon technology, two CI(G)S and one CdTe module. The reference value for each PV module parameter was calculated from the average of the results of all four laboratories, weighted by the respective measurement uncertainties. All single results were then analysed with respect to this reference value using the E_n number approach. For the four modules in crystalline silicon technology, the results agreed in general within ±0.5%, with all values within ±1% and all E_n numbers well within [−1,1], indicating further scope for reducing quoted measurement uncertainty. Regarding the three thin-film modules, deviations were on average roughly twice as large, i.e. in general from ±1% to ±2%. A number of inconsistent results were observable, although within the 5% that can be statistically expected on the basis of the E_n number approach. Most inconsistencies can be traced to the preconditioning procedure of one participant, although contribution of other factors cannot be ruled out. After removing these obvious inconsistent results, only two real outliers remained, representing less than 2% of the total number of measurands. The results presented show improved agreement for the calibration of PV modules with respect to previous international exercises. For thin-film PV modules, the preconditioning of the devices prior to calibration measurements is the most critical factor for obtaining consistent results, while the measurement processes seem consistent and repeatable.
© 2017 The Authors and National Renewable Energy Laboratory. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

As the amount of photovoltaic (PV) installations increases around the world, the reliability of such technology for energy supply becomes more and more crucial for both the energy demand and the financial investment that supports it. In order to maintain at the highest level possible the quality of PV production or even to improve it, reliable and accurate measurements are fundamental to set the reference. To achieve this objective, continuation of periodic intercomparisons on PV module calibration among international reference laboratories is of primary importance so that:

(1) The highest possible level of measurement confidence is assured;
(2) A calibration measurement "good practice" for traditional PV technologies is established, eventually including additional measurement conditions and parameters beyond the standard conditions;
(3) The calibration practices of thin-film PV technologies can be compared and improved;
(4) World-wide consensus on PV standards drafting and update can be established and enlarged, in order to respond to the latest improvements in the PV market and management thereof.

In this framework, a second intercomparison between four reference laboratories for PV calibration was carried out between February 2015 and February 2017, following the good practice of interlaboratory comparison already established with the previous calibration campaign (Dirnberger et al., 2014).

The four laboratories involved are the European Solar Test Installation (ESTI) of the European Commission's Joint Research Centre (Ispra, Italy), the CalLab PV Modules (CalLab) of the

* Corresponding author.
E-mail address: elena.salis@ec.europa.eu (E. Salis).

http://dx.doi.org/10.1016/j.solener.2017.07.081
0038-092X/© 2017 The Authors and National Renewable Energy Laboratory. Published by Elsevier Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Fraunhofer Institute for Solar Energy Systems (Freiburg, Germany), the Research Center for Photovoltaics of the National Institute of Advanced Industrial Science and Technology (AIST; Tsukuba, Japan) and the National Renewable Energy Laboratory (NREL; Golden, CO, USA). The Fraunhofer ISE laboratory coordinated the intercomparison and also provided the PV modules to be measured by purchasing them on the market.

The intercomparison aims to keep the comparability of calibration results obtained by these four reference laboratories up to date. This is achieved by including their updated measurement uncertainties as well as PV technologies that present different calibration challenges.

2. Methodology

2.1. Type of material

Seven PV modules of various technologies were included in the intercomparison: two of them were traditional c-Si modules, two high-efficiency c-Si modules, two CI(G)S modules and one CdTe module. The list is summarised in Table 1 with the main characteristics for each. Modules M1 and M2 had already been included in the previous intercomparison (respective codes M1 and M2 in (Dirnberger et al., 2014)) and were used here as monitoring modules between the four laboratories and between the periodical intercomparisons as well. The relative spectral responsivity for these modules is shown in Fig. 1.

2.2. Procedures

All seven PV modules were measured in turn by each partner and then sent to the next in sequence. CalLab initiated the intercomparison, followed in order by NREL, ESTI and AIST. The parameters to be compared were the short-circuit current I_{SC}, the open-circuit voltage V_{OC}, the power at maximum power point P_{MAX} and the fill factor *FF*. At the end of the intercomparison, the modules were returned to the Fraunhofer ISE, where they were re-measured. No intermediate measurements were performed at CalLab on thin-film modules, contrary to the previous intercomparison (Dirnberger et al., 2014).

The silicon devices (M3, M4) not included in the previous intercomparison were stabilised at CalLab (see CalLab's procedure hereafter). Modules M1 and M2 had already been included in the previous intercomparison, thus they did not need any additional stabilisation.

For the thin-film modules (M5, M6 and M7), a first measurement at the partners' premises was required as soon as the modules were taken out of the shipping crate and another after the preconditioning procedure was completed. The results from the latter were used for the final intercomparison between the four laboratories. As one of the main goals of the intercomparison was to compare calibration results for PV modules as they would be according to each reference laboratory's internal procedures, no common protocol was adopted for thin-film modules, and all partners applied their own procedures for both preconditioning and calibration, which are described hereafter for each laboratory.

2.2.1. CalLab

Measurements at CalLab PV Modules were performed with a PasanSunSim3b pulsed solar simulator with a flash duration of 10 ms. A detailed description of the measurement procedure and an uncertainty estimation is given in (Dirnberger and Kräling, 2013). The best-case uncertainty for P_{MAX} of crystalline silicon modules was 1.6% during intercomparison. Primary calibrated cells (PTB) with an uncertainty of 0.6% of the calibration value were used for all measurements. The spectral mismatch correction for c-Si modules was performed on the basis of the spectral response as measured at CalLab PV Cells, Fraunhofer ISE's accredited calibration laboratory for PV cells. The spectral response is measured on one single cell of a separate module of identical module type. The spectral response for single-junction thin-film modules was measured at CalLab with a Pasan SunSim3b system and bandpass filters on module level. All measurements were carried out as hysteresis and section measurements in order to monitor and control sweep speed effects. Hysteresis measurements are measurements from I_{SC} to V_{OC} (forward) and V_{OC} to I_{SC} (backward). This allows for detection of capacity-related over- or underestimations in the module's power. Section measurements are applied to reduce hysteresis effects. The I-V curve is not measured entirely during one flash, but swept in smaller voltage sections during several flashes, which decreases the voltage sweep rate. Temperature in the laboratory is strictly controlled so that the module has a temperature between 24 °C and 26 °C at the time of the measurement. The temperature is measured with four Pt100 probes at the back of the module. Electroluminescence images of the c-Si modules were taken at the end of the comparison. The modules M3 and M4 were stabilised using a steady-state solar simulator for 20 h at irradiance level of 1000 W/m^2 at V_{OC}. Both CI(G)S modules were exposed for 10 h at irradiance level of 1000 W/m^2 at V_{OC} using the steady-state solar simulator prior to the measurement. The CdTe module was preconditioned according to procedure developed and provided by manufacturer of the CdTe modules.

2.2.2. NREL

Measurements at NREL were performed using three different module test beds. These test beds include a Spectrolab model X200 class AAA continuous solar simulator, a Spire model 5600 class AAA pulsed solar simulator, and an outdoor module test bed under natural sunlight.

The four silicon modules in this intercomparison were measured using a new method that minimizes uncertainty by coupling measurements on all three test beds (Levi et al., 2017). First, the module is brought to thermal equilibrium at 25 °C indoors. V_{OC} versus irradiance is measured at 800, 900, and 1000 W/m^2 using the Spire 5600. The light level in the simulator is set using a c—Si reference cell with I_{SC} uncertainty of 1.4%. The flash duration is short

Table 1
Module list. The value reported for each electrical parameter is the nominal value given by the manufacturer, except for the modules M1 and M2 for which the results of the previous intercomparison are listed (Dirnberger et al., 2014).

Device	Technology	Cells No./type	Size [m × m]	I_{SC} [A]	V_{OC} [V]	P_{MAX} [W]	*FF* [%]
M1	High efficiency c-Si	240/Honeycomb	1.610 × 0.860	7.355	43.30	239.8	75.30
M2	Poly c-Si	72/6 in.	1.970 × 0.990	8.286	44.29	272.9	74.34
M3	High efficiency c-Si	96/5.1 in.	1.559 × 1.045	6.2	64.7	318	79.3
M4	Poly c-Si	60/6 in.	1.650 × 0.990	8.37	37.2	240	77.6
M5	CI(G)S	170/stripes	1.256 × 0.976	2.2	112	170	67.0
M6	CI(G)S	104/stripes	1.586 × 0.670	3.11	58.5	120	66.0
M7	CdTe	154/stripes	1.200 × 0.600	1.99	60.5	87.5	72.7

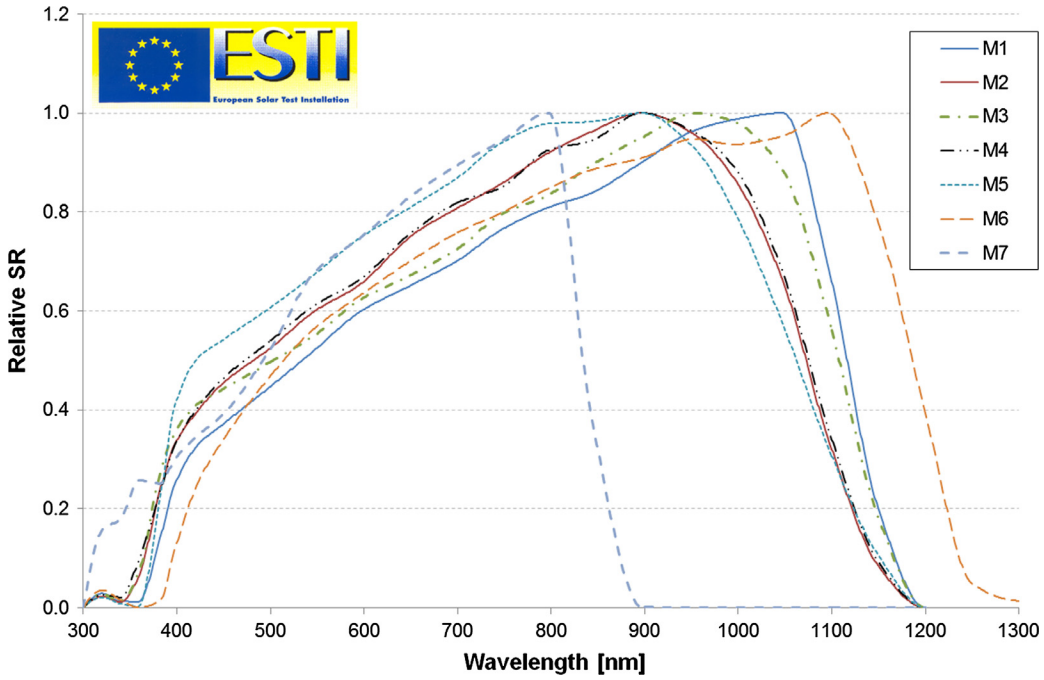

Fig. 1. Relative spectral responsivity of the seven modules included in the intercomparison. The curves shown have been obtained by interpolation of the data points as measured at ESTI, then normalised to the maximum.

enough (100 ms), that module heating is negligible. This procedure provides a calibration of V_{OC} versus irradiance within 0.3 °C of 25 °C.

This V_{OC} calibration is subsequently used to measure junction temperature in the module during measurements in the outdoor test bed and continuous simulator, reducing errors due to temperature uncertainty. Prior to outdoor measurement the module is cooled to ~15 °C, then mounted on the outdoor test bed and allowed to heat under natural sunlight while V_{OC} is monitored. When V_{OC} indicates the module is at ~24.5 °C the I-V curve is measured. Irradiance is measured using a c-Si primary reference cell calibrated at NREL with an uncertainty of 0.42%. The spectrum is measured using an ASD FieldSpec3 spectroradiometer. Because natural sunlight is expected to have uniform irradiance, this technique minimises uncertainty in I_{SC} related to spatial non-uniformity of the irradiance. The value of I_{SC} from the outdoor test bed is adjusted for spectral mismatch and scaled to irradiance of 1000 W/m^2 (Campanelli and Osterwald, 2016).

After the outdoor measurement, the module is again cooled to ~15 °C, then mounted on the Spectrolab X200 continuous simulator. Instead of using a reference cell to set the light intensity, the value of I_{SC} from the outdoor test bed is used to set the irradiance level of the simulator. This procedure minimises uncertainty in P_{MAX} by eliminating spectral error between the module and reference cell, and also minimizes error in I_{SC} due to spatial non-uniformity of the light source. After mounting, the module is illuminated and V_{OC} is monitored until it indicates the module is at ~24.5 °C. The I-V sweep is then performed, providing a measurement at 25 °C with minimal error due to uncertainties in temperature, spectrum, and non-uniformity of irradiance.

Thin-film modules were measured using only the continuous solar simulator against a primary reference cell. The modules were first exposed to natural sunlight outdoors for at least 0.2 kWh/m^2, corresponding to ~20 min exposure time. Exposure was performed under open-circuit condition without monitoring the device temperature. This procedure is designed for short-term metastable states and not for longer-term transients. After light-soaking the devices are cooled down to ~23 °C, then measured using the Spectrolab model X200 class AAA continuous simulator. The measurement was performed about 1.5 h after the completion of the preconditioning.

2.2.3. ESTI

The I-V curves of the seven modules were measured under natural sunlight with the same procedures as explained in (Dirnberger et al., 2014). The SR of each full module was measured with pulsed monochromatic light (Bahrs et al., 1997; Dirnberger et al., 2014).

Concerning the thin-film modules, a preconditioning procedure was followed in order to take the device to a stable condition before calibration. The CI(G)S modules were exposed for approximately 15–20 min to natural sunlight while keeping them under open-circuit conditions. The stabilisation was checked with a series of I-V curve measurements acquired under natural sunlight immediately before exposure to sunlight and after it, and on different days. For the CdTe the preconditioning procedure according to standard IEC 61646 Ed. 2.0 (IEC, 2008) was applied until the stability requirement was reached. Once the module was deemed stable, it was kept outdoors in the shade, but uncovered, for the time of setting up the measurement and cooling down the module with tap water. Then, it was calibrated straightaway.

2.2.4. AIST

Measurements at AIST used a Large Area Multi-junction module Solar simulator (LAMS; WACOM WPSS – 2.0 × 1.4–50 × 6, AM1.5GF), which generates long pulse lights of 100 ms to1000 ms by xenon short arc lamps. All measurements were carried out within a single pulse of light. Typical time needed for one I-V curve measurement was 300 ms. The irradiance was measured based on reference cells calibrated at AIST. Spectral mismatch correction was carried out for all the modules, based on their module spectral responsivity measured at AIST and the spectral irradiance of LAMS. The module temperature was measured with nine Pt100 sensors attached to the back sheet and confirmed to be 25 °C ± 0.5 °C

throughout the module. A detailed description of the measurement procedures is given in (Dirnberger et al., 2014).

As preconditioning prior to the measurements, the thin-film modules were light-soaked outdoors at 0.5–1.0 kW/m^2 for a total amount of 13 h divided on two consecutive days, and at module temperature in the range (40–60) °C under open-circuit conditions. The measurements were performed twice (the first one after light-soaking for 7 h on the first day and the second one after the total amount of 13 h completed on the second day) to confirm reproducibility of the performance of the modules for preconditioning. Each I-V measurement was performed immediately within an hour after the light-soaking procedure. The deviations of the performance of the modules between the two measurements were within ±0.2% and ±0.1% for CI(G)S and CdTe modules, respectively. These values are included in the measurement uncertainty.

2.3. Methodology for data analysis

The PV calibration results are compared for the four main electrical parameters: the short-circuit current I_{SC}, the open-circuit voltage V_{OC}, the power at maximum power point P_{MAX} and the fill factor *FF*.

A similar approach to the one used in the previous intercomparison (Dirnberger et al., 2014) is applied. The weighted mean (see Eq. (1)), which is taken as the best estimate of the real value, is calculated for each measurand I_{SC}, V_{OC}, P_{MAX} and *FF* from the results submitted by each laboratory, by weighting them with the corresponding stated uncertainty (UC). The overview of all the stated uncertainties for each laboratory is given in Table 2.

$$\bar{x} = \frac{\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}} \tag{1}$$

where σ_i is half of the UC at 95% of confidence (k = 2) stated for each measurement x_i, i.e. it is equal to the combined standard UC. The expanded combined UC (k = 2) of the weighted mean is calculated as follows:

$$UC_{\bar{x}} = \frac{2}{\sqrt{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}} \tag{2}$$

with $\bar{x}$ and σ_i as given above. An important consequence of these definitions is the fact that the smaller the UC stated by a laboratory the larger the weight it gains within the comparison towards the others. This causes the weighted mean to be pulled towards that laboratory's result. Furthermore, a very small UC introduced in the weighted mean calculation will lead to an even lower UC of the weighted mean itself (see Eq. (2)), because the UC of the latter is always lower than any single measurement's UC. Therefore, an accurate and sensible UC calculation is always of primary importance in order to not artificially bias the result of inter-laboratory comparisons, and ultimately the result of every calibration.

A statistical analysis based on E_n numbers calculation (ISO, 2010), instead of the *z*-score evaluation used in (Dirnberger et al., 2014), is then applied to the measurement results in comparison to their weighted mean.

The E_n number is defined as:

$$E_{n_i} = \frac{x_i - x_{ref}}{\sqrt{(UC_i)^2 + (UC_{ref})^2}} \tag{3}$$

where x_i is the measured value to be evaluated, x_{REF} is the reference value, UC_i is the expanded combined uncertainty (k = 2) stated for the value x_i and UC_{REF} is the expanded combined uncertainty for x_{REF}. For the case of the intercomparison presented here, x_i is the measurement result for one of the four measurands and from one of the laboratories, and x_{REF} is the weighted mean for that measurand. The *z*-score approach used in (Dirnberger et al., 2014) and the E_n number approach are therefore equivalent if the standard deviations of the first are equal to half of the measurements expanded combined uncertainties, i.e. if the coverage factor *k* of the uncertainty distributions is equal to 2.

The absolute value of the E_n number is often used for evaluating the performance of comparison results; here we prefer to keep the sign to see more easily in which direction the deviations are. According to the E_n number approach, the output of a measurement agrees with the reference value within the stated uncertainties if the corresponding E_n number is between −1 and 1. As the uncertainty of the difference between the measured and the reference values is the expanded combined uncertainty calculated from the individual expanded combined uncertainties, the confidence of the agreement between the measured and the reference values (i.e. the weighted mean in our case) is given with 95% probability.

3. Results

Table 3 shows the results of the intercomparison exercise. The weighted mean of each measurand, as calculated according to Eq. (1), is reported together with its expanded combined UC (k = 2) for each module.

Table 2
Stated UCs for all the measurands and every laboratory.

CalLab					ESTI				
Device	I_{SC} [%]	V_{OC} [%]	P_{MAX} [%]	*FF* [%]	Device	I_{SC} [%]	V_{OC} [%]	P_{MAX} [%]	*FF* [%]
M1	1.5	0.6	2.2	2.0	M1	0.8	1.4	1.8	0.72
M2	1.3	0.6	1.6	1.4	M2	0.8	1.4	1.8	0.72
M3	1.3	0.6	1.6	1.4	M3	0.8	1.4	1.8	0.72
M4	1.3	0.6	1.6	1.4	M4	0.8	1.4	1.8	0.72
M5	2.5	0.8	3.2	2.1	M5	0.8	1.4	1.8	0.72
M6	2.5	0.8	3.2	2.1	M6	0.8	1.4	1.8	0.72
M7	2.5	0.7	2.9	1.7	M7	0.8	1.4	1.8	0.72
NREL					AIST				
Device	I_{SC} [%]	V_{OC} [%]	P_{MAX} [%]	*FF* [%]	Device	I_{SC} [%]	V_{OC} [%]	P_{MAX} [%]	*FF* [%]
M1	0.7	0.6	1.1	1.4	M1	1.4	0.4	1.5	0.8
M2	0.7	0.6	1.1	1.4	M2	1.4	0.4	1.5	0.8
M3	0.7	0.6	1.1	1.4	M3	1.4	0.4	1.5	0.8
M4	0.7	0.7	1.2	1.5	M4	1.4	0.4	1.5	0.8
M5	3.2	0.6	3.2	1.2	M5	1.4	0.6	1.7	0.9
M6	3.2	0.6	3.2	1.2	M6	1.5	1.1	1.8	1.2
M7	3.2	0.7	3.2	1.2	M7	1.6	0.6	1.9	0.9

Table 3
Calculated weighted means and their expanded combined UCs (for k = 2) for each module.

Device	Weighted mean of the measurand				Expanded combined uncertainties (k = 2)			
	I_{SC} [A]	V_{OC} [V]	P_{MAX} [W]	*FF* [%]	UC (I_{SC})	UC (V_{OC})	UC (P_{MAX})	UC (*FF*)
M1	7.360	43.27	240.28	75.37	0.47%	0.28%	0.75%	0.48%
M2	8.303	44.35	273.30	74.28	0.46%	0.28%	0.71%	0.47%
M3	6.224	64.97	314.56	77.81	0.46%	0.28%	0.71%	0.47%
M4	8.655	37.00	238.71	74.62	0.46%	0.29%	0.74%	0.47%
M5	2.185	112.25	169.78	69.10	0.66%	0.36%	1.08%	0.49%
M6	3.105	58.63	120.11	65.49	0.66%	0.42%	1.11%	0.52%
M7	2.056	60.25	88.73	71.64	0.67%	0.37%	1.12%	0.48%

3.1. c-Si modules

Fig. 2 gives an overview of the results for c-Si modules in terms of percentage deviations for each module and each laboratory from the weighted means, which correspond therefore to the zero value. The shaded areas around the zero reflect the expanded combined uncertainty calculated for the weighted mean of each measurand and each module.

Short-circuit current - Isc

The deviations from the weighted mean in the I_{SC} of all the c-Si modules are between −0.5% and +0.4%, which is also the deviation range resulting from the results for modules M1 and M2. A slightly smaller spread is found when considering only the high-efficiency modules, for which the measured values differ from the reference one between −0.4% and +0.2%.

Open-circuit voltage - Voc

The V_{OC} values are within −0.5% and +0.4% from the reference value, with an improvement in the overall agreement compared to the previous intercomparison, where the individual measurements differed from the weighted mean within ±0.6% (Dirnberger et al., 2014). This may be explained by a general better control of the temperature during the I-V curve measurements also for those setups with constant illumination (e.g. under natural sunlight), as the uncertainties associated to this measurand are basically unchanged or even reduced with respect to the previous intercomparison.

Maximum power - Pmax

The deviations of the measured values from the reference one are between −0.9% and +0.8%, with ESTI reporting the lower limit for the M2 measurement and CalLab the higher value in the M4 measurement. Re-measurement of M4 by CalLab after the intercomparison showed a change of −0.31% in P_{MAX} compared to the initial measurement. This power loss was correlated to nine additional cell cracks in module M4 shown by Electroluminescence Imaging, performed at CalLab before and after the intercomparison on the four c-Si modules. The damage observed for module M4 is believed to have contributed to the spread in the measurement results.

It is curious to see that the widest spread is reported for the traditional c-Si modules. The high-efficiency ones, indeed, agree with the weighted mean within ±0.5%. Nonetheless, even the largest deviation range from c-Si modules is a significant improvement compared to the ±1.3% of the previous intercomparison (Dirnberger et al., 2014).

Fill factor- FF

The variation of the FF around the weighted mean is within −0.6% and +0.4%. Again, the high-efficiency modules have a smaller spread around the weighted mean compared to the traditional ones. The percentage deviations are indeed within ±0.2% for the high-efficiency devices and between −0.6% and +0.4% for the traditional ones.

Monitoring modules

Fig. 3 shows the percentage deviations for each laboratory and each measurand for the two monitoring modules M1 and M2. The values are listed in Table 4 together with the deviations between the weighted means as resulting from the data analysis of the two intercomparisons, and the individual percentage deviations for each laboratory.

The evaluation of the percentage deviations between the weighted means resulting from the two intercomparisons (Table 4) shows that the calibration of the modules is reproducible with much higher accuracy than that achievable by simply considering the UC of the individual laboratories. This is clearly visible from the very low value of the E_n numbers calculated between the calibration values for the measurands of both modules, reported in the table as well, and it is an important reason for maintaining periodic intercomparison between peer laboratories.

The comparison between the individual laboratories' deviations from the two measurement campaigns shows also a large deviation of NREL results compared to what is generally achieved by the others. Indeed, from the individual deviations from the weighted mean for the previous intercomparison (values listed at the bottom part of Table 4), it is evident that NREL had the largest deviations for these two modules, although the overall agreement was satisfactory according to the *z*-score analysis. The better agreement obtained within the current intercomparison may be related therefore mainly to a better alignment of NREL's results to the general measurements outcome. The high variation for the measurement from NREL and the better agreement with the other laboratories in the current intercomparison can both be explained by the change in methodology at NREL. While in the previous intercomparison just the results from one test bed (continuous solar simulator) were reported, here the consolidated data from three test beds (as described above) are used.

3.2. CI(G)S

Fig. 4 gives an overview of the results for CI(G)S modules in terms of percentage deviations for each module and each laboratory from the weighted means, which correspond to the zero value. The shaded areas around the zero reflect the expanded combined uncertainty calculated for the weighted mean of each measurand and each module.

Short-circuit current - Isc

The deviations in I_{SC} for the CI(G)S modules are within −1.9% and +1.7%, with NREL reporting the lowest value for the traditional CI(G)S and AIST the highest one for the module with extended SR. The latter is also the only value where the error bar does not cross the weighted average.

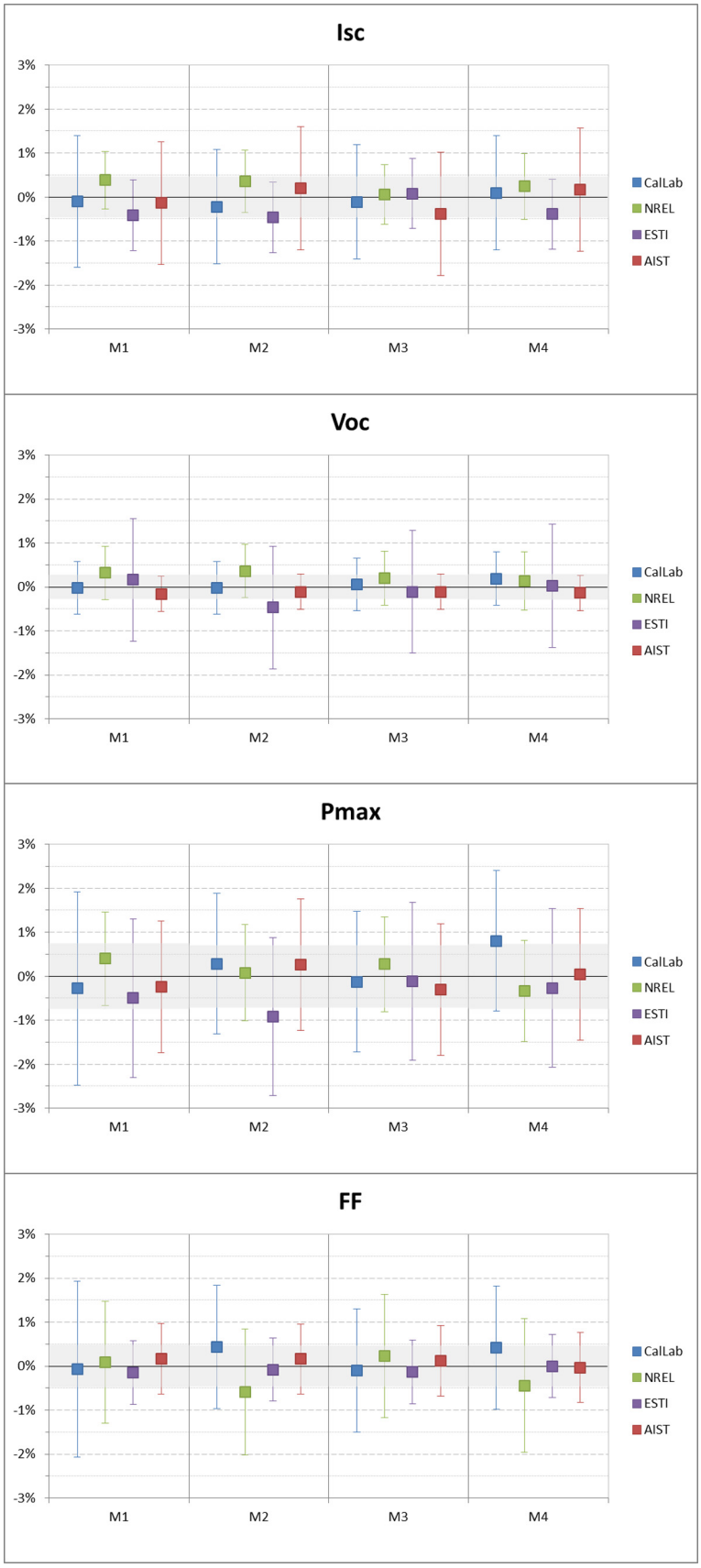

Fig. 2. Summary of results for c-Si modules. For each laboratory, the percentage deviation from the weighted mean is shown, with the error bars representing the stated expanded combined UC for the submitted result, and the shaded area around the zero represents the expanded combined UC of the weighted mean for each measurand.

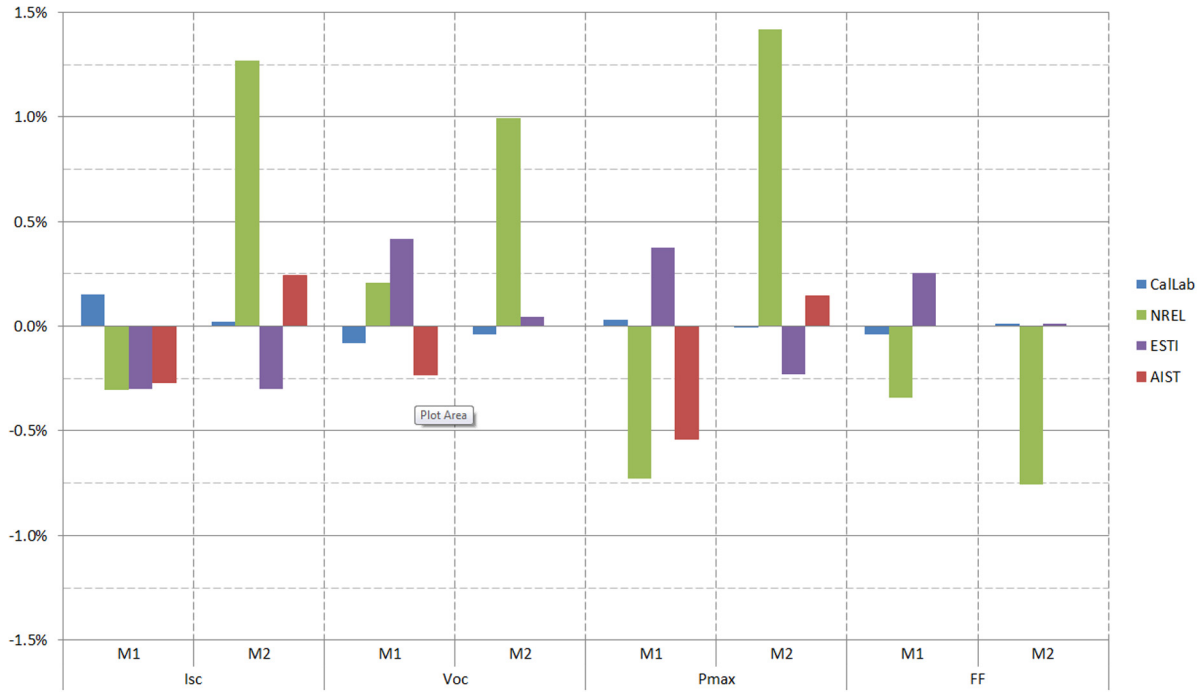

Fig. 3. Comparison of the monitoring modules. The deviations, explicitly reported in Table 4, are given for each laboratory as percentage with respect to the measurement results of the previous intercomparison (Dirnberger et al., 2014).

Table 4
Percentage deviations of measurement results from the current intercomparison with respect to results submitted by each laboratory for the previous publication (Dirnberger et al., 2014). The percentage deviation between the calculated weighted means of the current intercomparison and the previous one is also reported, together with the E_n numbers that can be calculated from them and a summary table of the previous percentage deviations.

Laboratory	I_{SC}		V_{OC}		P_{MAX}		*FF*	
	M1	M2	M1	M2	M1	M2	M1	M2
Percentage deviations in laboratories measurement results of present intercomparison compared to previous intercomparison (Dirnberger et al., 2014)								
CalLab	0.15%	0.02%	−0.08%	−0.04%	0.03%	0.00%	−0.04%	0.01%
NREL	−0.31%	1.27%	0.21%	0.99%	−0.73%	1.42%	−0.34%	−0.76%
ESTI	−0.30%	−0.30%	0.42%	0.05%	0.37%	−0.23%	0.25%	0.01%
AIST	−0.27%	0.24%	−0.23%	0.00%	−0.54%	0.15%	0.00%	0.00%
Weighted means from the two intercomparisons								
Dirnberger et al. (2014)	7.355	8.286	43.30	44.29	239.80	272.90	75.30	74.34
UC (2014)	0.90%	0.82%	0.26%	0.26%	1.16%	1.04%	0.51%	0.49%
Today	7.360	8.303	43.27	44.35	240.28	273.30	75.37	74.28
UC (2017)	0.47%	0.46%	0.28%	0.28%	0.75%	0.71%	0.48%	0.47%
% deviation	*0.07%*	*0.21%*	*−0.07%*	*0.14%*	*0.20%*	*0.15%*	*0.09%*	*−0.08%*
E_n number	0.07	0.22	−0.18	0.35	0.14	0.12	0.13	−0.12
Relative deviations from weighted mean from the previous intercomparison								
CalLab	−0.18%	−0.03	0.00%	0.14%	−0.12%	0.42%	0.06%	0.34%
NREL	0.76%	−0.68%	0.06%	−0.50%	1.33%	−1.18%	0.53%	0.08%
ESTI	−0.06%	0.05%	−0.31%	−0.39%	−0.67%	−0.56%	−0.31%	−0.18%
AIST	0.20%	0.17%	0.01%	0.01%	0.49%	0.25%	0.26%	0.08%

Open-circuit voltage - Voc

For the open-circuit voltage the deviations are between −1.3% and +0.6%, with CalLab value for the module M6 being the farthest value from the weighted average. CalLab's V_{OC} is lower than the other measurements for the two CI(G)S modules; this might be related to the specific preconditioning procedure applied to the modules.

Maximum power - Pmax

The P_{MAX} deviations from the reference value are within −0.8% and +0.4% for the traditional CI(G)S (M5), with E_n numbers within [−1,1] for all the partners. For the module with extended SR (M6) the spread between the measurements is larger and between −1.8% (CalLab) and +2.2% (AIST), which looks in fact to be singular compared to the other three measurements for this module. The distribution of percentage deviations for P_{MAX} seems to reflect mainly the I_{SC} behaviour, even if for CalLab the opposite deviations in I_{SC} and V_{OC} partly cancel out and produce a better agreement in P_{MAX} with ESTI and NREL. On the whole, considering the E_n number assessment, the agreement between the laboratories is good also for this technology, as AIST measurement results in E_n number equal to 1.00 (see Table 5a in Section 4).

Fill factor - FF

For the deviations in *FF*, the opposite situation occurs. Indeed, a better agreement is found for the extended-SR module, for which the percentage deviations from the weighted average range between −0.2% and +0.6%, with AIST slightly different compared to the other three measurements that look well aligned to each other. On the contrary, for the traditional CI(G)S module a more

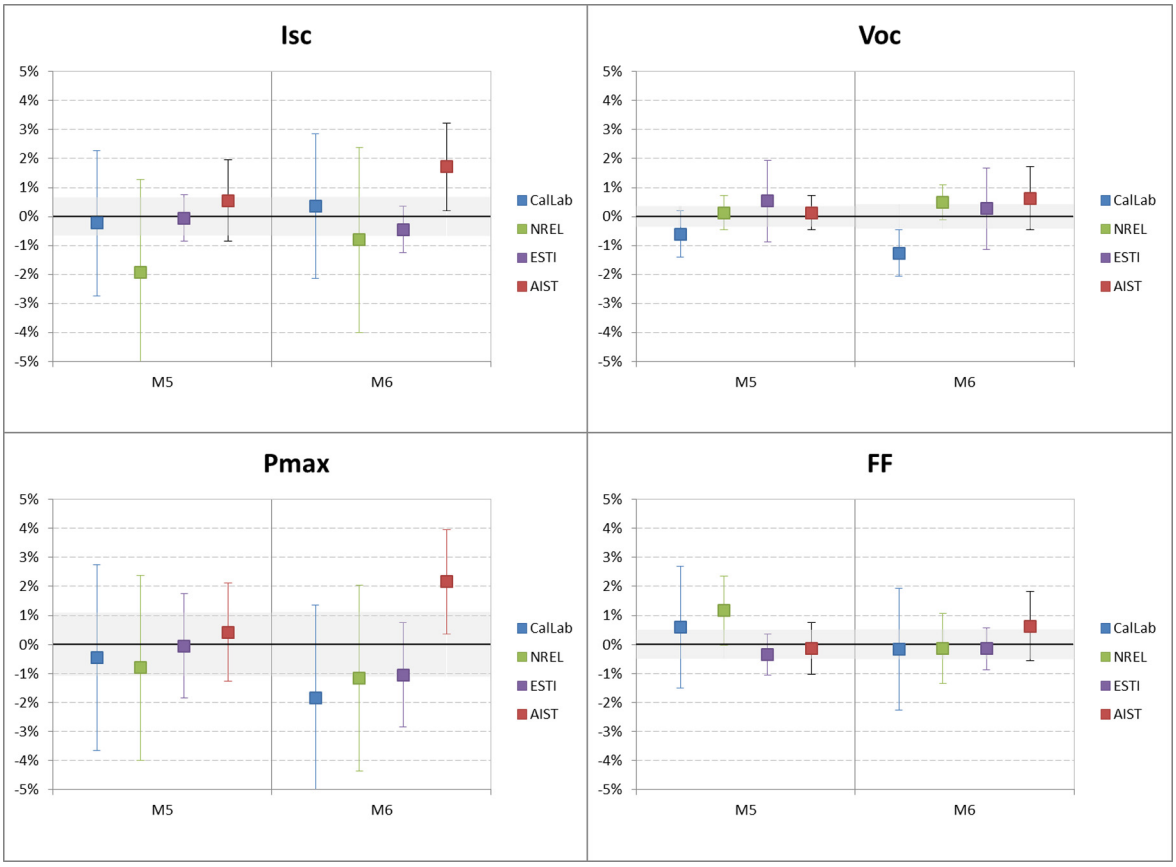

Fig. 4. Summary of results for CI(G)S modules. For each laboratory, the percentage deviation from the weighted mean is shown, with the error bars representing the stated expanded combined UC for the submitted result, and the shaded area around the zero represents the expanded combined UC of the weighted mean for each measurand. Note the different vertical scale as compared to Fig. 2.

variable situation is visible, with CalLab and NREL above the reference value, and ESTI and AIST below it, with an overall spread for the M5 module resulting between −0.4% and +1.2%. However, all the laboratories agree with the weighted mean value within the stated UCs, as confirmed by the fact that all E_n numbers are within $[-1,1]$ for this measurand (see Table 5a in Section 4).

3.3. CdTe

Fig. 5 gives an overview of the results for the CdTe module in terms of percentage deviations for each laboratory from the weighted mean, which corresponds to the zero value. The shaded area around the zero reflects the expanded combined uncertainty calculated for the weighted mean of each measurand.

As the proprietary protocol for preconditioning of the module was not made available to all partners of this intercomparison, also for the CdTe module the four laboratories applied the preconditioning procedure that they would follow for accredited PV calibrations.

Short-circuit current - Isc

The percentage deviations in short-circuit current from the weighted mean are imbalanced towards positive values, being within −0.2% and +0.4%. In comparison to the previous measurement campaign, the agreement between the laboratories has significantly improved, as proved also by a reduction of the expanded combined uncertainty of the weighted mean from a value of 1.0% in the previous intercomparison to 0.67% for the current one. As shown in (Dirnberger et al., 2014) by intermediate measurements at CalLab, the short-circuit current is almost unaffected by metastabilities that can occur in this type of module. Furthermore, the individual uncertainty applicable to the present I_{SC} measurements was either unchanged or even reduced by the laboratories (see Table 2 and (Dirnberger et al., 2014)). Therefore, the improvement shown here can be connected mainly to a strong reduction of the percentage deviation in the AIST measurement from the weighted mean, and so to an overall better agreement between the laboratories, which is a very positive outcome for this thin-film technology.

Open-circuit voltage - Voc

The percentage deviations in V_{OC} are on the contrary quite large compared not only to the I_{SC} ones for this module, but also to the same measurand for all the other modules. Indeed, the percentage deviations in V_{OC} for the CdTe module are the largest among the open-circuit voltage results. They range from −1.4% to +1.9%, with all the laboratories but CalLab below the weighted average, although AIST is very close to it. This could be related to the different preconditioning procedure applied by the laboratories, as the V_{OC} is more sensitive than the short-circuit current to the temperature and to the electrical-load conditions set during the light exposure and to the overall irradiation (Cueto and von Roedern, 2006). It is interesting to notice that, although ESTI and AIST have quite different UCs associated to this measurement, their results are closer to the average than those given by CalLab and NREL, and they both agree with the weighted average within their UCs. This is explained partly by the percentage deviations for CalLab's and NREL's results, which are almost symmetric around the average with the same stated UC for this measurand.

In addition, although this cannot be stated with full confidence due to the lack of intermediate measurements at CalLab, there seems to be a trend in the measurements with the temporal

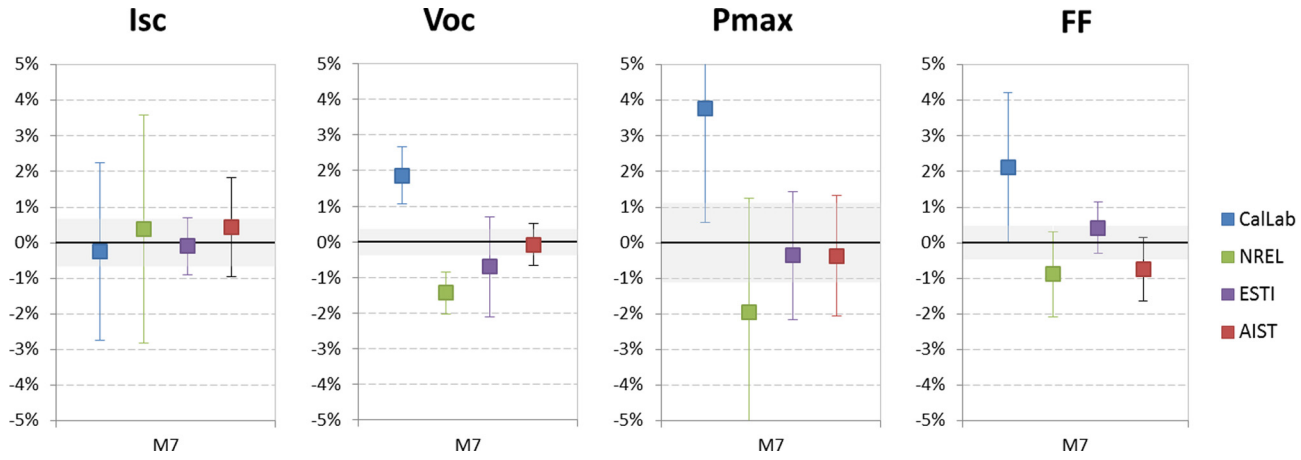

Fig. 5. Summary of results for the CdTe module. For each laboratory, the percentage deviation from the weighted mean is shown, with the error bars representing the stated expanded combined UC for the submitted result, and the shaded area around the zero represents the expanded combined UC of the weighted mean for each measurand. Note the vertical scale is the same as in Fig. 4, but different as compared to Fig. 2.

sequence of laboratories, as if the module were in fact not fully stable during the intercomparison and possibly insufficiently preconditioned. The first hypothesis is considered very likely, as CdTe is known to be sensitive to dark-storage, which influences negatively the open-circuit voltage (Cueto and von Roedern, 2006). In comparison to the previous measurement campaign, though, this trend seems to go in the opposite direction than the expected one, i.e. V_{OC} increases. As the laboratories' results in (Dirnberger et al., 2014) look well aligned to the preconditioned intermediate measurements at CalLab, with a long-term decrease in V_{OC}, we can expect a similar agreement also in this campaign, with a possible positive cumulative effect due to the actual preconditioning procedures that were applied in this measurement campaign.

Maximum power - Pmax

The deviations in P_{MAX} for the CdTe module are the largest of the whole intercomparison. They range from –2.0% to +3.8% and are affected mainly by the deviations in V_{OC}, although in case of CalLab there seems to be some negative contribution also from factors influencing the *FF*. The results for P_{MAX} agree with the weighted average within their stated UCs, with the exception of CalLab. This is represented in the E_n number assessment, for which NREL, AIST and ESTI have $|E_n|$ values well below 1, with CalLab's E_n value equal to 1.17. A possible influence of the spectral mismatch component, which may usually affect the calibration of a CdTe PV device due to the lack of a well-matched reference cell, is likely to be insignificant in this case due to the very good agreement between the I_{SC} measurements. The latter were indeed obtained with reference cells following different traceability chains to the SI, and thus through different calibration procedures and primary standards.

Fill factor- FF

The percentage deviations in *FF* are between −0.9% and +2.1%, showing a relatively large imbalance towards positive values, due largely to the very good agreement between NREL and AIST results and the still singular result submitted by CalLab. As consequence, CalLab result does not match the weighted average value within the stated UC and its E_n number assessment is negative, with E_n value above 1 (see Table 5a in Section 4).

4. Discussion and conclusions

Tables 5a and 5b summarise all the E_n numbers as calculated according to Eq. (3). Among all the measurements included in this world-wide intercomparison, 5% of E_n number values are out of the acceptable range of $[-1,1]$ (see Table 5a). This means that 95% of the measurements agree within the stated uncertainty with the weighted mean, which is assumed to be the best estimate of the real value. This amount reflects the 95% level of confidence that is represented by the UC statement itself, implying that the stated UCs satisfactorily represent the measurement accuracy capability of the peer laboratories involved in the intercomparison.

If we analyse in detail the different cases according to the technology, all the measurements reported for c-Si have corresponding E_n numbers within ±0.6. Additionally, these measurements have a stated UC that has been generally reduced by all the laboratories compared to the previous intercomparison (Dirnberger et al., 2014), with few exceptions. Therefore, c-Si PV calibration is proved to have reached a very good confidence level for both traditional and high-efficiency modules, with generally reduced uncertainties compared to just few years ago. In the end, this may also be translated into a better financial risk assessment for the final users' investment in the technology.

As regards thin-film modules, further improvement is necessary in order to obtain the same level of agreement in the measurements. The number of E_n values outside the $[-1,1]$ range correspond to slightly more than 6% of the CI(G)S measurements and up to 25% for the CdTe module, although the total number of measurements involved in both cases is not high.

In the CI(G)S case, one value ($E_n = -1.41$) is due to the measurement of V_{OC} by CalLab and the other ($E_n = 1.03$) to the I_{SC} measurement by AIST, both for module M6. However, it has to be noticed that AIST value is anyway very close to 1 and that it is reported for a module type that has been measured at AIST for the first time during this intercomparison. It is noted that factors other than the preconditioning procedure, such as possible persistent change in the performance which is not levelled by usual preconditioning, cannot be ruled out at the present stage. In general, improvements are also expected for the future for thin-film modules, as already shown for the c-Si ones.

In contrast, the CdTe module has a quite clear outcome. Three of the four E_n values outside the range $[-1,1]$ are reported for the CalLab's calibration, showing a significant deviation of all the parameters but the I_{SC} from the weighted average. This is likely due to the different preconditioning procedure applied at this laboratory compared to what has been used in the others. Unfortunately, despite the shared conviction of the importance of a common and validated procedure to be used in the intercomparison, the proprietary protocol was not made available to all the labora-

tories. While this fact may have contributed to the large variation in results between laboratories, it also reflects the reality better, as each independent reference laboratory should use an independently validated preconditioning method and cannot rely on proprietary methods. The latter are indeed developed mainly for specific commercial purposes, and do not necessarily meet the same requirements that are included in international and independently validated standards, which should be adopted, especially but not exclusively, by calibration laboratories.

As an overall result for the CdTe module, it has to be highlighted once more that the preconditioning procedure may strongly affect the result of PV calibration for this technology. Additionally, the applied preconditioning should lead to a final stabilisation level more similar to what is observed in the field, instead of setting the device to an artificial better-performing level with less correspondence with open-field performance. At the time of the measurements the standard IEC 61646 ed. 2 (IEC, 2008) was in force, which only provided a generic preconditioning for all thin-film technologies. In the meanwhile, this standard has been discontinued and replaced by technology-specific standards within the series IEC 61215 (IEC, 2016a,b,c,d,e). It is therefore strongly suggested that for future intercomparisons all laboratories should perform the preconditioning according to the Part of the standard IEC 61215 relevant for the module technology.

Looking at Figs. 2, 3 and Table 5a it is clear that the agreement for c-Si is very good. For this module technology, the quoted UCs seem to be even rather conservative and could be reduced. For example, the scatter in data for P_{MAX} is much narrower than the ranges of UCs quoted. Another example is the UC of V_{OC} for ESTI, which is much larger than the deviation from the weighted average. In fact, at ESTI the large UC for V_{OC} is due to the limited temperature control under natural sunlight combined with a conservative estimate of temperature deviations. However, based on this intercomparison (and further experience of several years with such measurements under natural sunlight) it is expected to give more realistic values for the temperature deviations and thereby officially reduce the UC for V_{OC} at ESTI in the near future. From a visual inspection of Fig. 2 it can be estimated that the following uncertainties could be halved without affecting the agreement with the weighted average: UC for I_{SC} of CalLab and AIST, UC for V_{OC} for ESTI, UC for P_{MAX} of all four laboratories and UC of *FF* for CalLab and NREL. Part of this will be due to the change and improvement of measurement procedure at NREL, which has improved international agreement between the laboratories as shown here.

As the c-Si modules can be assumed to be perfectly stable, the good agreement between the laboratories observed for these modules indicates that the measurement procedures themselves are well under control and give this high level of agreement internationally. The lower level of agreement for the thin-film technologies is then probably due to the preconditioning and/or some metastability of these modules, as the other measurement procedures in general follow those used for c-Si modules. Possibly the spectral mismatch and its correction would be the only aspect significantly different between thin-film and c-Si modules. Discrepancies between the laboratories should in this case show up mostly in I_{SC} (and as a consequence in P_{MAX}). However, from Figs. 4, 5 and Table 5a the agreement in I_{SC} is satisfactory. In fact, CdTe has the SR most different from traditional c-Si, which is commonly used as reference device, and would therefore be expected to show most likely problems due to spectral mismatch correction; but it does not. On the other hand, the disagreement between laboratories for thin-film modules is mostly in V_{OC} and *FF* (and consequently also in P_{MAX}). It is well known that for the thin-film technologies the changes with preconditioning mainly occur in *FF* and V_{OC}, whereas I_{SC} is relatively unaffected. So, again we can conclude that the discrepancies between laboratories are most likely due to the preconditioning, rather than actual measurement procedures.

Based on the evaluation of the overall results for the CdTe module, the results of CalLab for this module can be considered outliers due to different preconditioning procedure. Removing these results (see Table 5b) eliminates not only three cases of E_n outside the range [−1,1] (for V_{OC}, P_{MAX} and *FF* of CalLab for M7), but also the outlier of NREL in V_{OC} for this module (due to the recalculation of weighted mean). Hence, there are only two outliers left: V_{OC} for M6 at CalLab and I_{SC} for M6 at AIST (the latter being just above 1 and with rounding to one decimal as considered in the IEC standard it would be equal to 1.0). These two data points are considered real outliers. They amount to less than 2% of all data points, but also to less than 5% of just the data points for the thin-film modules as a whole. This is then in line also for the thin-film modules with expectations that the analysis of E_n numbers should yield values in the range [−1,1] for 95% of all data points.

As a consequence, the UCs for the thin-film modules might need to be generally reconsidered by the laboratories, possibly including components related to the preconditioning itself. The observed deviations ($|E_n| > 1$) might be due to the laboratories only/mainly considering the UCs of the actual measurement procedure, but excluding variations due to the different state the module is actually in during the calibration measurement. Metastabilities are much less relevant in the field, where modules are permanently installed and have a regular irradiance. For calibration laboratories the situation is different, as modules are generally kept in the dark (during transport and storage), thus making the effect of metasta-

Table 5a
E_n numbers calculated for each reported measurand based on Eq. (3). The values outside the acceptance range [−1,1] are in bold.

	$E_n(Isc)$	$E_n(Voc)$	$E_n(Pmax)$	$E_n(FF)$	$E_n(Isc)$	$E_n(Voc)$	$E_n(Pmax)$	$E_n(FF)$
	CalLab				ESTI			
M1	−0.06	−0.03	−0.12	−0.03	−0.46	0.11	−0.26	−0.17
M2	−0.16	−0.03	0.16	0.30	−0.50	−0.33	−0.48	−0.09
M3	−0.08	0.09	−0.08	−0.07	0.09	−0.08	−0.06	−0.15
M4	0.07	0.28	0.45	0.28	−0.43	0.02	−0.14	0.00
M5	−0.09	−0.70	−0.13	0.27	−0.04	0.36	−0.03	−0.40
M6	0.14	**−1.41**	−0.55	−0.08	−0.44	0.19	−0.50	−0.17
M7	−0.09	**2.31**	**1.17**	**1.17**	−0.09	−0.48	−0.17	0.48
	NREL				AIST			
M1	0.47	0.48	0.30	0.06	−0.09	−0.33	−0.14	0.18
M2	0.43	0.54	0.06	−0.39	0.14	−0.23	0.15	0.17
M3	0.08	0.30	0.21	0.16	−0.26	−0.22	−0.18	0.12
M4	0.28	0.19	−0.25	−0.28	0.12	−0.27	0.02	−0.03
M5	−0.60	0.19	−0.24	0.88	0.35	0.19	0.21	−0.14
M6	−0.25	0.67	−0.35	−0.11	**1.03**	0.53	1.00	0.48
M7	0.12	**−1.84**	−0.59	−0.70	0.25	−0.12	−0.17	−0.74

Table 5b
E_n numbers re-calculated for module M7 when excluding the results from CalLab. No value outside the acceptance range [−1,1] is now present.

	$E_n(Isc)$	$E_n(Voc)$	$E_n(Pmax)$	$E_n(FF)$	$E_n(Isc)$	$E_n(Voc)$	$E_n(Pmax)$	$E_n(FF)$
	CalLab				ESTI			
M7	–	–	–	–	−0.14	−0.02	0.11	0.68
	NREL				AIST			
M7	0.10	−0.96	−0.40	−0.55	0.22	0.79	0.10	−0.56

bility on the module performance more relevant. Therefore, preconditioning procedures play a vital role in this context as demonstrated by the development of international standards specific to more numerous technologies (IEC, 2016a,b,c,d).

In general, the conclusion from this intercomparison is that the agreement for c-Si PV modules between international reference laboratories (with independent traceability chains to SI units) has been further improved compared to previous exercises. The observed agreement is even better than that suggested by quoted UCs, which should trigger a reduction of the UCs quoted by each laboratory. For thin-film technologies the indications are that for these the actual measurement procedures agree internationally, with discrepancies in results that are due mainly to the metastable nature of these modules. Therefore, preconditioning procedures need to be well-established and rigorously implemented. With the new standards of the series IEC 61215 (IEC, 2016a,b,c,d,e) published by now, the expectation is that future intercomparisons applying IEC requirements for module's preconditioning will show an improved agreement also for thin-film technologies.

The present calibration campaign, included in the required activities for laboratories accredited for calibration according to the ISO 17025 (ISO, 2005), emphasises the importance of a regular comparison between peer laboratories in order to assess and share the best measurement practice in PV calibration, and to spot possible areas of improvement in the measurement procedures as well as in the UCs associated to specific technologies. As further outcome of this measurement exercise, it seems evident to the authors that in the future more effort should be put into assessing the measurement capabilities for well-established thin-film technologies like amorphous Si, CdTe and CI(G)S, while keeping c-Si devices more as historical references between the intercomparisons. Also, additional parameters like the temperature coefficients and the performance at different irradiance levels, which are important for the developing energy rating needs, should become regular subjects of these international intercomparisons for both silicon and thin-film PV devices.

Acknowledgments

ESTI authors thank Ewan Dunlop, ESTI laboratory manager, and Tony Sample for helpful discussion on the measurements and the overall intercomparison. They also acknowledge help by Germana Trentadue in the PV performance measurements under natural sunlight.

The work at NREL was supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Solar Energy Technologies Office under Contract No. DE-AC36-08GO28308 with Alliance for Sustainable Energy, LLC. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

The NREL authors acknowledge important contributions by Alan Anderberg in performing module measurements and Larry Ottoson in primary calibrations and module test bed hardware and software.

The work at AIST was supported by New Energy and Industrial Technology Development Organization (NEDO) under Minister of Economy, Trade and Industry (METI), Japan.

References

Bahrs, U., Zaaiman, W., Mertens, M., Ossenbrink, H., 1997. Automatic large area spectral response facility. In: WIP (Ed.) 14th European Photovoltaic Solar Energy Conference and Exhibition. Barcelona, Spain, pp. 2296–2298.

Campanelli, M.B., Osterwald, C.R., 2016. Effective irradiance ratios to improve I-V curve measurements and diode modeling over a range of temperature and spectral and total irradiance. IEEE J. Photovolt. 6 (1), 48–55.

Cueto, JAd., von Roedern, B., 2006. Long-term transient and metastable effects in cadmium telluride photovoltaic modules. Prog. Photovoltaics Res. Appl. 14 (7), 615–628.

Dirnberger, D., Kräling, U., 2013. Uncertainty in PV module measurement–Part I: calibration of crystalline and thin-film modules. IEEE J. Photovolt. 3 (3), 1016–1026.

Dirnberger, D., Kräling, U., Müllejans, H., Salis, E., Emery, K., Hishikawa, Y., Kiefer, K., 2014. Progress in photovoltaic module calibration: results of a worldwide intercomparison between four reference laboratories. Meas. Sci. Technol. 25 (10), 105005.

IEC, 2008. IEC 61646. Thin-Film Terrestrial Photovoltaic (PV) Modules – Design Qualification and Type Approval. International Electrotechnical Commission, Geneva.

IEC, 2016a. IEC 61215-1-1 Terrestrial Photovoltaic (PV) Modules - Design Qualification and Type Approval - Part 1–1: Special Requirements for Testing of Crystalline Silicon Photovoltaic (PV) Modules. International Electrotechnical Commission, Geneva.

IEC, 2016b. IEC 61215-1-2. Terrestrial Photovoltaic (PV) Modules - Design Qualification and Type Approval - Part 1–2: Special Requirements for Testing of Thin-Film Cadmium Telluride (CdTe) Based Photovoltaic (PV) Modules. International Electrotechnical Commission, Geneva.

IEC, 2016c. IEC 61215-1-3. Terrestrial Photovoltaic (PV) Modules - Design Qualification and Type Approval - Part 1–3: Special Requirements for Testing of Thin-Film Amorphous Silicon Based Photovoltaic (PV) Modules. International Electrotechnical Commission, Geneva.

IEC, 2016d. IEC 61215-1-4. Terrestrial Photovoltaic (PV) Modules - Design Qualification and Type Approval - Part 1–4: Special Requirements for Testing of Thin-Film Cu(In, GA)(S, Se)2 Based Photovoltaic (PV) Modules. International Electrotechnical Commission, Geneva.

IEC, 2016e. IEC 61215–1 Terrestrial Photovoltaic (PV) Modules - Design Qualification and Type Approval - Part 1: Test Requirements. International Electrotechnical Commission, Geneva.

ISO, 2005. ISO/IEC 17025:2005 General Requirements for the Competence of Testing and Calibration Laboratories. International Organization for Standardization, Geneva.

ISO, 2010. ISO/IEC 17043:2010 Conformity Assessment – General Requirements for Proficiency Testing. International Organization for Standardization, Geneva.

Levi, D.H., Osterwald, C.R., Rummel, S., Ottoson, L., Anderberg, A., 2017. Self-Reference Procedure to Reduce Uncertainty in Module Calibration, 44th IEEE Photovoltaic Specialists Conference (PVSC) 2017. Washington, DC.