

Round robin performance testing of organic photovoltaic devices

Suren A. Gevorgyan^{a,*}, Oihana Zubillaga^b, José María Vega de Seoane^b, Maider Machado^b, Elif Alturk Parlak^c, Nesrin Tore^c, Eszter Voroshazi^d, Tom Aernouts^d, Harald Mülleijans^e, Giorgio Bardizza^e, Nigel Taylor^e, Wiljan Verhees^f, Jan M. Kroon^f, Pasquale Morvillo^g, Carla Minarini^g, Francesco Roca^g, Fernando A. Castro^h, Stéphane Crosⁱ, Balthazar Lechêneⁱ, Juan F. Trigo^j, Cecilia Guillén^j, Jose Herrero^j, Birger Zimmermann^k, Subarna Babu Sapkota^k, Clemens Veit^k, Uli Würfel^k, Pabitra S. Tuladhar^l, James R. Durrant^l, Stefan Winter^m, Sanna Rousuⁿ, Marja Välimäkiⁿ, Volker Hinrichs^o, Sarah R. Cowan^p, Dana C. Olson^p, Peter Sommer-Larsen^a, Frederik C. Krebs^a

^a CLOP, Department of Energy Conversion and Storage, Technical University of Denmark, Frederiksborgvej 399, 4000 Roskilde, Denmark

^b TECNALIA Research and Innovation, Mikeletegi Pasealekua 2, 20009 San Sebastian, Spain

^c TÜBİTAK, National Metrology Institute, Tübitak Gebze Campus, Gebze, 41470 Kocaeli, Turkey

^d IMEC, Kapeldreef 75, 3000 Leuven, Belgium

^e Joint Research Centre, European Solar Test Installation, Via Fermi 2749, I-21027 Ispra, VA, Italy

^f ECN/Solliance, High Tech Campus 5, 5656 AE Eindhoven, The Netherlands

^g ENEA, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile, Piazzale E. Fermi 1, 80055 Portici, Italy

^h Materials Division, National Physical Laboratory, Hampton Road, Teddington, Middlesex TW11 0LW, UK

ⁱ National Institute for Solar Energy (INES), CEA/DRT/LITEN/DTS/LMPV, 50 Avenue du Lac Léman, 73377 Le Bourget du Lac Cedex, France

^j CIEMAT, Renewable Energy Division, Edif. 42, Avda. Complutense 40, 28040 Madrid, Spain

^k Fraunhofer Institute for Solar Energy Systems ISE, Heidenhofstrasse 2, 79110 Freiburg, Germany

^l Centre for Plastic Electronics and Department of Chemistry, Imperial College London, London SW7 2AZ, UK

^m Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany

ⁿ VTT Technical Research Centre of Finland, Kaitoväylä 1, 90570 Oulu, Finland

^o Helmholtz-Zentrum Berlin für Materialien und Energie, Lise-Meitner Campus, Hahn-Meitner-Platz 1, 14109 Berlin, Germany

^p National Renewable Energy Laboratory, 1617 Cole Blvd, Golden, CO 80401, USA

ARTICLE INFO

Article history:

Received 13 March 2013

Accepted 24 September 2013

Available online

Keywords:

Organic photovoltaic

Round robin

I–V characterization

Standard testing conditions

Intercomparability

ABSTRACT

This study addresses the issue of poor intercomparability of measurements of organic photovoltaic (OPV) devices among different laboratories. We present a round robin performance testing of novel OPV devices among 16 laboratories, organized within the framework of European Research Infrastructure Project (SOPHIA) and European Energy Research Alliance (EERA). Three types of OPVs with different structures, dimensions and encapsulations are studied and compared with reference Si solar cells certified by accredited laboratories. The agreement of the measurements of these among different laboratories is analyzed by focusing on testing procedures, testing equipment and sample designs. A number of deviations and pitfalls are revealed and based on the analyses, a set of recommendations are suggested for improving the agreement among the measurements of such OPV technologies.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Tackling the global challenges of climate and energy resources requires developing efficient research strategies, which can lead to creation of technologies with the necessary capacity to address these challenges. The European Research Infrastructure Project (SOPHIA)

[1] and the European Energy Research Alliance (EERA) [2] share the common goal to accelerate and optimize the research and developments of photovoltaics (PVs) by improving the interoperability of the European leading research programs and organizations specialized in these technologies. The field of organic photovoltaics (OPVs) has a particular focus within SOPHIA and EERA. One of the initial tasks among others is to improve and harmonize the procedures, parameters and equipment for OPV testing and increase the consistency of test reports among the partners of SOPHIA and EERA.

* Corresponding author.

E-mail address: surg@dtu.dk (S.A. Gevorgyan).

The most powerful tools, when it comes to establishing agreement on measured parameters and validation of test procedures, are the round robin (RR) or interlaboratory study (ILS) methodologies. They not only enable reaching consensus on measured values, but also determine the spread in data and can be used to identify variations inherent in a procedure and in measured parameters. Such intercomparable studies have long been used for inorganic photovoltaic technologies involving accredited laboratories [3–10], while only recently some practices were recorded for OPVs with however no involvement of certified tests by accredited laboratories [11–19]. For energy technologies such as organic solar cells these methodologies are particularly useful and have demonstrated that a significant spread in performance is observed, which questions the accuracy and validity of commonly reported noncertified OPV performances. In the early examples the experiments were typically interlaboratory studies, where devices prepared at one location were tested before and after shipment to a separate laboratory for testing [11]. Later on interlaboratory studies on the stability of devices [12–17] were followed by true round robins, where a single device or a series of devices traveled between laboratories [18]. This has been used to determine the spread in data observed between competent laboratories while having one operator traveling along with the device, which in spite of the effort to reduce variation lead to as much as 25% spread [18]. While many reasons may be speculated for the large spread, the root of the issue seems to be lying within the basics of designs and higher sensitivity of OPVs towards the testing equipment and procedures, when compared to their inorganic counterparts. Examples are specific sensitivity of OPVs towards light spectrum (UV light doping of certain layer inside the device) [20] and degradation of active layer generated by light (photosensitivity of polymeric materials) [21].

The presented work addresses the issue of the large spread by focusing primarily on the design of the tested samples, the testing procedures and measuring equipment. The study encompasses accurate testing and reporting of a number of novel OPV technologies in a round robin cycle among 16 laboratories, organized within the framework of SOPHIA and EERA. The analyses of the

reported results and the pitfalls of the studies are presented and a set of recommendations are outlined.

2. Experimental

2.1. Participating laboratories

The participants of the RR tests were chosen on voluntary basis among SOPHIA and EERA partners during project meetings. Total of 16 laboratories participated in the RR studies. Table 1 presents the list of the laboratories and their affiliation with SOPHIA and EERA. The participants not linked to SOPHIA or EERA are marked as guests.

2.2. Round robin history

The initiative was created at the kick off meeting of SOPHIA and EERA projects on September 2011, where the group from Denmark was assigned to coordinate the RR studies. The list of the 16 participating laboratories was established and Groups #1 (G1) and #2 (G2) were created with 8 and 7 participating laboratories, respectively, for performing measurements in RR cycles within each group. A protocol was created and distributed to each member prior to RR studies and the members were requested to follow the protocol strictly. The protocol contained instructions for performing current–voltage (I – V) characterization of the samples with respect to standard reporting conditions (SRC). This implied that the temperature of the devices had to be kept at 25 °C during I – V testing (since temperature coefficients for photovoltaic parameters of OPVs were unknown and thus, correction of temperature after I – V tests was not possible). The recommended technique was ruling a quick I – V before the sample heats up. The SRC additionally implied calculating the mismatch factor for each sample with respect to the reference device. The copy of the protocol is provided in Section S1 in the Supporting Document. The start of the RR tests was set to February 2012. Four types of test samples (including a Silicon reference cell) were manufactured by different laboratories and after initial testing the samples were sent to the first laboratories in each

Table 1

List of participants in the round robin studies and their affiliation to EERA and SOPHIA.

Laboratory name	Location	Contact persons	Affiliation with SOPHIA/EERA	Activities
CLOP, Technical University of Denmark	Roskilde, Denmark	Suren Gevorgyan, Peter Sommer-Larsen	SOPHIA & EERA	RR Coordination, Measurements, Data Analyses
National Renewable Energy Laboratory	Golden, CO, USA	Dana C. Olson	Guest	Reference Test Sample Supplier, Measurements
Tecnalia Research & Innovation	San Sebastian, Spain	Oihana Zubillaga	SOPHIA	Measurements
TÜBİTAK, National Metrology Institute	Kocaeli, Turkey	Elif Alturk Parlak	EERA	Test Sample Supplier
IMEC	Leuven, Belgium	Eszter Voroshazi	SOPHIA & EERA	Measurements
ESTI – European Commission Joint Research Centre	Ispira, Italy	Giorgio Bardizza, Harald Müllejans, Nigel Taylor	SOPHIA	Measurements
ECN	Petten/Eindhoven, The Netherlands	Jan Kroon	SOPHIA & EERA	Data coordination
ENEA, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile	Portici, Italy	Pasquale Morvillo, Carla Minarini, Francesco Roca	SOPHIA & EERA	Data encrypting
National Physical Laboratory	Teddington, UK	Fernando A. Castro	EERA	Measurements
VTT Technical Research Center of Finland	Oulu, Finland	Sanna Rousu, Tommi Vuorinen	SOPHIA & EERA	Test Sample Supplier, Measurements
Fraunhofer ISE	Freiburg, Germany	Birger Zimmermann	SOPHIA & EERA	Test Sample Supplier, Measurements
Imperial College London	London, UK	Pabitra Shakya Tuladhar	EERA	Measurements
CIEMAT	Madrid, Spain	Juan F. Trigo	EERA	Measurements
CEA INES, Institut National des Energies Solaires	Le Bourget du lac cedex, France	Stéphane Cros, Balthazar Lechêne	SOPHIA & EERA	Measurements
Helmholtz-Zentrum Berlin	Berlin, Germany	Volker Hinrichs	SOPHIA & EERA	Measurements
Physikalisch-Technische Bundesanstalt	Braunschweig, Germany	Stefan Winter	Guest	Measurements

group. Packages of samples were then formed, which would subsequently be transported to and characterized by each laboratory in the cycle until the last laboratory would be reached. A timetable was formed with timeframe for characterization and transport set to two weeks for each participant and the dates were correlated with the participants. This was important, since it would allow the participants to prepare in advance (such as booking of equipment, avoiding interference with local holidays, etc.) and reduce possibilities of delays. In spite of this, some delays were recorded during the tests. Table 2 shows the timetable of the actual characterization dates at each laboratory. The entire study lasted ca. 5–6 months.

The participants additionally received a special reporting form to fill in the measured data. In order to keep the blind format of the studies, each laboratory reported the measured data to Nigel Taylor from the European Commission Joint Research Centre (as an impartial representative). An individual code was then assigned to each report and sent onward to Denmark for data analyses. The reported data of some of the laboratories additionally included the I – V curves and the spectral response data, as well as the spectral distributions of the light sources used for testing. Those were used to assess the accuracy of the setup calibration for performing measurements.

2.3. Sample set

Test samples of organic photovoltaic devices were supplied by Fraunhofer ISE, VTT and TUBITAK (Table 1). While ISE and TUBITAK prepared glass encapsulated single solar cell devices, VTT samples were modules with 8 serially connected solar cells encapsulated in flexible barrier material. Additionally, NREL provided two reference Si solar cells with KG5 filters together with the performance values certified by Fraunhofer ISE Callab (all PV parameters) NREL (all PV parameters) and PTB (short circuit current only). The following labels will be used throughout the text to refer to each device type: Calibrated Si solar cell – “Si”; Fraunhofer ISE OPV device – “A”; VTT OPV module – “B”; TUBITAK OPV device – “C”. Two sets of devices were prepared, one set for each group. Each set contained 7 samples to be measured, one Si, two A, two B and two C samples. Fig. 1 shows the photograph of the samples of each type. The measured solar cells (whole module for B) are highlighted with dashed red squares. The contacts are highlighted with green circles. C samples were designed for only 2-probe measurement, while the other three had a possibility for 4-probe connection. The details of the samples are outlined below:

Sample A:	
Manufacturer:	Fraunhofer ISE
Sample outline:	3 single cells on a rigid substrate. 1 solar cell used for RR measurements
Device structure:	Inverted structure - Glass Substrate/Cr/Al/Cr/P3HT:PCBM/PEDOT:PSS/Grd Au
Coating technique:	Spin-coating
Encapsulation:	Rigid glass encapsulation
Sample dimensions:	Solar cell active area – 1.1 cm ² ; Sample dimensions – 5 cm × 5 cm
Temperature:	Thermocouple attached to the front of the adjacent solar cell
Connections:	4 – terminals (connection via soldered wires)
Sample B:	
Manufacturer:	VTT
Sample outline:	Flexible module with 8 serially connected solar cells. Module was mounted on a rigid backplate.
Device structure:	Normal structure - PET Substrate/ITO/PEDOT:PSS/P3HT:PCBM/Ca/Ag
Coating technique:	Roll-to-roll screen and gravure printing
Encapsulation:	Flexible barrier encapsulation, commercial grade WVTR 10 ^{−3}

(continued)

Sample dimensions:	Module active area–15 cm ² ; Sample dimensions – 4.6 cm × 5.1 cm
Temperature:	–
Connections:	4 – terminals (connection via exposed parts of the printed electrodes)
Sample C:	
Manufacturer:	TUBITAK
Sample outline:	6 solar cells. 2 out of 6 solar cells used for RR measurements.
Device structure:	Normal structure - Glass Substrate/ITO/PEDOT:PSS/PCDTBT:PCBM/Ca/Ag
Coating technique:	Spin-coating
Encapsulation:	Rigid glass encapsulation
Sample dimensions:	Individual solar cells active area–0.04 cm ² ; Sample dimensions – 12 cm × 6 cm
Temperature:	PT100 sensor attached to the bottom of the encapsulation
Connections:	2 – terminals (connection via integrated electric switch circuitry)
Reference Si sample:	
Manufacturer:	–
Sample outline:	ISE Callab calibrated device with WPVS design.
Device structure:	Monocrystalline Si solar cell (with KG5 filter)
Coating technique:	–
Encapsulation:	Rigid glass/metal encapsulation body
Sample dimensions:	Solar cells active area–1 cm ² ; Sample dimensions – 7 cm × 7.5 cm
Temperature:	Integrated RTD sensor
Connections:	4 – terminals (connection via LEMO port)

3. Results and discussion

Due to the very sensitive nature of the OPVs towards the testing conditions, accurate characterization of those is a rather complex process. The diversity of the design and geometric dimensions of the devices (Fig. 1) were additional challenges for the experimenters. In spite of this, most of the groups succeeded in adjusting their measuring equipment to accommodate and measure all the samples. S2 in the Supporting Document presents the reported values of the photovoltaic parameters of all the measurements. Since Si solar cells provide sufficient stability for testing in round robin process, while OPVs may present degradation patterns, the two types of samples were analyzed individually and the results were compared between the two in order to establish the effect of instabilities of OPVs on the accuracy of the measurements.

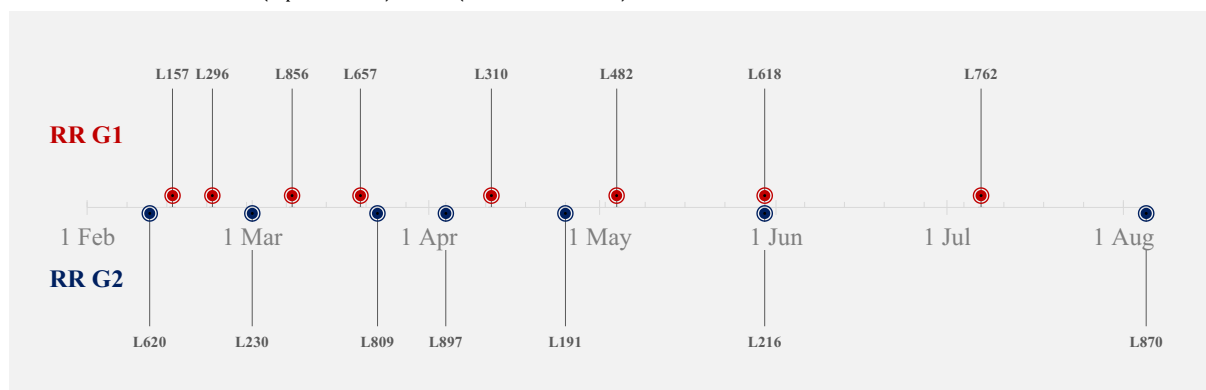
3.1. Reference Si solar cells

Each sample package contained one reference Si solar cell with WPVS (World Photovoltaic Scale) design [8] specially built for reference testing procedures. The rigidity and stability of Si minimized the chances of variations in the measurements caused by instabilities in the sample itself. Thus, any performance variations among the reported data could mainly be ascribed to variations in testing conditions and procedures. The performance of Si solar cells was initially certified by NREL and ISE Callab and after the RR experiments one of them was certified by PTB (short circuit current only), but the values were made available only after the RR cycles.

To establish the spread of Si measurements among the participants, the data of the solar cell parameters was plotted as deviations (in percent values) from the average of the certified values of NREL & ISE & PTB. Fig. 2 shows the deviations of short circuit current (I_{sc}), open circuit voltage (V_{oc}), fill factor (FF) and maximum power (P_{max}) for both Group 1 (G1) (left) and Group 2 (G2) (right). The values of the data are presented in S3 in the Supporting

Table 2

The timetable of the measurements of G1 (top red circles) and G2 (bottom blue circles).



Document. The red patterned bars highlight the measurements, which were performed with at least one of the following deviations from the protocol:

- Mismatch factor was not taken into account during I_{sc} measurement/reporting.
- Device temperatures above 25 °C (± 2 °C) were reported.
- 2 instead of 4-probe connection was used during measurements (except C samples, which could only be measured by 2-probe technique).

The table in Fig. 2 shows the agreement among the laboratories in the form of standard deviation values. The values are under 4% for G1 and under 7% for G2. The larger spread in G2 is explained by the fact that a larger number of group members deviated from the protocol. Fig. 3 shows the corresponding I – V curves for G1 (left)

and G2 (right). The arrows highlight the measurements with 2- and 4-probe methods. The use of the 2-probe method resulted in a higher serial resistance, which reduced FF, as can also be seen in Fig. 2 for the laboratories L482, L618 and L897.

The groups also measured the spectral response (SR) of the samples to calculate the mismatch factor. In order to compare the data among the laboratories each SR curve was normalized to its integral area within the range of 300–800 nm. According to the certified data of ISE Callab the difference of spectral response between the two Si samples was below 5% for the normalized SR values above 0.2%. Thus, it was chosen to combine the data for G1 and G2 in one plot. Fig. 4 (top) shows the normalized data. An easier way for revealing measurement errors however is the calculation of the deviations of the normalized data from the average of all curves, which is shown in Fig. 4 (bottom). The analysis of SR data in such a way allows for each laboratory to re-assess the measurement accuracy of their instrumentation and identify any errors in the SR measurements, which may result in wrong estimation of the mismatch factor. As an example, the measurements of L618 revealed a somewhat noisy signal and L620 had a slight miss-calibration of the wavelength step in the setup, both however being too minor for having any significant effect on the accuracy of the mismatch factor determination.

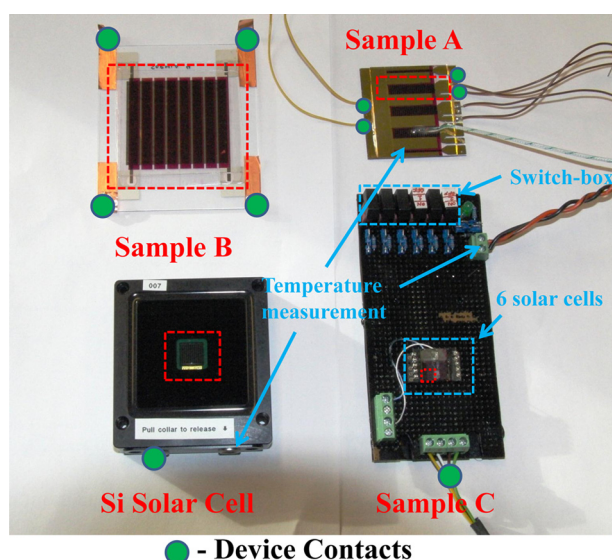


Fig. 1. Photo of the samples used in the RR cycle. The red dashed boxes highlight the solar cells in the samples. Sample A contains three individual solar cells, sample C contains 6 individual solar cells, sample B is a module of serially connected 8 solar cells and Si is a single monocrystalline Si solar cell. The green circles highlight the electric terminals of the samples. The blue arrows point to the temperature measuring units on the three samples and the switch box of sample C used for contacting each individual cell. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. OPV samples

3.2.1. Sample stability and I – V characterization

The stability of the samples is of great importance when round robin experiments are in question. Fig. 5 shows the stability curves of the PV parameters of each sample, plotted based on the reported data for G1 (left) and G2 (right) (the original data is provided in S2 in the Supporting Document). The measured points are connected with solid lines for guiding the eye. The data are normalized to the measurements of the first laboratory in the RR cycle. Some of the samples were re-tested by the laboratories of origin at the end of the RR cycle, but the measurements were performed at a very late stage. Thus, most of the samples except A types, significantly deteriorate due to exposure to environmental stresses. The results of the re-testing are not shown here. Fig. 5 shows also the curves of the Si samples for comparison. The following conclusions can be made from the stability curves:

- One of the C samples in each group failed at the beginning of the measurements.
- One of the A samples in G1 failed in the middle of the cycle.

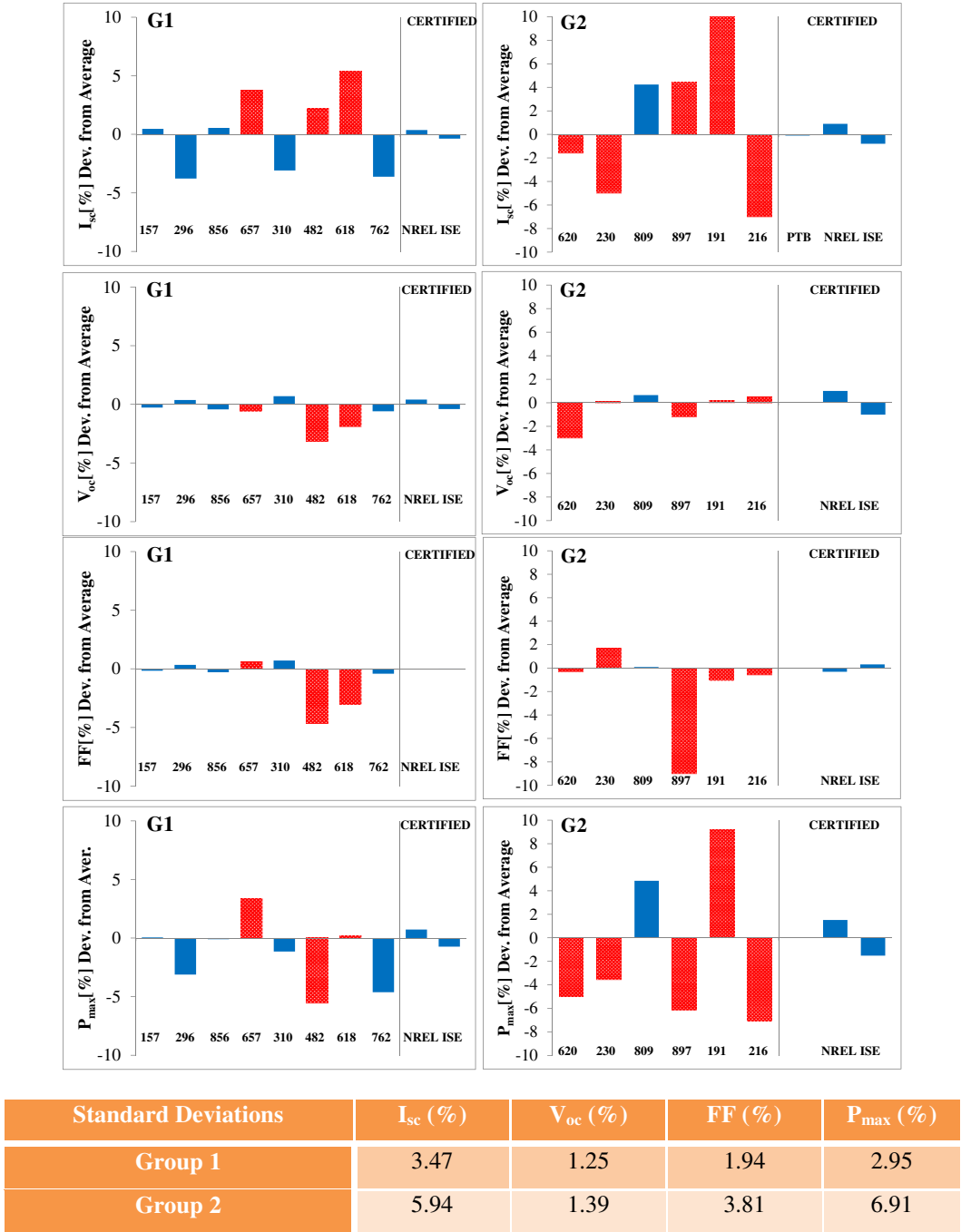


Fig. 2. Deviations of the reported PV parameters of Si samples for all the laboratories from the average of NREL & ISE & PTB for G1 (left) and G2 (right). The red patterned bars represent the measurements performed with some deviations from the protocol. The table shows the standard deviations of the reported values among the participants. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- c) Decay of V_{oc} (up to 10%) and FF (up to 20%) could be recognized for some of the samples during the course of the RR tests.
- d) Most of the samples experienced a significant drop in the performance while tested by the last participant in G1.

To understand the aforementioned issues the I – V curves were analyzed. I – V s of all the samples are provided in [S4 in the Supporting Document](#). The failure of the two C samples was associated with the sudden shorting of the devices possibly caused by stresses during the transportation. As an example [Fig. 6](#) shows the

I – V curves of one of the failed C samples. The failure of the A sample was associated with the burned region in the electrode grid identified during the post-RR analyses of the sample. Such burning was possibly caused by the application of excessive voltage to the device during the characterization by one of the participants. The decrease of V_{oc} and FF in the other samples was ascribed to the changes in the slope of the I – V curves at the negative bias, which most likely is caused by mechanisms creating shunt resistance inside the devices. To quantify this effect the slope of each I – V curves was defined in a region of ± 0.2 V and referred to as shunt resistance (R_{sh}). [Fig. 7](#) shows the R_{sh} change relative to the initial performance. The

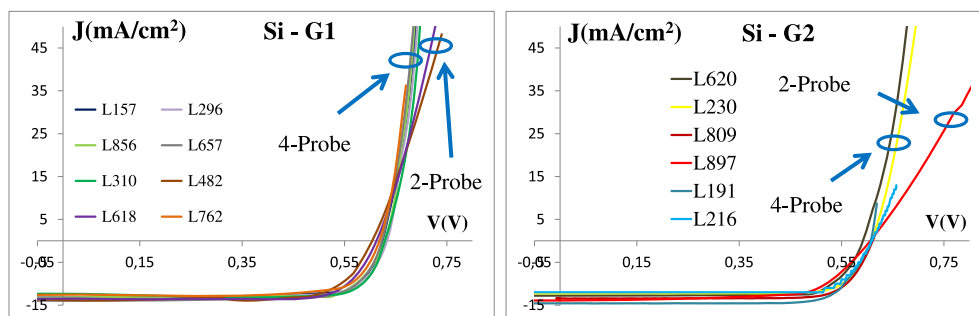


Fig. 3. I – V curves of Si samples measured in G1 (left) and G2 (right). The arrows are highlighting the I – V s measured with 2-probe and 4-probe methods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

values were extracted from the I – V curves. From the figure it is obvious that while experiencing large variations, R_{sh} shows a general decrease for most of the samples. As an example, Fig. 8 demonstrates the change of R_{sh} in the I – V s of sample A2 of G1 during the course of RR characterization. The possible reasons for such a change can be ascribed to the gradual electro-migration of the electrode material inside the devices, forced by the electric fields applied during I – V testing, which create shunt paths inside the device [14]. I – V s of the samples in G1 also revealed a significant drop in the performance of most of the samples during the testing by the last participant in the RR cycle. Such a synchronic drop in many samples suggests that the entire set of samples experienced possible stresses during transportation to or extended storage by the last laboratory prior to characterization.

Apart from the aforementioned issues, the short circuit currents of the OPV samples did not show any obvious decay patterns when compared to Si samples, as will also be shown later in the manuscript. It is also worth mentioning that B modules with flexible

encapsulation performed almost as well in terms of stability, as the glass encapsulated A and C samples.

I – V curve analyses additionally helped revealing a number of issues related to the measurement procedures. Similar to Si samples, the difference between 2- and 4-probe measurements resulted in variation of serial resistance among the OPV samples, but to a lesser degree, possibly due to different geometrical structures of the latter. Additionally, one of the participants used a measuring setup with low sensitivity in the photocurrent range of the samples. This resulted in generation of I – V curves with “stair-step” form and thus, large error in the recorded values of PV parameters. Anomalous kinks in the I – V curves were identified for another participant, which resulted in wrong determination of the PV parameters. Those were associated with possible weak contacting between the measuring setup and the sample.

Additionally, a large diversity in the voltage sweep range was recorded among the I – V s especially pronounced in the measurements of G2. This was a result of the fact that the instructions for the voltage sweep range were provided not in the protocol, but separately by the device manufacturers. This has led to confusion and some participants performed the I – V testing according to their own discretion. As a result, in some cases the sweep range was kept to a minimum (between 0 and V_{oc}). Such a short sweep range reduced the time of light exposure of the sample. As a consequence, the C samples, which required initial photo-doping for optimal performance, received insufficient doping and thus, showed a reduced performance.

3.2.2. Data deviations

S2 in the Supporting Document provides the original reported values of sample temperature, mismatch factor, V_{oc} , I_{sc} , FF and P_{max} for all the samples measured by all the participants. The following approach was used to estimate the variations of these parameters. As a first step, any data of the failed or severely degraded samples was removed from the general plot. This was followed by calculation of the average values among all the participants for each device and normalization of the data to that average. The data was then readjusted to a new average calculated using only the data that deviated from the original average by less than 10%. Fig. 9 shows the results of these calculations for G1 (left) and G2 (right) (the calculated values are presented in S5 in the Supporting Document). The values for each device in the plots are positioned in the order of the RR cycle. Similar to Si solar cells (see Section 3.1) the red patterned bars highlight the measurements with deviations from the protocol. From Fig. 9 it is obvious that rather large number of participants had one or more of the earlier described deviations (Section 3.1), which is mostly ascribed to the interplay between the sophisticated designs of the samples, non-rigid protocol and lack of appropriate testing instrumentation.

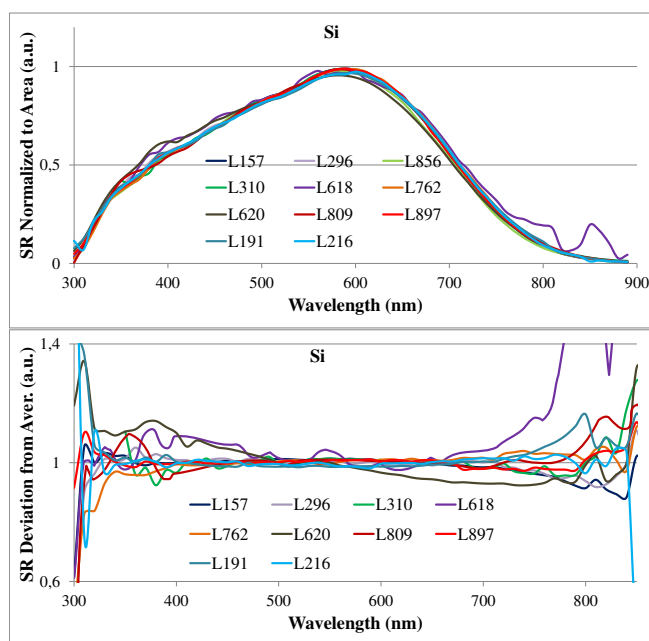


Fig. 4. The SR data of Si samples normalized to the integral area of the curves in the range of 300–800 nm (top) and the deviations of the normalized curves from the average of all the curves as a function of wavelength (bottom). The plots include both G1 and G2 data. For L856 the calculations of the deviations in the bottom plot were not performed due to the difference in the wavelength step format.

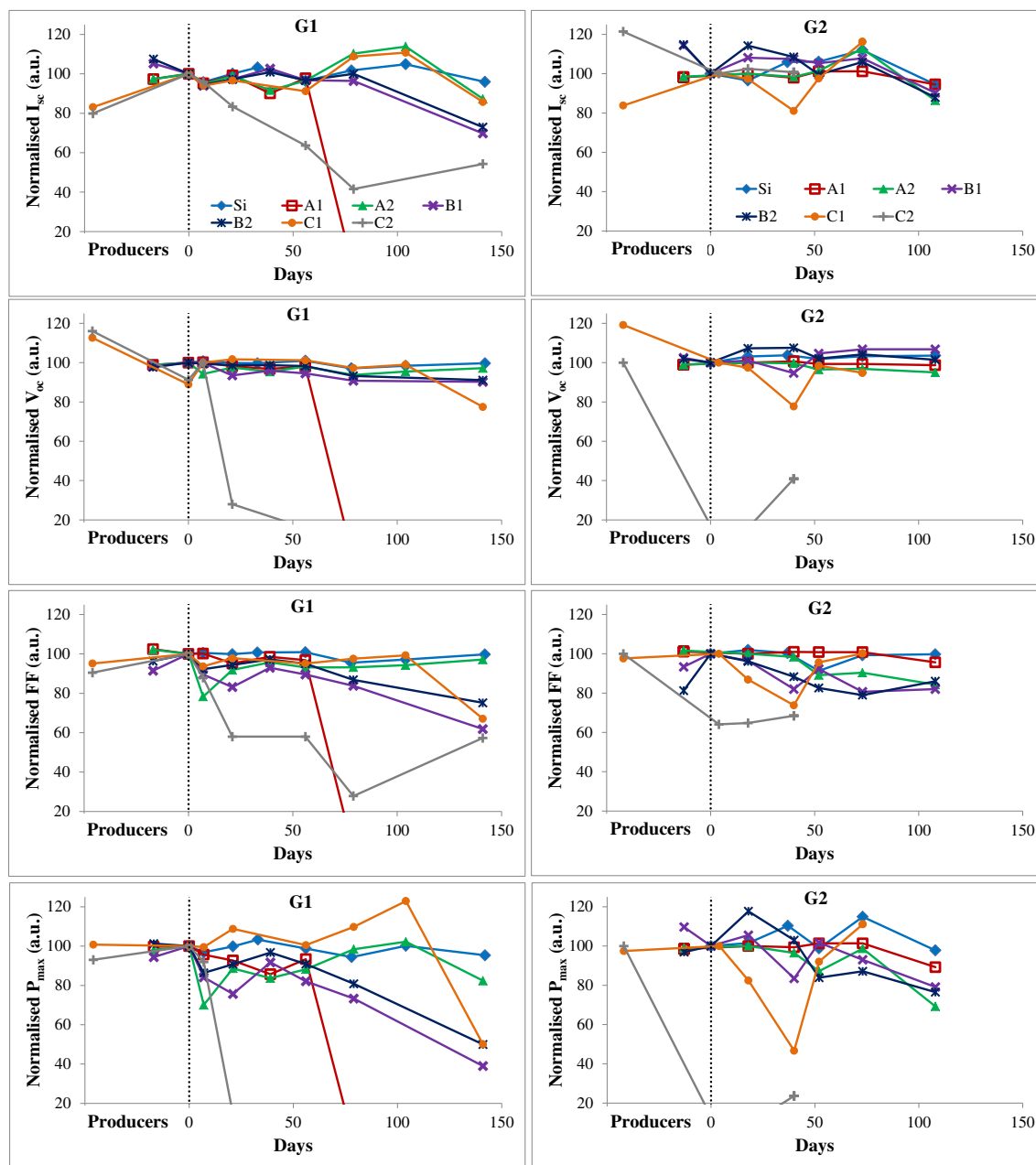


Fig. 5. Stability curves of the PV parameters for the samples in G1 (left) and G2 (right). The vertical stripe separates the measurements of the producers from the RR cycle measurements. The data are normalized to the first measurements in the RR cycle.

One of the particular issues was that in most cases the temperature of the samples was not controlled due to either the lack of appropriate equipment or the sophisticated design of the sample. The common technique was to perform a fast I – V scan before the sample heats up. This however was not sufficient due to the instant heating of the device during light absorption and in many cases temperatures higher than 25 °C were reported.

The earlier discussed slight decay of V_{oc} and FF can be recognized in the plots of the deviations in Fig. 9. The plots of the same deviations, grouped versus the laboratories, are provided in S6 in the Supporting Document.

Fig. 10 shows the standard deviations calculated for each device. The values for Si samples are shown as well for comparison. Table 3 shows the corresponding values for both groups. Despite of all the deviations, there was an agreement on the performance of the

reference Si samples with variations in the range of 3–7% and for some of the OPVs below 5%. However, for most of the OPVs the range was between 5 and 15%. From all parameters, I_{sc} stands out as the most significant source of disagreement for the majority of the measured samples, followed by discrepancy in FF values. V_{oc} measurements seem to be much more reproducible, except for C samples, which could possibly be due to contacting issues.

To study the deviations of I_{sc} the following method was used. A set of X – Y columns were created based on pairs of I_{sc} values (normalized to the average) produced from the data of measurements of all devices (excluding the data of failed or degraded samples) by one laboratory. The data were then graphed on an X – Y scatter plot. The same procedure was performed for each laboratory and the groupings were assembled in one plot (this technique was earlier used at PEP'93 intercomparison studies [6]). Fig. 11(A,B)

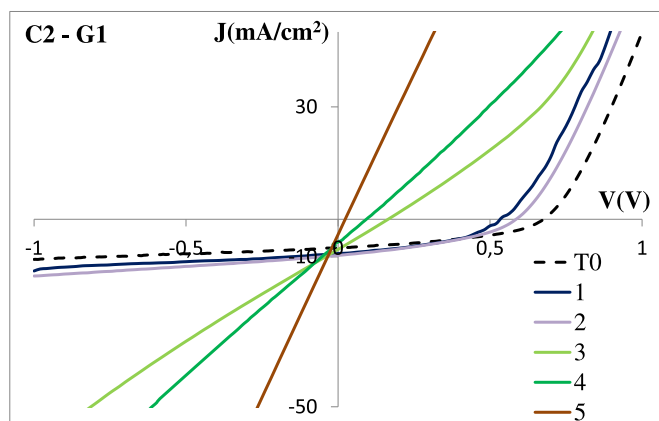


Fig. 6. I – V curves of C2 sample measured by G1. The numbering of the curves represents the course of RR characterization.

shows the two plots for G1 (left) and G2 (right). The value 100 in the plot corresponds to the average defined by the data within 10% deviations from the original average, as was described at the beginning of this section. The size of each grouping is proportional to the laboratory's standard deviations, while the position of the grouping along the diagonal line defines the laboratory ranking (deviation from the average) [6]. A number of conclusions can be made from Fig. 11(A,B):

- The spread of data in G1 was less than in G2, which can be explained by the fact that a larger number of groups deviated from the protocol in G2.
- Laboratories L482 and L618 in G1 showed significantly larger spread and somewhat elevated ranking. This may be related

to the following deviations from the protocol that were identified for these laboratories:

L482 – mismatch factor not taken into account, I – V testing with 2-probe method.

L618 – I – V testing with 2-probe method, sample temperatures above 25 °C.

- A large spread was also recorded for the laboratories L191 and L216 in G2, with the ranking of the first being significantly above and the latter below the average. The following issues were identified for these laboratories:

L191 – I – V testing with 2-probe method, sample temperatures above 25 °C.

L216 – I – V testing with 2-probe method, low sensitivity of source measuring unit, flashing light used for illumination.

If all the data, which were measured with deviations from the protocol, are removed from the plots, then the spread of I_{sc} is significantly reduced, as can be seen in Fig. 11(C,D).

Additionally, the four laboratories with large spreads were among the last in each group to perform the RR test. Thus, possible degradation of samples could be a partial reason for large deviations, which however is highly unlikely, as no clear patterns of I_{sc} degradation were identified in the stability data.

Additionally, the deviations of reference Si samples and the deviations of the average of all the OPVs were compared for each laboratory, as shown in Fig. 12 (the deviations are presented with line-charts for guiding the eye). The deviations of OPV measurements for each laboratory can be found in S6 in the Supporting Document. The comparison showed almost the same dynamics of deviations between Si and OPVs for many of the laboratories including the last two measurements in both G1 and G2, possibly

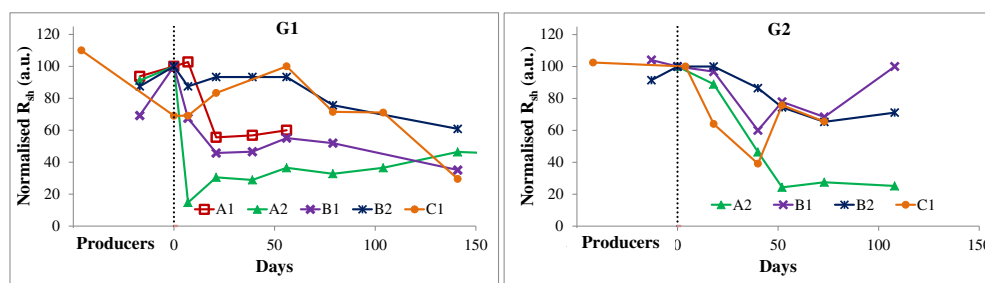


Fig. 7. R_{sh} shunt resistance of the samples (defined in Section 3.2.1) Vs time for G1 (left) and G2 (right). The R_{sh} values are extracted from the I – V curves and normalized to the values of the first measurements in the RR cycle.

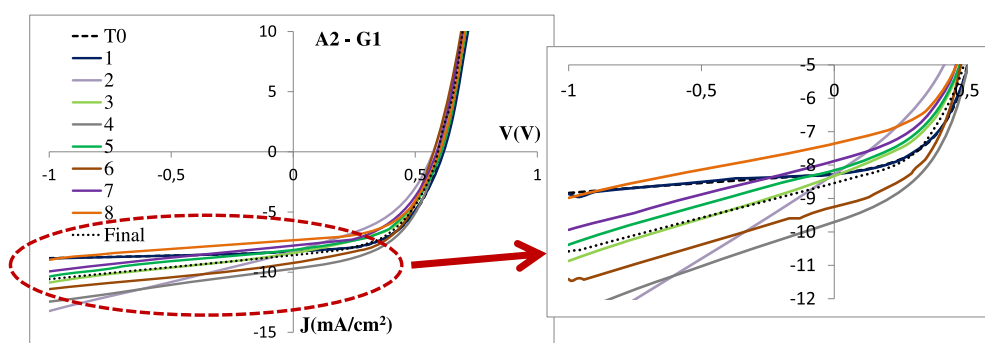


Fig. 8. I – V curves of A2 measured by G1 (left) and the enlarge section of I – V s of A2 with the demonstration of shunt resistance change (right). The numbering of the curves represents the course of RR characterization.

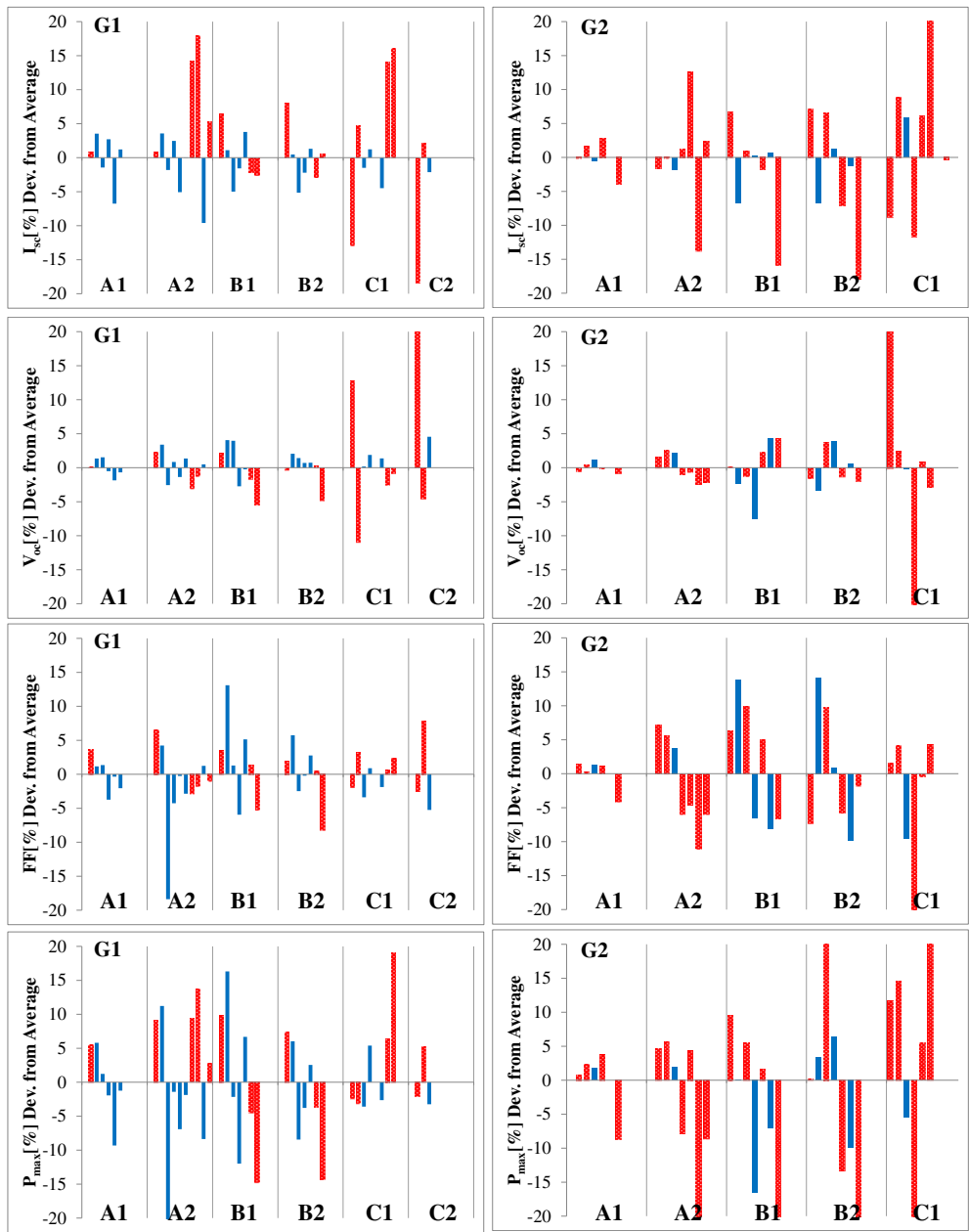


Fig. 9. Deviations of the measurements from the average for all the functioning OPV devices in G1 (left) and G2 (right). The average was calculated from the data which deviated less than 10% from the original average. The bars marked with red correspond to the measurements performed with certain deviations from the protocol. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

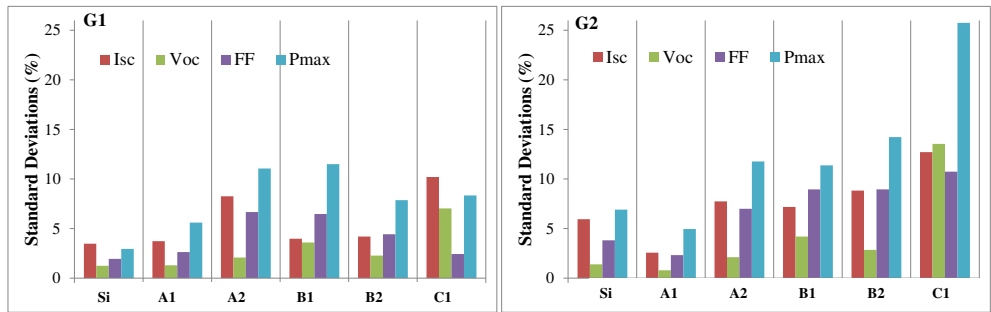


Fig. 10. The standard deviations of the PV parameters for all the functioning devices.

Table 3

The values of the standard deviations of PV parameters for G1 and G2.

Device	I_{sc} (%)	V_{oc} (%)	FF (%)	P_{max} (%)
Si – G1	3.5	1.2	1.9	3.0
A1 – G1	3.7	1.3	2.6	5.6
A2 – G1	8.7	2.2	7.1	11.7
B1 – G1	4.0	3.6	6.5	11.5
B2 – G1	4.2	2.3	4.4	7.9
C1 – G1	10.2	7.0	2.4	8.3
Si – G2	5.9	1.4	3.8	6.9
A1 – G2	2.6	0.8	2.3	5.0
A2 – G2	7.7	2.1	7.0	11.8
B1 – G2	7.2	4.2	9.0	11.4
B2 – G2	8.8	2.8	9.0	14.2
C1 – G2	12.7	13.5	10.7	25.7

confirming that the degradation of OPVs is not the reason for the large deviation of the last measurements.

The spectral response measurements of the OPV samples were analyzed in a similar fashion as for Si samples (see Section 3.1). S7 in the Supporting Document shows the corresponding data. The SR data was used to identify any possible errors and re-assess the accuracy of the instrumentation. The groups additionally reported the spectra of the light sources used for I – V testing. The data are provided in S7 in the Supporting Document. The analyses revealed deviations among the spectra especially pronounced in the UV region. However, no clear trend could be identified between the reported sample performance and the UV content of the light sources.

3.3. Pitfalls

The analyses of the reported data allowed identifying a number of issues and pitfalls that affected the accurate characterization of the devices.

3.3.1. Issues related to sample design

1. The change of the slope of I – V s at negative bias (possibly due to generated shunts) in most of the samples during the RR cycles resulted in slight degradation of V_{oc} and FF.
2. One of the participants reported difficulties with illuminating the entire large active area (15 cm^2) of the B modules during the characterization.
3. The sophisticated design of the C samples did not allow the temperature control during testing. Very small active areas of these devices ($\sim 0.04\text{ cm}^2$) resulted in complications during quantum efficiency measurements; the sophisticated electrical circuitry integrated in the device made the contacting additionally complex. Two out of four C samples experienced catastrophic failure at the start of the experiments possibly due to contact failure.

3.3.2. Issues related to testing protocol

1. The protocol did not contain instructions related to I – V sweep range and step size, nor did it have information related to light

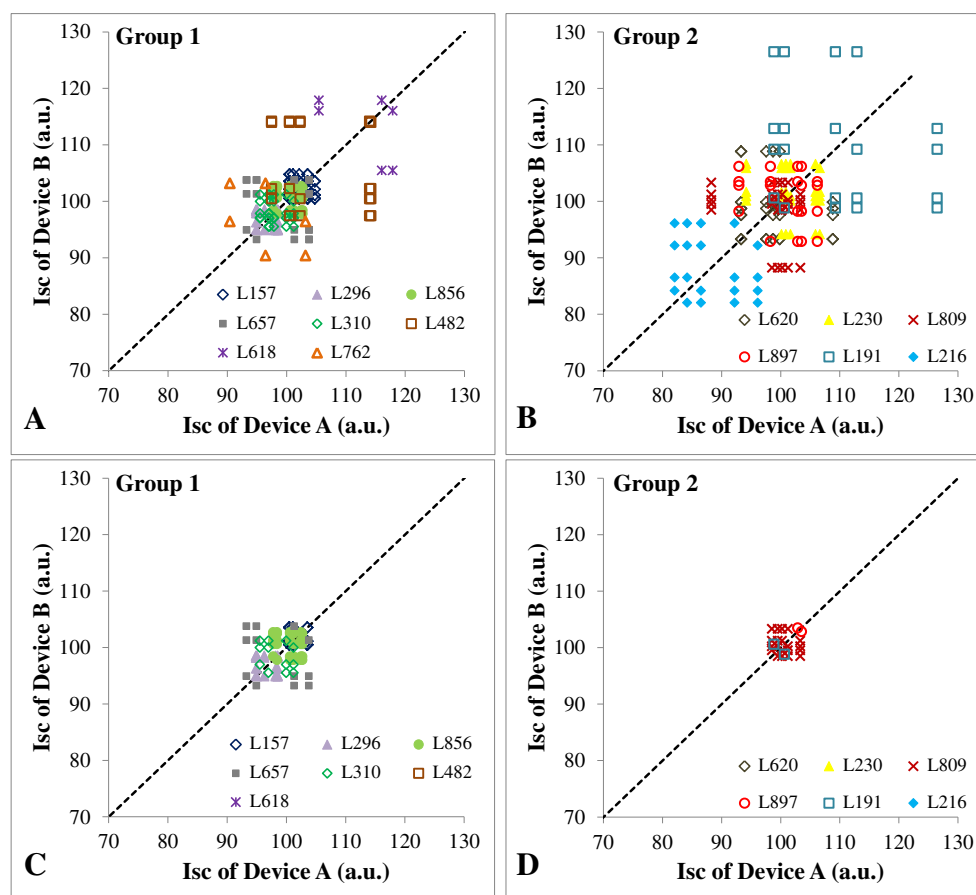


Fig. 11. (A) and (B) show the scatter plots of I_{sc} for all the laboratories in G1 and G2, respectively. The data was plotted using X–Y pairs of I_{sc} values from all the measurements of each laboratory. The I_{sc} values normalized to the average were used. (C) and (D) show the same data after removing the measurements that were performed with deviations from the protocol by G1 and G2, respectively.

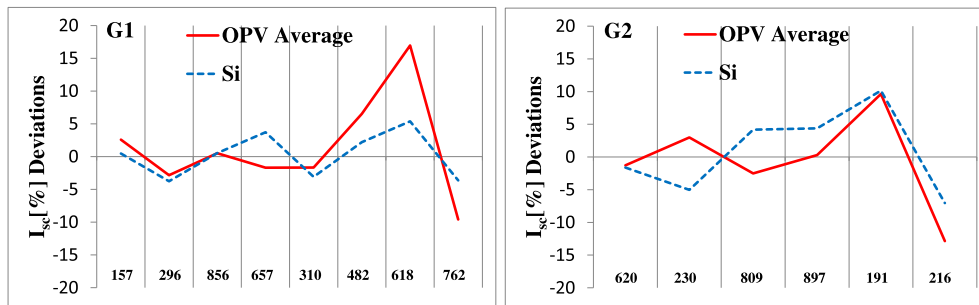


Fig. 12. Comparison of deviations of Si samples and average of OPV devices for each laboratory in G1 (left) and G2 (right). The deviations are presented via line-charts for guiding the eye.

soaking of the samples prior to characterization. The experimenters were referred instead to the instructions given by the device manufacturers. This has led to certain confusion, resulting in variations in the characterization procedures among the participants.

2. 2-probe measuring technique was initially recommended for all the samples in the protocol. However, the participants were later recommended to choose the method that suited the best for each sample design. This has eventually led to variation in the choice of 2- or 4-probe methods among the measuring techniques used by the participants.
3. While minimum of 3–4 I – V scans were recommended, no instructions were provided on whether the average or the best data should be reported.
4. A quick I – V scanning was recommended to avoid heating of the samples. However, this technique proved not to be useful based on two reasons:
 - a) The variation in the speed of I – V characterization interfered with the photo-doping process of the samples.
 - b) The samples studied in this work had a design that without effective cooling of the sample would result in an instant temperature increase upon light absorption and thus, keeping the temperature close to 25 °C would require unrealistic speeds for I – V scanning.

3.3.3. Issues related to measurements

1. Delays from the RR cycle plan were recorded in some circumstances.
2. According to the protocol the samples had to be returned to the laboratories of origin for the final re-testing. However, the transportation back to the laboratories of origin and the re-testing of the samples by the manufacturers were delayed resulting in significant decay of the samples. Thus, no valid comparison of initial and final testing could be made for most of the samples.
3. Anomalies in the reported I – V curves were recorded, which were generated by either low sensitivity of the measuring unit or possible weak contacting between the measuring setup and the sample.
4. Despite the required minimum of 3–4 I – V curves by the protocol, only one I – V scan per device was performed in some cases.
5. Mismatch factor was not calculated by a number of participants.
6. Some of the spectral response data had large deviations from the average (shown in S7 in the Supporting Document).

It is believed that if all the issues identified in this work will be addressed in future RR studies, the agreement among the measurements will significantly be improved.

4. Conclusions and recommendations

Round robin characterization of OPV devices was performed among 16 laboratories within SOPHIA and EERA projects. Three types of OPV samples were characterized, two with glass encapsulation and one with flexible barrier material. Fraunhofer ISE, NREL and PTB certified Si solar cells were used as a reference. Analyses of the measurements were performed and the standard deviations among the participants were established. While the standard deviations among the measurements of the reference Si devices were within 3–7%, variation between 3 and 15% were recorded for OPVs. The RR studies helped identifying a number of critical issues related to sample designs, testing protocol and procedures that affected the agreement between the measurements. It is believed that addressing all these issues will result in more harmonized measurements among different parties.

Based on the analyses and after-study discussion the following recommendations are suggested:

1. RR characterization of OPV technologies at least once every 2 years is recommended.
2. A special sample design must be considered, that will meet the needs for round robin testing. The manufacturers and the testing laboratories must agree on a design that is simple enough to assure that every participating laboratory can characterize it according to the protocol and rigid enough to provide stable reproducible performance. In particular:
 - a) Effective thermal contact is important to allow control of the sample temperature during light exposure.
 - b) Sample active area dimensions must be within 1 cm² and 10 cm² range, with the maximum length not exceeding 10 cm.
 - c) Sample height must be ≤ 2 cm.
 - d) Wires soldered to sample terminals must be used to avoid any contact failures.
3. The samples must be tested for stability. If initial instabilities are identified, the sample must be stabilized before being subjected to RR tests.
4. The quantity of the tested samples of each type must be sufficient to provide good data statistics and eliminate the risk of catastrophic failures. Minimum 5 samples of each type are recommended.
5. A reference sample must be provided, that will only travel with the other samples, but will not be characterized. This will help distinguishing between the stresses of transportation and characterization.
6. The RR cycle schedule must be planned and agreed with each participant to minimize any possible delays. The participants must strictly follow the RR time schedule.

7. In addition to the steps for I – V characterization presented in these studies the protocol must also state:
 - a) I – V scanning step and range provided for each sample.
 - b) Information related to sample light soaking prior to I – V test.
 - c) Information related to the methodology for sample contacting.
 - d) Instructions for taking into account the height difference between the reference device and the tested specimen.
8. The participants must get acquainted with the testing protocol and report possible deviations and issues to the coordinators before the start of the RR test. During the RR test the participants must strictly follow the protocol.
9. Retesting of the samples by the manufacturer or alternatively by the first laboratory of the round robin after the cycle must be performed.
10. 4-probe testing method must always be used, if the device design allows it.
11. Spectral mismatch correction must be performed, since large deviations of spectral irradiance strongly affect measurement results.

Acknowledgments

European Research Infrastructure (SOPHIA) and European Energy Research Alliance (EERA) are acknowledged for the support; This work has been supported by EUDP (j.no. 64011-0002); Komlan Anika is acknowledged for performing the measurements at ESTI; This work has been supported by the UK Department for Business, Innovation and Skills; Antonio Romano and Aniello Borriello are acknowledged for technical support at ENEA.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.renene.2013.09.034>.

References

- [1] <http://www.sophia-ri.eu>.
- [2] <http://www.eera-set.eu/>.
- [3] Ossenbrink H, Krebs K, Steenwinckel RV. Results of the 1984/1985 international round robin calibration summit working group on technology, growth and employment. IEEE; 1985. p. 943–4.
- [4] Metzendorf J, Wittchen T, Heidler K, Dehne K, Shimokawa R, Nagamine F, et al. Objectives and results of the PEP'87 round-robin calibration of reference cells and modules. In: Proceedings of the 21st IEEE photovoltaic specialists conference. Orlando, FL: IEEE; 1990. p. 952–9.
- [5] Metzendorf J, Wittchen T, Heidler K, Dehne K, Shimokawa R, Nagamine F, et al. The results of the PEP'87 round-robin calibration of reference cells and modules. ISBN 3-89429-067-6; November 1990. Final Report, PTB Technical Report PTB-Opt-31, Braunschweig, Germany.
- [6] Osterwald CR, Anevsky S, Barua AK, Dubard J, Emery K, King D, et al. Results of the PEP'93 intercomparison of reference cell calibrations and newer technology performance measurements. In: Proceedings of the 25th IEEE photovoltaic specialists conference. Washington, DC: IEEE; 1996. p. 1263–6.
- [7] Osterwald CR. Results of 1992 ASTM cell and module measurement intercomparison, vol. 1. IEEE; 1993. p. 1102–6.
- [8] Osterwald CR, Anevsky S, Bücher K, Barua AK, Chaudhuri P, Dubard J, et al. The world photovoltaic scale: an international reference cell calibration program. Prog Photovolt Res Appl 1999;7:287–97.
- [9] Betts TR, Gottschalg R, Infield DG, Kolodenny W, Prorok M, Zdanowicz T, et al. Round robin comparison of European outdoor measurement systems. In: Proceedings of the 21st European photovoltaic solar energy conference, 4–8 September 2006. p. 2447–51.
- [10] Herrmann W, Mau S, Fabero F, Betts T, van der Borg N, Kiefer K, et al. Advanced intercomparison testing of PV modules in European test laboratories. In: Proceedings of the 22nd European photovoltaic solar energy conference, 3–7 September 2007. p. 2506–10.
- [11] Krebs FC, Gevorgyan SA, Gholamkhass B, Holdcroft S, Schlenker C, Thompson ME, et al. A round robin study of flexible large-area roll-to-roll processed polymer solar cell modules. Sol Energy Mater Sol Cells 2009;93:1968–77.
- [12] Gevorgyan SA, Medford AJ, Bundgaard E, Sapkota SB, Schleiermacher HF, Zimmermann B, et al. An inter-laboratory stability study of roll-to-roll coated flexible polymer solar modules. Sol Energy Mater Sol Cells 2011;95:1398–416.
- [13] Søndergaard RR, Makris T, Lianos P, Manor A, Katz EA, Gong W, et al. The use of polyurethane as encapsulating method for polymer solar cells – an inter laboratory study on outdoor stability in 8 countries. Sol Energy Mater Sol Cells 2012;99:292–300.
- [14] Rösch R, Tanenbaum DM, Jørgensen M, Seeland M, Bärenklau M, Hermenau M, et al. Investigation of the degradation mechanisms of a variety of organic photovoltaic devices by combination of imaging techniques—the ISOS-3 inter-laboratory collaboration. Energy Environ Sci 2012;5:6521–40.
- [15] Andreasen B, Tanenbaum DM, Hermenau M, Voroshazi E, Lloyd MT, Galagan Y, et al. TOF-SIMS investigation of degradation pathways occurring in a variety of organic photovoltaic devices – the ISOS-3 inter-laboratory collaboration. Phys Chem Chem Phys 2012;14:11780–99.
- [16] Tanenbaum DM, Hermenau M, Voroshazi E, Lloyd MT, Galagan Y, Zimmermann B, et al. The ISOS-3 inter-laboratory collaboration focused on the stability of a variety of organic photovoltaic devices. RSC Adv 2012;2:882–93.
- [17] Teran-Escobar G, Tanenbaum DM, Voroshazi E, Hermenau M, Norrman K, Lloyd MT, et al. On the stability of a variety of organic photovoltaic devices by IPCE and *in situ* IPCE analyses – the ISOS-3 inter-laboratory collaboration. Phys Chem Chem Phys 2012;14:11824–45.
- [18] Larsen-Olsen TT, Machui F, Lechene B, Berny S, Angmo D, Søndergaard R, et al. Round-robin studies as a method for testing and validating high-efficiency ITO-free polymer solar cells based on roll-to-roll-coated highly conductive and transparent flexible substrates. Adv Energy Mater 2012;2:1091–4.
- [19] Angmo D, Gonzalez-Valls I, Veenstra S, Verhees W, Sapkota S, Schieffer S, et al. Low-cost upscaling compatibility of five different ITO-free architectures for polymer solar cells. J Appl Polymer Sci 2013;130:944–54.
- [20] Lilliedal MR, Medford AJ, Madsen MV, Norrman K, Krebs FC. The effect of post-processing treatments on inflection points in current–voltage curves of roll-to-roll processed polymer photovoltaics. Sol Energy Mater Sol Cells 2010;94:2018–31.
- [21] Reese MO, Nardes AM, Rupert BL, Larsen RE, Olson DC, Lloyd MT, et al. Photoinduced degradation of polymer and polymer–fullerene active layers: experiment and theory. Adv Funct Mater 2010;20:3476–83.