

8-346-15 Yoshi Ohno

110 labs in world's largest interlaboratory comparison of LED test labs – improving testing competency to support market transformation

Yoshi Ohno

National Institute of Standards and Technology
100 Bureau Drive, Stop 2140,
Gaithersburg, MD 20899-2000
USA
Email: ohno@nist.gov

Koichi Nara

Research and Innovation Promotion Headquarters
National Institute of Advanced Industrial Science and Technology
Central 2, 1-1-1-, Umezono, Tsukuba, Ibaraki, 305-8568,
Japan
Email: narakoichi@gmail.com

Peter Bennich

The Swedish Energy Agency
Box 380 69
S-100 64 Stockholm
Sweden
Email: peter.bennich@energimyndigheten.se

Michael Scholand

IEA 4E SSL Annex
c/o N14 Energy Limited
42 Conway Road
London, N14 7BE
United Kingdom
Email: mscholand@n14energy.com

Nils Borg

IEA 4E SSL Annex
c/o Borg & Co AB
Sveavagen 98, 4 tr
113 50 Stockholm
Sweden
Email: nils@borgco.se

Abstract

Light Emitting Diode (LED) technology has been gaining significant market share in general illumination applications around the world, but reliable measurement of the performance of LED lamps and luminaires has been challenging. Furthermore, the lack of a global accreditation scheme has driven up costs of getting LED products tested. The IEA 4E SSL Annex therefore initiated a programme starting in 2011 looking to support a global laboratory accreditation scheme. The problems were the lack of an international test method and the difficulty in proficiency testing. To assist accreditation bodies to establish testing accreditation programs for LED lighting products, the 2013 Interlaboratory Comparison (IC 2013) was developed and conducted. IC 2013 was designed to serve as proficiency testing by complying with ISO/IEC 17043, covering several test methods used in different regions, thus serving many regions of the world. IC 2013 brought together and compared test results of 110 laboratories around the world on their ability to measure photometric, colorimetric, and electrical

quantities. The sample of products took into account the most important products in the different regional markets, but all test samples included integrated LED lamps and LED luminaires. While most laboratories were within the expected levels of agreement, a few extreme outliers were observed which suggests problems meeting the test method requirements. Identification of these large deviations for a few laboratories demonstrates the importance of proficiency testing. Overall, this large-scale interlaboratory comparison clarified the level of agreement in measurement results for these products, and will help governments and manufacturers around the world to learn the uncertainties of the measurements of LED lighting products and ensure that new LED products sold to consumers and companies are of high quality and meet the claimed performance.

1. Introduction

Performance testing of regulated or labelled products is the most fundamental requirement of any regulatory or endorsement programme. Reliable and repeatable testing of products is the cornerstone on which policy makers are able to construct market transformational initiatives to push and pull the market towards more environmentally responsible products. When new technologies are introduced to a market, there can be periods of uncertainty and issues arising around measurement, putting consumers and markets at risk. Reliable measurement of the performance of LED lamps and luminaires has been challenging. For instance, LED light sources behave differently when tested compared to conventional light sources and many laboratories have experienced difficulties in calibrating their equipment and establishing procedures that generate reliable and consistent test results.

The Solid-State Lighting (SSL) Annex was established in 2010 under the framework of the International Energy Agency's (IEA) Energy Efficient End-Use Equipment (4E) Implementing Agreement. The IEA 4E SSL Annex (simply "SSL Annex" hereafter) works to assist governments of member countries in promoting SSL as an effective means to reduce energy consumption worldwide.¹

Starting in 2011, the SSL Annex launched an initiative that sought to address the lack of a global laboratory performance assessment scheme. This initiative was designed to help support harmonisation of SSL testing around the world by developing an approach to compare and assess the measurement capabilities of lighting test laboratories, and to support accreditation programmes for laboratories testing LED lighting products.

The problems were the lack of an international test method and the difficulty in proficiency testing required for testing accreditation programs for LED lighting products. To assist accreditation bodies to establish such testing accreditation programs, the SSL Annex developed and conducted an interlaboratory comparison (IC) program, called "IC 2013", which was designed to serve as proficiency testing by complying with ISO/IEC 17043 [2], covering several test methods used in different regions, thus to serve many regions of the world. IC 2013 was conducted between October 2012 and August 2013. It involved 54 laboratories from 18 countries directly in this study, comparing measurements of photometric, colorimetric, and electrical quantities of several different types of SSL products. In addition, measurement data from the proficiency testing of 35 laboratories in the National Voluntary Laboratory Accreditation Program (NVLAP) Energy Efficient Lighting Products (EELP) for SSL and in the (US) National Institute of Standards and Technology's (NIST) Measurement Assurance Program (MAP) were linked to IC 2013. Also, data from an additional 21 laboratories from the Asia Pacific Laboratory Accreditation Cooperation (APLAC) proficiency test T088, were linked to IC 2013, bringing the total number of laboratories to 110 involving 123 sets of data, making this the world's largest interlaboratory comparison of LED test labs.

As SSL technology continues to advance at a rapid pace, industry researchers are making improvements in efficacy, light quality, and operating life. In addition, manufacturing improvements and market competition are putting downward pressure on prices, making SSL products less expensive relative to fluorescent and other conventional light sources. These two trends will benefit consumers and result in SSL penetrating many more lighting applications across all end-use sectors. That said, one factor that could impact the speed of penetration is the quality of the SSL products placed on the market. The IC 2013 sought to conduct an international testing

¹ The SSL Annex works internationally to support efforts at a national and regional level by addressing the main challenges with SSL technologies. The Annex member countries believe that there are significant advantages in engaging in an international collaboration in order to develop a consensus on harmonised approaches to SSL performance and quality (Bennich, Ref. [1]). Sponsoring governments of the SSL Annex during the period when IC 2013 was conducted included Australia, Denmark, France, Japan, Republic of Korea, the Netherlands, Sweden, the United Kingdom and the United States. China also participates as an expert member of the SSL Annex. The work of the SSL Annex spans a wide range of initiatives which can be found on the Annex's website (<http://ssl.iea-4e.org/>), including guidance for policy makers, quality and performance tiers and support for laboratory accreditation.

scheme to help ensure that performance testing across the range of lighting products will provide accurate and comparable test results from any of the different laboratories participating in the comparison.

2. Lack of an international test standard

When the SSL Annex was started in 2010, the founding members noted that there was a lack of agreement on test standards for LED products which contributed to a lack of international accreditation schemes. This lack of an international test standard for LED products created a barrier to trade that increases the overhead costs of doing business in different countries. Ultimately, these factors slowed down the market penetration of LEDs and resulted in lost energy savings. Figure 1 depicts the situation that existed in 2011 when the IC 2013 was launched. The figure shows similar systems of quality assurance for LED lighting products using testing accreditation that have been or are being established in four key regions (or countries). In each of the regions shown, if accreditation is introduced, proficiency testing (PT) is required to assess the laboratory's competence for measuring the performance of LED lamps. However, due to different test methods used in the different regions, different PTs are used and no region will accept (i.e., recognise) the test results from other regions.

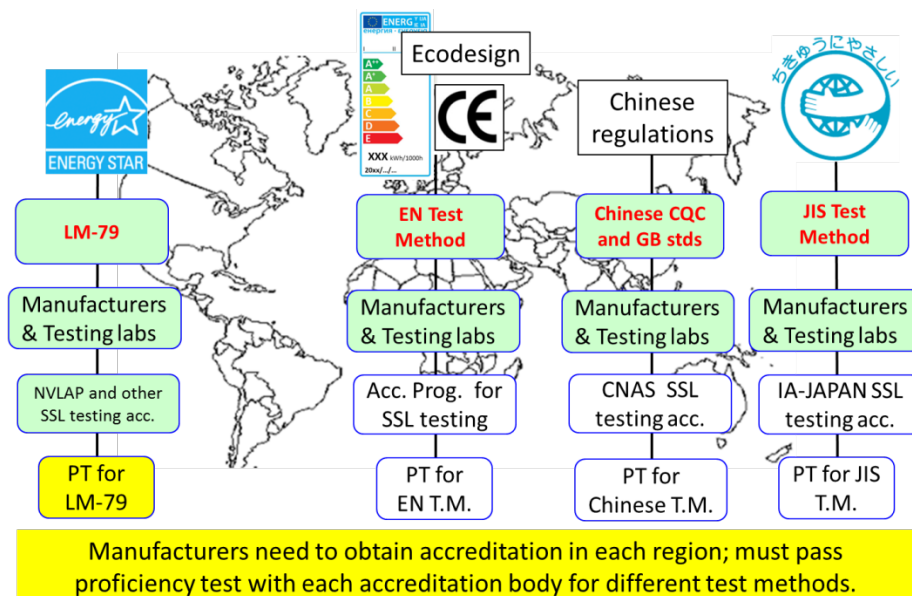


Figure 1. Problems associated with poor harmonisation in SSL testing, no mutual recognition of test results

The SSL Annex sought to address the situation depicted in Figure 1. With IC 2013, the Annex established a consistent foundation on which LED lighting products can be tested and compared, to provide an opportunity for mutual recognition of accreditation programmes across the various regional markets. Figure 2 represents an illustration of the ideal scheme that the SSL Annex is working to achieve.

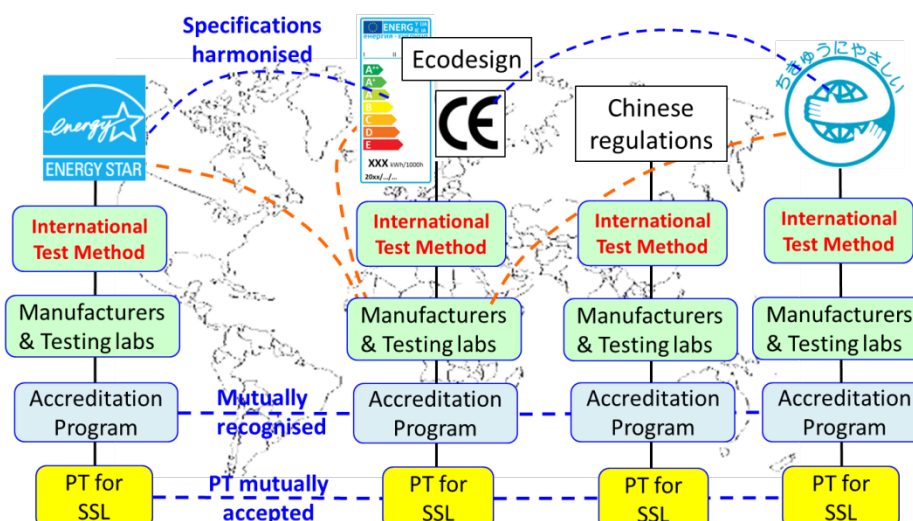


Figure 2. The ideal scheme - harmonisation of SSL testing; mutual recognition of test results

3. Interlaboratory comparison testing

The IC 2013 was led jointly by Dr. Yoshi Ohno of the National Institute of Standards and Technology (NIST) and Dr. Koichi Nara of International Accreditation Japan (IAJapan), National Institute of Technology and Evaluation (NITE). Dr. Ohno (Leader of Task 2, SSL Testing, in SSL Annex) led the work to develop the scheme and technical specifications of IC 2013 using one common test method that encompassed those used in the four regions. Dr. Nara (Leader of Task 3, Accreditation, in SSL Annex) led the work to enable the IC 2013 to be accepted as PT by accreditation programmes worldwide, by ensuring compliance to international standards on proficiency testing for accreditation.

3.1 Harmonise SSL performance testing

In an ideal world, harmonised test methods would be applied for performance testing. This would generate many benefits, including:

- Shared (and thus lower) development costs for preparing a test methods, especially for emerging products such as SSL;
- Comparative test results for products sold domestically and in neighbouring economies, enabling different countries to co-operate on market surveillance;
- The ability to transpose and adapt analyses from other markets to determine appropriate domestic efficiency requirements;
- Adopting minimum performance thresholds and applying them as a starting point in a domestic regulatory programme;
- Adopting a common set of upper thresholds that can be used for market pull programmes such as labelling and incentive schemes; and
- Faster and less expensive testing – for compliance and other purposes – as harmonised testing creates a larger choice of laboratories that can conduct product tests.

The testing of SSL products is in many ways different from testing conventional light sources, and a laboratory that can reliably test conventional light sources may not necessarily be able to provide accurate test results for SSL products.

However, despite the benefits of harmonisation, IC 2013 did not aim to create a harmonised lamp test method to be applied all over the world. Many regions prefer to apply their own regional test methods for a number of reasons and any effort to harmonise product test methods would be the responsibility of national and international standardisation bodies. Furthermore, the ideal solution would be to have one international test method that is used for testing of lamps as well as for proficiency testing. In this case, all laboratories undertake to establish their ability – or proficiency – to carry out testing according to a common (harmonised) test standard.

In the absence of a harmonised test method², the SSL Annex sets out on the more modest goal of creating an interim globally harmonised test method for proficiency testing. Regional test methods have many similarities but also differ in many areas. The SSL Annex interim harmonised test method is based on the most stringent requirements of the various LED test methods from around the world. Thus, the harmonised SSL Annex test method for proficiency testing can be used by any laboratory in the world to establish its ability to carry out tests according to any of the major regional/international test methods around the world. This process is explained in detail in the next section.

IC 2013 was conducted through four so-called Nucleus Laboratories: the National Institute of Standards and Technology (NIST) in the USA; the National Lighting Test Centre (NLTC) in China; the National Metrology Institute of The Netherlands (VSL) in the Netherlands; and the National Metrology Institute of Japan in the National Institute of Advanced Industrial Science and Technology (NMIJ, AIST) in Japan. Each of these nucleus laboratories acted as the regional hub of a group of participating laboratories. In August 2012, the Annex published the results of the star-type comparison testing between the four Nucleus Laboratories [4]. The test results showed that these four laboratories had acceptable agreement within the stated expanded uncertainties, confirming their measurement accuracy and their collective capability to measure the performance of the LED lamp test samples.

The IC 2013 was carried out according to the time table shown in Table 1.

Table 1. Timeline of the 2013 interlaboratory comparison

Item	Date
Announcement and opening of application period for participants	22 October 2012
Closure of the application period	30 April 2013
Measurements conducted by the participants	November 2012 – August 2013
Participants Results Reports and Regional Interim Reports issued	January 2014
Final Report of IC 2013 issued to participants	30 June 2014
Final Report of IC 2013 issued to the public	10 September 2014

3.2 Standards and accreditation

The IC2013 was carefully designed to be in compliance with ISO/IEC 17043, “Conformity assessment -- General requirements for proficiency testing” [2]. This was done to ensure that the work would be acceptable to Accreditation Bodies (ABs) around the world, as evidence of proficiency testing for any SSL method of measurement currently being used or in draft form. The four nucleus laboratories are all national metrology institutes or national testing institutes and possess their own primary measurement standards and have developed and validated a measurement method for SSL testing. All of their associated measurement services have been accredited by International Laboratory Accreditation Cooperation Mutual Recognition Agreement (ILAC/MRA) signatories or peer reviewed against ISO/IEC 17025 [5] and registered to Appendix-C, CIPM/MRA. Therefore, the basic competence of the institutions related to the measurements in this IC was well-established.

Prior to IC 2013, it was difficult for many ABs to carry out SSL proficiency testing (PT) as part of their ISO/IEC 17025 [5] accreditation. While some ABs like NVLAP (in the USA) and CNAS (China National Accreditation Service of Conformity Assessment) already provided SSL PTs or comparison tests for SSL testing, other ABs did not, as SSL testing is not within the scope of the ILAC/MRA, and therefore it is not possible for these PTs to be accepted globally. Thus, the fragmentation and disparities in SSL test methods has held back the accreditation and proficiency testing of SSL laboratories.

The ultimate aim of the SSL Annex is that in the future, a robust international programme for SSL accreditation and PT can be established when an international published test method is available. The International Commission on Illumination (CIE) just recently published a Draft International Standard DIS 025/E:2014[3]. However, this standard was not available at the time IC 2013 started, and even now, it will take some time before this new standard is adopted by regulations and accreditation programmes. Figure 3 depicts how IC 2013 can

² The CIE (International Commission for Illumination) has worked many years to develop a new harmonised test method, which is now (in 2015) a draft international standard, however at the time of the IC 2013, it was still being prepared by the committee. CIE Draft International Standard DIS 025/E:2014; Test Method for LED Lamps, LED Luminaires and LED Modules [3].

serve as an interim solution to address some of the problems of the past while also preparing the market for the adoption of the new CIE international test method.

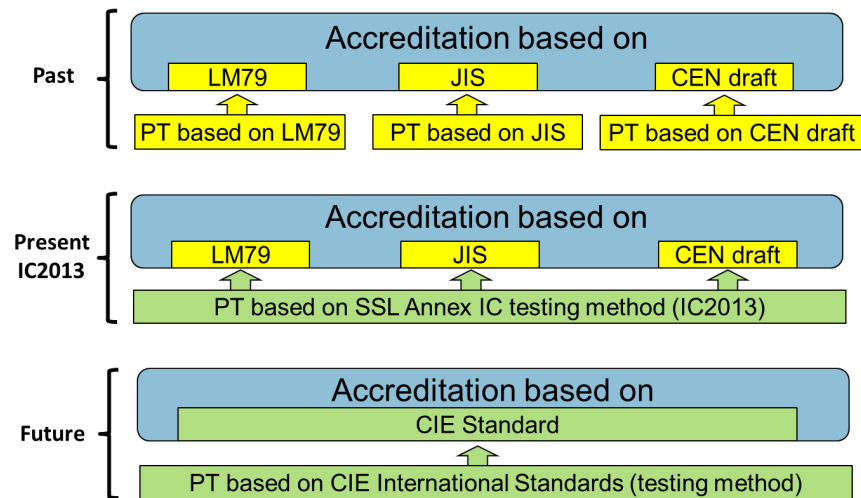


Figure 3. SSL Annex IC2013 serves as an interim solution for laboratory accreditation

In general, PTs for testing accreditation are designed for a specific test method for specific products. However, the SSL Annex experts were able to create a harmonised test method based on the most stringent requirements of the various LED lighting products test methods around the world [6]. Through the use of the IC 2013 test method for measurement of LED lighting products, it is hoped that IC 2013 will be regarded as a PT or valid evidence of the competence of the laboratory for regional test methods, thus enabling the accreditation of laboratories for testing LED products.

4. Test methodology of IC 2013

The detailed design of IC 2013 is described in the SSL Annex Interlaboratory Comparison Generic Protocol [7], and was, as mentioned earlier, based on using four nucleus laboratories that served as reference laboratories to the participants. All the results of the four Nucleus laboratories were then combined for the results presented in the IC 2013 Final Report. The test method used was the SSL Annex IC Test Method 1.0 [6], which covers test methods for many performance characteristics of SSL products. The test method was written in such a way that the measurement requirements encompass those in IES LM-79-08; EN 13032-4 (Draft) prepared by CEN TC169 WG7 and CIE TC2-71 (excluding the parts on LED modules); the test method drafts included in the Annexes of IEC performance standards (drafts) on LED lamps and LED luminaires; the test methods covering LED lamps and LED luminaires in the Japanese standards: JIS C 7801:2009, JIS C 8105-5:2011, and JIS C 7801 Amendment 1: 2012; and the test methods covered in the Chinese standards: GB standards Drafts for self-ballasted LED reflector lamps, and CQC3127-2010, CQC3128-2010, CQC3129-2010, CQC3130-2011. The SSL Annex IC Test Method adopted the most stringent requirements across all of these test methods for the measurement of LED products. In this way, any laboratory complying with this IC Test Method should also be recognised as satisfying the requirements of all of the aforementioned test methods.

The IC 2013 was designed to have five common test artefact types and some additional (optional) test artefacts that were used as shown in the table below.

Table 2. Test artefacts used in IC 2013

Five artefacts common to all regions	Three optional artefacts used by different regions
<ul style="list-style-type: none"> Incandescent lamp (I-AC) 	<ul style="list-style-type: none"> Incandescent lamp – DC operation (I-DC), using the same lamp as I-AC
<ul style="list-style-type: none"> Omnidirectional LED lamp (OD) 	<ul style="list-style-type: none"> Tubular type LED lamp (TL)
<ul style="list-style-type: none"> Directional LED lamp (D) 	<ul style="list-style-type: none"> Remote-phosphor type LED lamp (RP)
<ul style="list-style-type: none"> High correlated colour temperature (HCCT) LED lamp or luminaire (> 5000 K, preferably ≈ 6500 K) 	
<ul style="list-style-type: none"> Low power-factor LED lamp (LPF) ($PF < 0.6$, preferably $PF \approx 0.5$) 	

As depicted in Figure 4, IC 2013 was designed as a star type comparison (see Annex A of ISO/IEC 17043 [2] for the types of PT schemes). The artefacts were first measured by the reference laboratory, sent to a participant laboratory for testing, and then returned and tested again by the reference laboratory (note: a small modification to this procedure was allowed in the Regional Protocol for AIST, NMIJ in Japan). If the observed drift (difference between the first and second measurements by reference laboratory) exceeded $0.8 \times \text{SDPA}$ (Standard Deviation for Proficiency Assessment)³, then the associated results of the artefact were discarded and a replacement artefact was sent for re-measurement. Each nucleus laboratory developed its Regional Protocol in compliance with the SSL Annex IC Generic Protocol, with slight (regional) variations in the artefact types selected.

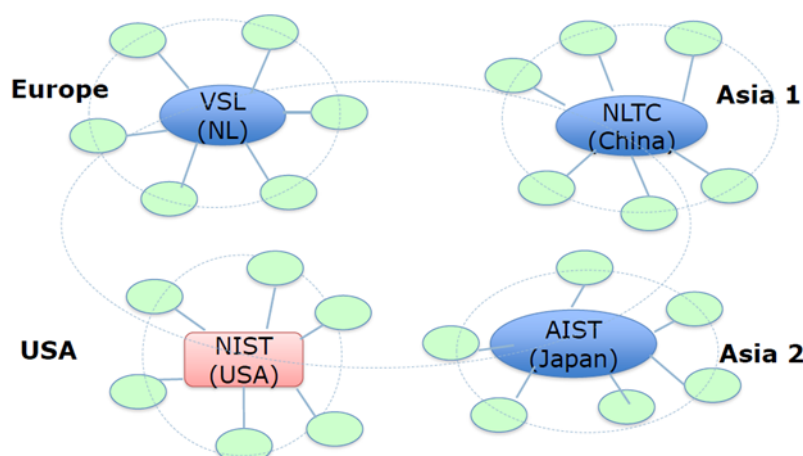


Figure 4. Star type configuration; green ovals represent participants in IC 2013.

To ensure confidentiality in the testing, each participant laboratory was given a “Lab Code” to identify them, but this code assignment was only known by the laboratory contact and the relevant nucleus laboratory. The Lab Codes were assigned to laboratories randomly so the participants in the four regions were all mixed together in the results presented in the Final Report.

The following measurement quantities were measured and compared in the IC 2013 for each of the samples tested:

- 1) Total luminous flux (lm)
- 2) RMS voltage (V) and current (A)
- 3) Electrical active power (W)
- 4) Luminous efficacy (lm/W)
- 5) Chromaticity coordinates x, y
- 6) Correlated colour temperature (K)
- 7) Colour rendering index (CRI) Ra
- 8) Power factor (optional)

³ See Section 7 of Ohno et al., Solid State Lighting Annex 2013 Interlaboratory Comparison Final Report [8]

4.1 Evaluating Participants Results

There are two ways of evaluating the measurement results by the participants: the E_n number and z' score. The E_n number and z' score are used for different purposes. The concept of the E_n number is to test whether the claimed measurement uncertainties of a laboratory are valid, and this is suitable when the uncertainty is in the scope of accreditation and needs to be certified (typically the case for *calibration laboratory accreditation*). The z' score, on the other hand, is to test whether the laboratory's results are within an acceptable range of variation, and is suitable for *testing laboratory accreditation* (supporting product certification activities) which examines the laboratory's competence and compliance to the reference test method. For laboratory accreditation programmes having both purposes (i.e., serving for product certification activities as well as certifying the reported uncertainties), the use of both the E_n number and z' score would be appropriate.

The z' score is calculated for all results, and is determined by:

$$z' = \frac{x - X}{\sqrt{\hat{\sigma}^2 + u_X^2 + u_{\text{drift}}^2}}, \quad (1)$$

where x is the value measured by the participant and X is the assigned value measured by the reference laboratory, $\hat{\sigma}$ is the Standard Deviation for Proficiency Assessment (SDPA) value and, in this IC test, is the generic standard uncertainty of a participant's measurement; u_X is the standard uncertainty of the reference value (average of uncertainties of measurement of the comparison artefacts by four nucleus laboratories reported in the Nucleus Laboratory Comparison Report published in 2012 [4]). The value of u_{drift} is the uncertainty contribution from the expected drift in the test artefacts (controlled to within $0.8 \times \text{SDPA}$)⁴ and calculated by:

$$u_{\text{drift}} = \frac{0.8 \cdot \hat{\sigma}}{2\sqrt{3}}. \quad (2)$$

If these equations are given in relative uncertainties $\hat{\sigma}_{\text{rel}}$, $u_{X,\text{rel}}$, $u_{\text{drift,rel}}$, then z' is calculated by:

$$z' = \frac{(x - X) / X}{\sqrt{\hat{\sigma}_{\text{rel}}^2 + u_{X,\text{rel}}^2 + u_{\text{drift,rel}}^2}}, \quad (3)$$

The values of u_X or $u_{X,\text{rel}}$ and $\hat{\sigma}$ or $\hat{\sigma}_{\text{rel}}$ were pre-determined, and are listed in Table 2 in the SSL Annex IC Generic Protocol [7]. The values of u_X or $u_{X,\text{rel}}$ were determined as the averages of the nucleus laboratories' measurement uncertainties reported in the preceding comparison among the nucleus laboratories [4]. The values of $\hat{\sigma}$ or $\hat{\sigma}_{\text{rel}}$ were determined as expected generic uncertainties of measurement of each quantity by the participants. The determined values of these parameters used in IC 2013 and their uncertainty budget are provided in in Appendix 1 of the Final Report; and also in the Regional Interim Reports distributed in each region.

If the uncertainties of measurements were reported by the participant, E_n numbers are calculated according to

$$E_n = \frac{x - X}{\sqrt{U_{\text{lab}}^2 + U_{\text{ref}}^2}}, \quad (4)$$

where:

x : value measured by the participant

X : assigned value (average of reference laboratory measurements, before and after)

U_{lab} : expanded uncertainty ($k=2$) of a participant's result

U_{ref} : expanded uncertainty ($k=2$) of the assigned value

⁴ See Section 2 of Ohno et al., Solid State Lighting Annex 2013 Interlaboratory Comparison Final Report [8]

U_{ref} is calculated by

$$u_{\text{ref}} = \sqrt{\left(\frac{u_1 + u_2}{2}\right)^2 + \left(\frac{X_1 - X_2}{2\sqrt{3}}\right)^2} \quad (5)$$

and

$$U_{\text{ref}} = 2 u_{\text{ref}}, \quad (6)$$

where X_1 and X_2 are measured values by the reference laboratory, before and after the participant's measurement, respectively, and u_1 and u_2 are their absolute standard uncertainties at the first and second measurements.

Equation (5) above assumes that the two measurements by the reference laboratory are fully correlated. The second term in the square root is a square of the standard uncertainty associated with the drift of the artefacts as measured by the reference laboratory (taken as a rectangular distribution [9]).

Reporting measurement uncertainties in compliance with ISO/IEC Guide 98-3 [9] was required by the SSL Annex Test Method 1.0 [6] for all results. However, results without uncertainty values were also accepted (noted as a non-compliance to the IC test method), since these results may still qualify for certain testing accreditation programmes that do not require uncertainty values in test reports.

The assigned value is a value attributed to a particular property of a proficiency test item [7]. Assigned values in IC 2013 were given by the nucleus laboratories, and were calculated as the mean of the measurements by the nucleus laboratory taken for each quantity before sending and after return of artefacts from each participating laboratory. The criteria used to analyse and evaluate the performance are given by the z' score (defined in ISO 13528 [10]) and E_n number (defined in ISO 13528 [10] and ISO/IEC 17043 [2]). However, it should be noted that the E_n numbers were not calculated if the uncertainties were not reported by the participant.

Generally, $|E_n| > 1.0$ is considered to be unsatisfactory. This means that the difference in the quantities measured by the nucleus laboratory and the participant laboratory is greater than the expanded uncertainty of the comparison. In the case of the z' scores, a value of $2.0 < |z'| < 3.0$ is considered to be questionable, and $|z'| \geq 3.0$ is generally considered to be unsatisfactory, but the judgment as to whether the result is acceptable will depend on the accreditation body.

5. Results

The results of IC 2013 were first analysed in each region by comparison with the assigned values as measured by the regional reference laboratory (nucleus laboratory), then the results of all the regions were combined based on the equivalence among the nucleus laboratories. The differences between the participants' measurement values and the assigned values, z' scores and E_n numbers, of all measured quantities of all the artefacts, for all participants, were determined. The numerical results of these values for all conditions were reported in the Participants Results Reports, and the graphical representations of all these results in each region were reported in the Regional Interim Reports (issued by each nucleus laboratory) for accreditation purposes. The IC 2013 Final Report presents graphical results of the differences in all conditions, but the z' scores and E_n numbers are presented only for certain conditions for technical discussions [8].

In this paper, due to length constraints, just a few figures were taken from the IC 2013 Final Report to provide examples of the results presented. Interested readers are encouraged to download a copy of the Final Report for the complete set of results.⁵ In the report, the (relative) differences of results between the participant (Lab) and the nucleus laboratory (Ref), defined by (Lab - Ref) or (Lab - Ref)/Ref, for all quantities, all participants, and all artefact types are presented in graphical form. The horizontal axis on these graphs (from 1 to 123) indicates individual data results from the 110 participant laboratories (laboratory numbers), where the order of laboratories from the four different regions has been mixed.

The error bars in the figures show the uncertainties of measurement (expanded uncertainty with a coverage factor, $k=2$) by the participants, and are shown only when the uncertainties are reported (note: all NVLAP-linked results do not include uncertainties). There are a few cases where laboratories reported uncertainties but not for colour quantities. Note that when a data point does not have an error bar, it means, in most cases, no uncertainty was reported, but in some cases, the reported uncertainty value is so small that the bar is hidden behind the data point. The dashed lines in the figures (in pink) show the average values of the uncertainties ($k=2$) by the

⁵ <http://ssl.ica-4e.org/task-2-ssl-testing/2013-ic-final-report>

reference (nucleus) laboratories for all the points plotted in each figure. The reference (nucleus) laboratory uncertainty values are similar to but not necessarily the same as those reported in the Nucleus Laboratory Comparison Report [4] and in the IC Generic Protocol [7] due to slightly different products used in the IC 2013 in each region.

Figure 5 shows the relative differences in luminous flux (light output) measurements between the nucleus laboratories and the participant laboratories for omni-directional (OD) LED lamps. There are two outliers depicted in this graph, laboratory data sets #61 and #84 which both had relative differences exceeding -20%.

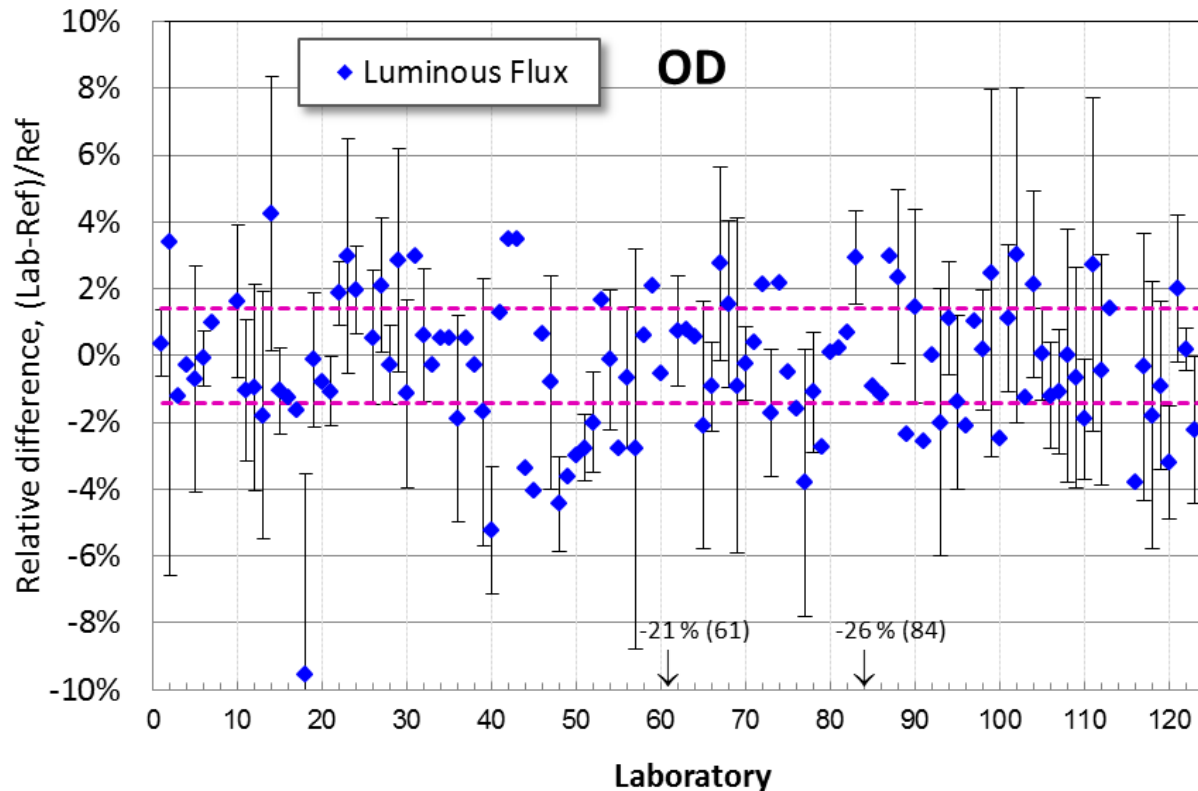


Figure 5. Relative differences of total luminous flux for omnidirectional LED lamp (OD)

Figure 6 provides a plot of the relative differences in RMS current measurements for the low power factor lamps. The measurement variations in RMS current turned out to be much larger than expected, and the reported uncertainties were significantly underestimated. It is evident that there are some additional uncertainty components that were overlooked. The variations of the LED lamps are significantly larger than those of the incandescent lamp by more than an order of magnitude.

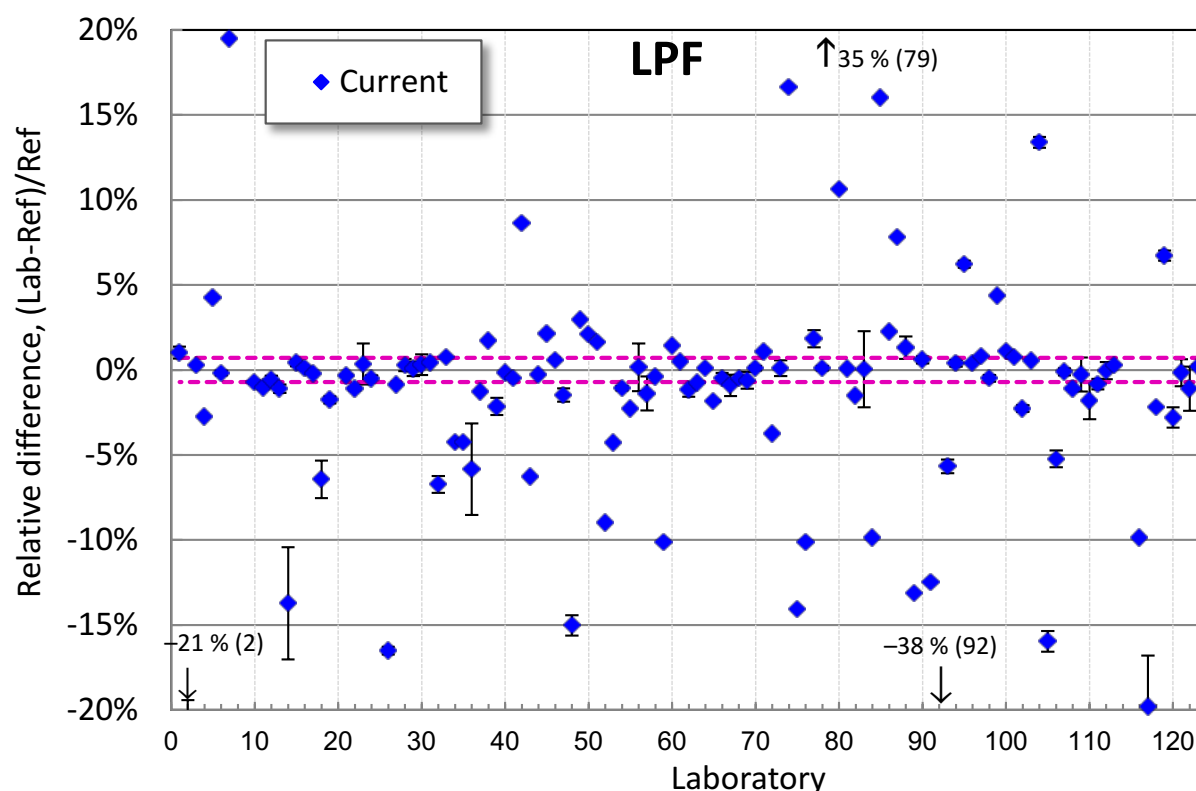


Figure 6. Relative differences of RMS current for low power factor lamp (LPF)

More than 80 graphs showing comparisons of the measurements between the nucleus laboratories and the participants are available in the Final Report [8].

6. Analysis and discussion

The results (measured values of quantities and reported uncertainties) were compared between different artefact types, between different quantities, and between different instrument types (sphere system and goniophotometer). Correlations between different quantities were also analysed. From these analyses, the following points have been concluded.

The results for total luminous flux of the LED test lamp artefacts (see Table 2 for description of these items) measured by most of the laboratories agreed to within $\pm 4\%$ (OD) to $\pm 5\%$ (D, LPF, HCCT), which is at an expected level of agreement. The results for chromaticity x, y measured by most of the laboratories agreed to within ± 0.005 for all artefact types, which is also at the expected level of agreement. These results verified the levels of the uncertainties estimated for the measurements of these products by laboratories using a well-established test method. It also confirmed that the test method used in this IC, a consolidation of several available test methods (SSL Annex Interlaboratory Comparison Test Method 1.0 [6]) was effective in reducing the variation in results.

On the other hand, a few extremely large deviations in the results were observed, for example, up to 30% in luminous flux or up to 0.2 in chromaticity x, y in each artefact type. These extreme results are most likely caused by some major flaws at the participant laboratories in meeting the requirements in the test method. These large deviations by some laboratories demonstrate the importance of proficiency testing, as these laboratories would not have become aware of their problems without participating in such an interlaboratory comparison.

The electrical measurement results also identified some issues. The variations in the results of RMS current for LED lamps were primarily within $\pm 3\%$ (OD, D, HCCT) to $\pm 15\%$ (LPF) with some deviations much larger than expected (up to 38%), resulting in high values of z' and E_n for many participants. This result indicates that the generic uncertainty for RMS current to determine the denominator of z' was underestimated, and this should be considered by accreditation bodies when they use these results. The results also indicate that uncertainties reported by many participants were significantly underestimated. However, it was found that the variations in

measured RMS current did not affect photometric and colorimetric values significantly, suggesting that agreement in RMS current is not very critical. This is explained by the finding that deviations in RMS current were strongly correlated with power factor in the direction to cancel the changes in active power, though not in all cases. The variations in measured power factor were also larger than expected, mostly within ± 0.02 (OD, D, HCCT) to ± 0.1 (LPF). These large variations in the electrical measurements may be caused by differences in the characteristics of the AC power supplies used by the participants, in particular, the output impedance, some of which are discussed in Ref. [11]. This is one of the remaining issues for the current test methods for LED lighting products, and future improvements are expected.

The uncertainties reported by the participants were found to be in a very large range (often more than two orders of magnitude) and were often significantly underestimated. Some laboratories reported unreasonably small uncertainties (e.g., 0.0001 in chromaticity x, y) or unreasonably large uncertainties (e.g., 10 % in luminous flux or 0.02 in chromaticity x, y). Several laboratories (not those linked) did not report uncertainties at all or did not report uncertainties of colour quantities (i.e., chromaticity x, y , CCT, CRI). These observations indicate that uncertainty evaluation (especially for colour quantities) is still very difficult for the SSL industry, and reported uncertainties are often not reliable. Practical methods and tools for uncertainty evaluation of measurements, as well as educational documents and training for the SSL industry on practical uncertainty evaluation are urgently needed.

In addition to the differences of participants' results from the reference values, both z' scores and E_n numbers were calculated in this IC test, for possible use by ABs. The IC results show that some laboratories would pass on E_n number but fail on z' score or vice versa. In particular, there were some cases where laboratories claiming large uncertainties would pass on the E_n number though the deviations in their results were very large. Thus, the use of E_n number alone can be problematic when measurement variations need to be limited by the accreditation programme. In practice, the E_n number is suitable for the purpose of assessing the validity of claimed uncertainties (e.g., in calibration laboratory accreditation). The z' score is suitable for the purpose of testing laboratory accreditation, which examines a laboratory's competence and compliance to a test method which is developed to limit measurement variations as is often required in product certification activities. For laboratory accreditation programmes serving both purposes (i.e., serving for product certification activities as well as certifying the reported uncertainties), the use of both the E_n number and z' score would be appropriate. In this study, it was found that the E_n number could be problematic where laboratories had difficulty in uncertainty evaluation, as shown in IC 2013 for colour quantities. It was also found that the z' score could be problematic if the denominator values were not appropriately specified, as was the case of RMS current measurements in this IC. The results of IC 2013 may be utilised for future SSL proficiency testing using z' score or a similar metric.

7. Conclusions

A large scale international interlaboratory comparison, IC 2013 by IEA 4E SSL Annex, was conducted successfully. This IC test was an attempt to establish a common proficiency test for SSL products test laboratories that could serve accreditation programmes supporting different regulations and government programmes using different regional test methods. For this purpose, the special test method, SSL Annex Interlaboratory Comparison Test Method 1.0 [6], was needed and developed by the SSL Annex. A solution for international harmonisation of SSL testing and accreditation would be to use the international test method for SSL products, which will soon be published by the CIE. Countries would then choose whether to harmonise to this test method standard based on their own needs and regulatory requirements, enabling worldwide mutual recognition of SSL product testing and laboratory accreditation.

Since IC 2013 was intended to be used as proficiency testing for SSL testing accreditation, it was designed in compliance with ISO/IEC 17043 [2]. Measurements of photometric, colorimetric, and electrical quantities were compared using at least four different types of LED lamp. IC 2013 included 54 laboratories as direct participants from 18 countries. In addition, the recent results of 35 US laboratories in NVLAP PTs and NIST Measurement Assurance Program were linked to IC 2013. Data from an additional 21 laboratories in the APLAC Proficiency Test T088 were also linked to IC 2013, making it a comparison of test results from 110 laboratories and 123 sets of data.

The results of the comparison clarified the levels of agreement of measurements of photometric, electrical, and colorimetric quantities of LED lamps and luminaires, which will help governments and manufacturers around the world to learn the uncertainties of the measurements of SSL products and ensure that new products sold to consumers and companies are of high quality and meet the claimed performance.

The results also revealed some cases of extremely large errors (e.g., 25 % error in luminous flux). Identification of these large deviations for a few laboratories demonstrates the importance of proficiency testing, as these laboratories would not have been aware of their particular problems without having participated in this interlaboratory comparison. Testing laboratories can verify their measurement uncertainties and gain confidence in their measurements only by interlaboratory comparisons.

Some of the IC 2013 participants already applied for the testing accreditation program for SSL products in their countries using the IC 2013 results as proficiency testing, which verifies that the original aim to assist accreditation bodies is being achieved.

The CIE test method (DIS 025 [3]) has been published and the final standard CIE S 025 is to be published shortly. This will serve as the international test method for SSL products, and should be used by SSL testing accreditation program worldwide. However, it will take some before this can be adopted by regulators and accreditation bodies in different regions, and thus, the approach taken by IC 2013 may be needed again till the test method is fully harmonised internationally.

The IC 2013 provided many laboratories in several countries with new knowledge and experience in PT for the measurement of SSL products. It also established a basis to promote SSL laboratory testing accreditation worldwide in support of regulations and government programmes to further accelerate the development of SSL. This is hoped to support and speed up the introduction of energy efficient LED technology around the world.

References

- [1] Bennich, P., Borg, N., Scholand, M., “On the bright side of life: International efforts to accelerate market adoption of LEDs while avoiding the pitfalls of CFLs”, Proceedings from eceee Summer Study 2013, Paper #6-491-13, Stockholm, Sweden
- [2] ISO/IEC 17043:2010, Conformity Assessment – General Requirements for Proficiency Testing
- [3] CIE DIS 025: 2014, Test method for LED lamps, luminaires, and modules
- [4] IEA 4E SSL Annex: Summary Report of Nucleus Laboratory Comparison, http://ssl.iea-4e.org/files/otherfiles/0000/0057/4E_SSL_Annex_-_Nucleus_Laboratory_Comparison_Report_final.pdf
- [5] ISO/IEC 17025:2005, General requirements for the competence of testing and calibration laboratories
- [6] IEA 4E SSL Annex: Interlaboratory Comparison Test Method 1.0, http://ssl.iea-4e.org/files/otherfiles/0000/0059/SSL_Annex_2013_IC_Test_Method_v.1.0.pdf
- [7] IEA 4E SSL Annex: Interlaboratory Comparison Generic Protocol, http://ssl.iea-4e.org/files/otherfiles/0000/0053/SSL_Annex_IC_Generic_Protocol.pdf
- [8] Ohno, Y., Nara, K., Revtova, E., Zhang, W., Zama, T., Miller, C., “Solid State Lighting Annex 2013 Interlaboratory Comparison Final Report”, <http://ssl.iea-4e.org/task-2-ssl-testing/2013-ic-final-report>
- [9] ISO/IEC Guide 98-3: 2008, Uncertainty of measurement -- Part 3: Guide to the expression of
- [10] ISO 13528:2005, Statistical methods for use in proficiency testing by interlaboratory comparisons uncertainty in measurement (GUM:1995)
- [11] Martinsons, C., Zong, Y., Miller, C., Ohno, Y., Olive, F., Picard, N., “Influence of Current and Voltage Harmonic Distortion on the Power Measurement of LED Lamps and Luminaires,” Proc. CIE Centenary Conference, Paris, CIE x038:2013, pp. 290 – 299, (2013)

Acknowledgments

The authors would like to thank the governments of Australia, China, Denmark, France, Japan, Republic of Korea, Sweden, the United Kingdom and the United States for their sponsorship of the first term (2010-2014) of the SSL Annex, during which time the 2013 Interlaboratory Comparison was conducted.