

To be published in Applied Optics:

Title: A Multilateral Spectral Radiance Factor Scale Comparison
Authors: Christian Strothkämper, Alejandro Ferrero, Annette Koo, Priit Jaanson, Guillaume Ged, Gael
Accepted: 31 January 17
Posted 06 February 17
Doc. ID: 279674

Published by

OSA[®]

The Optical Society

A Multilateral Spectral Radiance Factor Scale Comparison

C. STROTHKÄMPER,^{1*} A. FERRERO,² A. KOO,³ P. JAANSON⁴, G. GED,⁵ G. OBEIN,⁵ S. KÄLLBERG,⁶ J. AUDENAERT,⁷ F. B. LELOUP,⁷ F.M. MARTÍNEZ-VERDÚ,⁸ E. PERALES,⁸ A. SCHIRMACHER,¹ J. CAMPOS²

¹ Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany

² Instituto de Óptica "Daza de Valdés", Agencia Estatal CSIC, c/ Serrano 121, 28006 Madrid, Spain

³ MSL Measurement Standards Laboratory of NZ, Callaghan Innovation, 69 Gracefield Rd, Lower Hutt 5010, NEW ZEALAND

⁴ Centre for Metrology MIKES, VTT Technical Research Centre of Finland Ltd, Tekniikantie 1, 02150 Espoo, Finland

⁵ Conservatoire National des Arts et Métiers, 61 rue du Landy, 93210 La Plaine St Denis, France

⁶ SP Technical Research Institute of Sweden, P.O. Box 857, SE-501 15 Borås, Sweden

⁷ Light & Lighting Laboratory, KU Leuven, Gebroeders De Smetstraat 1, B-9000 Gent, BELGIUM

⁸ Color & Vision Group, Department of Optics, Pharmacology and Anatomy, University of Alicante, Carretera de San Vicente del Raspeig s/n, 03690 Alicante, Spain

*Corresponding author: christian.strothkaemper@ptb.de

The field of spectral radiance factor (SRF) measurements has seen growing interest in recent years. Scale conformity has so far only been established between the national metrology institutes (NMIs) of Germany and the USA. This study aims at a bigger, multilateral scale comparison. For this purpose, a total of six NMIs participated in a scale comparison of gonio-spectrophotometers based on neutral and colored diffusely reflecting ceramics samples. In addition, two universities, providing a home-built gonio-reflectometer and two widely used commercially available color measurement instruments, respectively, were involved. The wavelength range of the scale comparison covers the visible wavelength range from 380 nm to 780 nm. Results indicate systematic issues and the uncertainty evaluation of the NMIs requires further work, though for the greatest part of the covered spectral range the agreement is good.

OCIS codes: (290.1483) BSDF, BRDF, and BTDF; (120.1840) Densitometers, reflectometers; (120.5700) Reflection

DOI:

Introduction

In recent years the demand for measurements and calibrations of the spectral radiance factor [1] (SRF) has been rising. One reason is an increased demand in R&D applications, such as material characterization of satellite components [2] or calibration of reflectance standards for space-based earth observation [3,4]. The industrial interest arises for example from the necessity of characterizing goniochromatic coatings, whose visual appearance changes significantly depending on the combination of irradiation and viewing directions [5], typically due to the use of interference [6] or diffraction pigments.

Measurements of the SRF using arbitrary angles of incidence and reflection located within the half-space

above the sample surface under test is possible with some lab-based, custom-built gonio-reflectometers [7,8,9,10,11,12,13]. Commercially available multi-angle spectrophotometers are usually restricted to a small number of measurement geometries such as those recommended in standards of the Deutsches Institut für Normung e.V. (DIN) [14] and ASTM International [15]. In contrast to NMI devices they lack a (GUM compatible [16]) uncertainty evaluation.

SRF reflectance scales are currently not included in any key comparison arranged by the Consultative Committee of Photometry and Radiometry (CCPR) of the International Committee for Weights and Measures (CIPM), though in CCPR K5, the NMIs of Finland and the United Kingdom reported measurements of diffuse

Description of the measurand and measurement facilities

The quantity that is derived from the measurements and compared between the different instruments is the spectral radiance factor (SRF), defined as:

$$\beta(\theta_i, \varphi_i, \theta_r, \varphi_r; \lambda) = \frac{L_r(\theta_i, \varphi_i, \theta_r, \varphi_r; \lambda)}{L_{\text{PRD}}}, \quad (1)$$

where L_r denotes the radiance of a sample and L_{PRD} the radiance of the perfectly reflecting diffuser under identical conditions of illumination and viewing. The angles in eq. (1) refer to a spherical coordinate system with a polar angle θ (measured against the sample surface normal) and an azimuthal angle φ (measured against a marking at the sample edge) to describe the directions of illumination (θ_i, φ_i) and observation (θ_r, φ_r) [8]. So the angle of incidence of the illumination on the sample is θ_i and the standard geometries 45:0 and 0:45 are shorthand for (45,0,0,0) and (0,0,45,0).

The radiance of the PRD is independent of the geometry, i.e.

$$L_{\text{PRD}} = \frac{\partial^2 \Phi(\theta_r, \varphi_r)}{\partial \Omega \partial A \cos \theta_r} = \text{const} \quad (2)$$

Because the PRD does not absorb any radiation, the irradiance $E_i = \partial \Phi / \partial A$ on the sample is identical to the integral of the radiance from the surface over a solid angle of 2π in case of a PRD. This gives the simple relation

$$E_i = \int L_{\text{PRD}} \cos \theta_r d\Omega = L_{\text{PRD}} \cdot \pi \quad (3)$$

so that the SRF can also be written as

$$\beta = \pi \frac{L_r}{E_i}. \quad (4)$$

Instruments that do absolute measurements (cf. Table 2) calculate the radiance L_r from the sample and the irradiance E_i on the sample from the photodetector signals of the light from the sample surface (S_r) and the light source (S_i) and geometrical parameters of the instrument like apertures and distances. The photodetector signals are corrected for dark currents and non-linearity of the detector. An example calculation from the PTB instrument [8] yields the formula

$$\beta = \frac{\pi d_i^2}{A_i \cos \theta_i} \frac{S_r}{S_i}, \quad (5)$$

where A_i denotes the aperture of the light source and d_i the distance of the light source aperture to the sample surface. For the instruments of MIKES, CNAM, MSL a different formula holds true

$$\beta = \frac{\pi d_r^2}{A_r \cos \theta_r} \frac{S_r}{S_i}, \quad (6)$$

where A_r denotes the aperture of the detector and d_r the distance of the detector aperture to the sample surface (cf. ref. [21] for an example derivation).

In the case of relative measurements, a calibrated white standard, whose SRF β_{cal} is known, is required for the measurements and the SRF of the sample is obtained as

$$\beta = \frac{S_r}{S_{\text{WS}}} \beta_{\text{cal}}. \quad (7)$$

Here, S_{WS} is the photodetector signal of the light reflected from the white standard in the same geometry as the sample. The incident light from the light source is not directly measured. Eq. (7) is valid for the evaluations made by SP, CSIC and KU Leuven. The commercial instruments perform relative measurements but the exact method of their data treatment is not described; it is presumed to be equivalent to the method described in eq. (7) with the values of β_{cal} stored (and periodically renewed) in the instrument.

Figure 2 exemplifies the spectral radiance factor of the different sample types. The three grey scale standards exhibit a relatively uniform spectral distribution but differ in amplitude. As such they should be differently affected by a possible nonlinearity of measuring device, noise, respectively, dark signals. The three colored samples show a non-uniform spectral distribution and are thus much more influenced by the wavelength uncertainty ($\frac{\partial \beta}{\partial \lambda} u_\lambda$). Apart from that, each sample can be expected to exhibit a different angular behavior of the spectral radiance factor, which due to the angular positioning uncertainty, translates in different overall uncertainties.

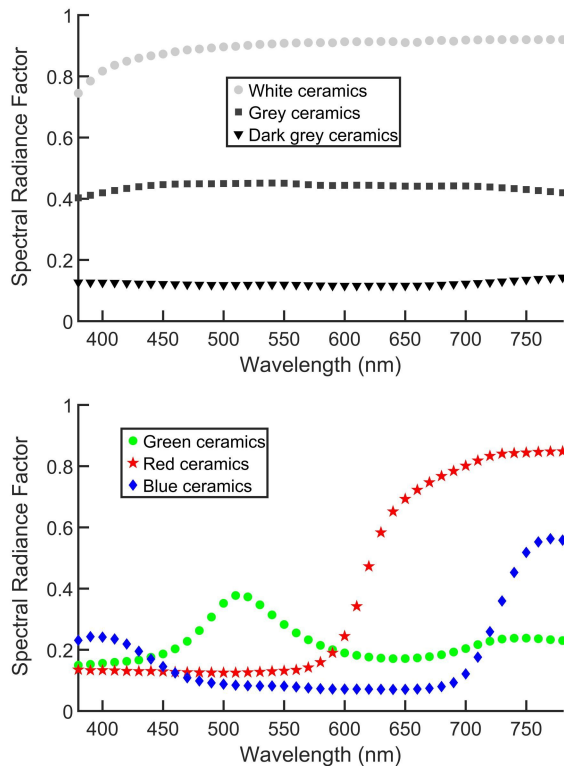


Figure 2 Spectral reflection characteristics of the ceramic samples.

An overview of the basic parameters of the measurement facilities compared in the study is presented in Table 2.

The GONIO facility at PTB is described in detail in [8]. In brief, a special light source [22] based on a halogen lamp with a spatially homogenous beam profile is set up on a ring mount so that it can be rotated around the sample. The sample is mounted on a robot arm in the center of the ring mount. Light reflected or scattered from the sample is detected by a stationary imaging system. Since the robot arm can position the sample in any spatial orientation and the light source can be rotated around the robot arm, nearly any directional configuration of incident light and observation with respect to the sample surface can be realized. The illumination is unpolarized and no polarizer is used in the detection unit. Since the refraction efficiency of the gratings in the monochromator is generally strongly dependent on the polarization, this setup requires that the sample does not polarize the reflected light.

The instruments at CNAM, MSL, and CSIC are also robot based and work in a very similar way, except that the illumination unit is stationary and the detection unit is attached to a ring mount. The CNAM goniospectrophotometer, is composed of an illumination unit put on a fixed table, a robot arm which acts as the sample holder and a detection unit that can be rotated

around the robot arm on a large ring with a diameter of 2 m. A detailed description can be found in [23]. For illumination a broadband Xenon arc is focused at the entrance slit of a single Czerny-Turner monochromator with a bandwidth of 3 nm. When it leaves the monochromator the light is modulated by an optical chopper and any partial polarization created by the gratings is removed by a Lyot depolarizer. The detection is made of the combination of a Silicon photodiode, a current voltage converter and a lock-in amplifier.

The instrument at MSL is based on a robot arm sample holder, a stationary illumination system, and a detection unit traversable on a ring mount. Illumination is done by monochromatic, collimated and polarized light derived from a tungsten halogen lamp after passing a double monochromator and the radiance of the sample is detected without further wavelength or polarization filtering.

The GEFE gonioreflectometer of CSIC [11,12] performs relative measurements using as a reference a white standard with traceability to the NPL. The instrument has a fixed irradiation system, the sample is mounted on a robot arm and the detection unit is placed on a ring, rotating around the robot arm. The light source is a xenon lamp or an incandescence lamp (depending on the spectral range), that is placed before a single monochromator if monochromatic instead of broadband illumination is desired. In order to irradiate the specimens uniformly and with an almost collimated beam, a long Köhler optical system is used. When broadband irradiation is applied, a spectroradiometer (Konica-Minolta CS-2000 A) is used to measure spectral radiance in the visible range (380 nm - 780 nm). Alternatively, for the NIR range, irradiation is selected as monochromatic and an InGaAs photodiode is used to measure the radiance.

At the MIKES-Aalto gonioreflectometer [24] the illumination system is stationary and both, sample and detector are mounted on two co-axial rotary stages. Hence, this instrument can perform only in-plane measurements. A xenon arc or a quartz-tungsten halogen lamp serves as the light source, a double monochromator selects the wavelength of irradiation, which is polarized. The collection system consists of a precision aperture giving a solid angle of 2.2 msr, a lens, a silicon photodetector and a low noise current preamplifier.

SP uses a commercial spectrophotometer (Perkin Elmer Lambda 900) with a Labsphere accessory for 0/45 geometry. The main reference is a glossy white standard with traceability to MIKES. An uncertainty

component is added for reference drift, as well as for the repeatability and nonlinearity.

UA possesses two commercial instruments, the MA98 from X-Rite and the BYK-mac from BYK-Gardner.

The gonireflectometer of KU Leuven is a home-built instrument [8], comprising a stationary Xe broadband illumination source and a rotatable detection unit, which incorporates a detector head (lens and cylindrical integrating cavity) coupled to an Oriel spectrometer with interchangeable grating. At the spectrometer exit plane, a cooled CCD detector is mounted. The detector head can be turned around the sample under test by aid of two stepping motors, thereby providing an almost full 3D spatial coverage. The sample under test is mounted on a sample holder which comprises two manually adjustable rotation stages, allowing for implementation of an adjustable light incidence angle.

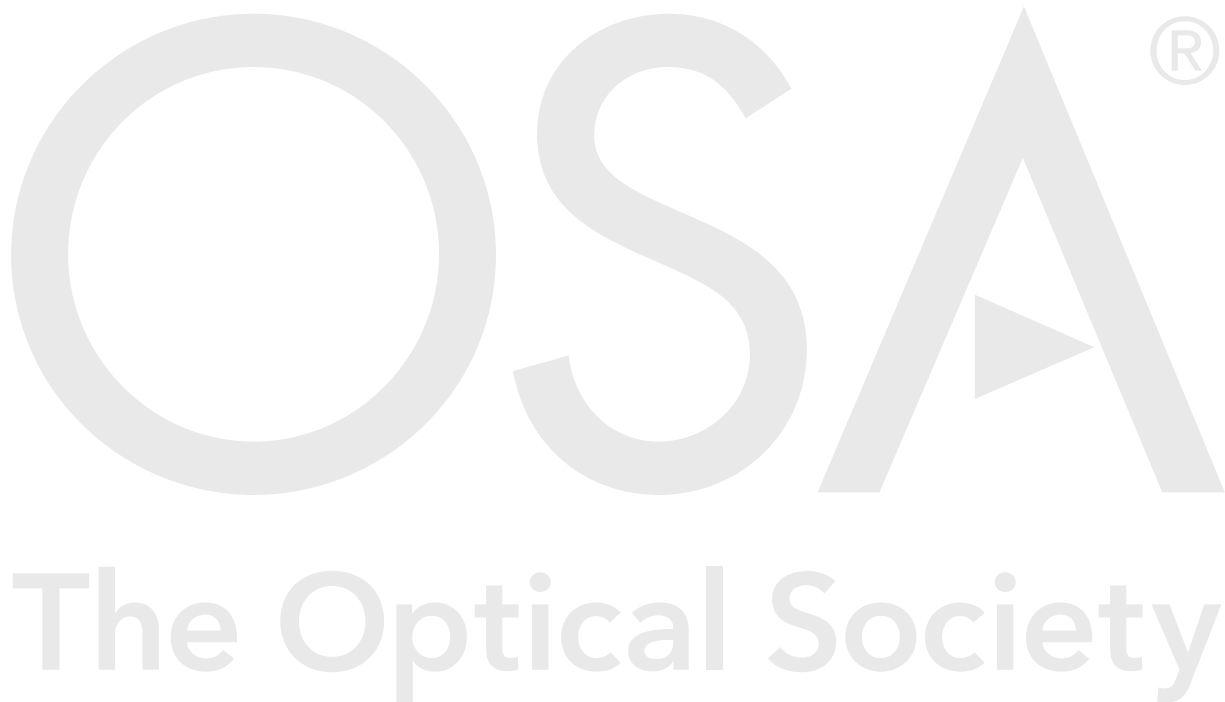


Table 2 Basic parameters of the different measurement facilities.

Participant	PTB	MIKES	CSIC	MSL	CNAM	SP	UA (BYK-mac)	UA (MA98)	KU Leuven
Measurement mode	absolute	absolute	relative	absolute	absolute	relative	relative	relative	relative
Irradiation spectrum	broadband	monochrom.	broadband	monochromatic	monochrom.	monochrom.	broadband	broadband	broadband
Irradiation polarization	unpolarized	polarized	almost unpolarized	polarized	unpolarized	unpolarized	unpolarized	unpolarized	unpolarized
Detection scheme	Monochromator and Si-Photodiode	Si-Photodiode	Spectroradiometer Konika-Minolta CS-2000	Si-Photodiode	Si-Photodiode	Photomultiplier	undisclosed	Undisclosed	Monochromator and Si-CCD
Aperture angle of irradiation	3°	2.3°	0.5°	1.6°	0.68°	1°	unknown	unknown	0.26°
Measured Area (0:45 geometry)	314mm ²	120 mm ²	20 mm ² at normal incidence	346 mm ²	81 mm ²	36 mm ²	529 mm ²	144 mm ²	133 mm ²
Aperture angle of detection	0.64°	3.03°	2.50°	3.43°	1.09°	5°	unknown	unknown	1.66°
Spectral bandwidth	3 nm	5 nm	3 nm – 4 nm	2.6 – 5.2 nm	3 nm	5 nm	10 nm	10 nm	4 nm
Number of Calibrations	9	3	9	3	3	5	3	3	10

Table 3 Overview which uncertainty contributions are taken into account by each participant

	PTB	MIKES	CSIC	MSL	CNAM	SP	KU Leuven
Repeatability of I/U measurement	+	+	+	+	+	+	+
Nonlinearity of I/U measurement	+	+	+	+	+	+	+
Drift of light source intensity	+	+	+	+	+	+	+
Repeatability of angular adjustment	+	+	+	+	+	+	-
Trueness of angular adjustment	+	+	+	+	+	+	-
Wavelength	+	+	+	+	+	+	+
Solid Angle of Illumination	+	-	+	+	-	-	-
Solid Angle of Detection	-	+	-	+	+	-	-
Uniformity of irradiation on sample surface	+	-	+	-	-	-	-
Inhomogeneity of sample surface	-	-	-	-	-	-	-
Angular positioning of sample surface	+	-	-	+	+	-	-
Polarization dependence	+	-	-	+	-	-	-
Fluorescence	-	-	-	-	-	-	-

Methodology of the comparison

Both sample sets are evaluated separately following the scheme in ref. [25] (Procedure A), which is also briefly stated below. The evaluation is based on the assumptions that the samples are stable, that the institute's measurements are independent [26] and that each institutes measurand can be described by a Gaussian distribution with mean and standard deviation equal to the institutes measured value of β and the standard uncertainty. For each sample s , each of the two possible geometries g and every wavelength λ a weighted mean is calculated

$$y_{s,g,\lambda} = \frac{\sum_n w_{n,s,g,\lambda} x_{n,s,g,\lambda}}{\sum_n w_{n,s,g,\lambda}} \quad (8)$$

from the measurements $x_{n,s,g,\lambda}$ of the participants' instruments (indexed by n) using as weights $w_{n,s,g,\lambda}$ the inverse squares of the standard uncertainties $u_{n,s,g,\lambda}$. The data acquired by the commercially available instruments and by the home-built instrument of KU Leuven are not used for the calculation of the mean, because these instruments do not offer a fully GUM compatible uncertainty evaluation [16].

From the weighted mean and its standard uncertainty

$$u(y_{s,g,\lambda}) = \left(\sum_{n=1}^N w_{n,s,g,\lambda} \right)^{-1/2} \quad (9)$$

χ^2 values are calculated:

$$\chi_{s,g,\lambda}^2 = \sum_{n=1}^N \frac{(x_{n,s,g,\lambda} - y_{s,g,\lambda})^2}{u_{n,s,g,\lambda}^2} \quad (10)$$

The data are regarded as consistent if $\chi_{s,g,\lambda}^2$ is mostly within the 95% quantile of the χ^2 distribution with $N-1$ degrees of freedom (N is not constant but depends on sample and geometry). Yet, 5 % of the data can always be expected to lie outside that limit due to statistical variability. If the data are consistent, it is meaningful to calculate the degree of equivalence of instrument n as $(d_{n,s,g,\lambda}, U(d_{n,s,g,\lambda}))$, whose components are defined as (cf. 4. (c) in ref. [25]):

$$d_{n,s,g,\lambda} = x_{n,s,g,\lambda} - y_{s,g,\lambda} \quad (11)$$

$$U(d_{n,s,g,\lambda}) = 2 \sqrt{u_{n,s,g,\lambda}^2 - u^2(y_{s,g,\lambda})}$$

On average, 95 % of the data fulfill $|d| < U$, if deviations $|d| > U$ are due to statistical variability. If more than 5 % of the data show $|d| > U$ or if, on the other hand, mainly $|d| \ll U$, this indicates that the data are either inconsistent or that the standard uncertainties are significantly overestimated.

Results and Discussion

Comparison of NMI instrumentation

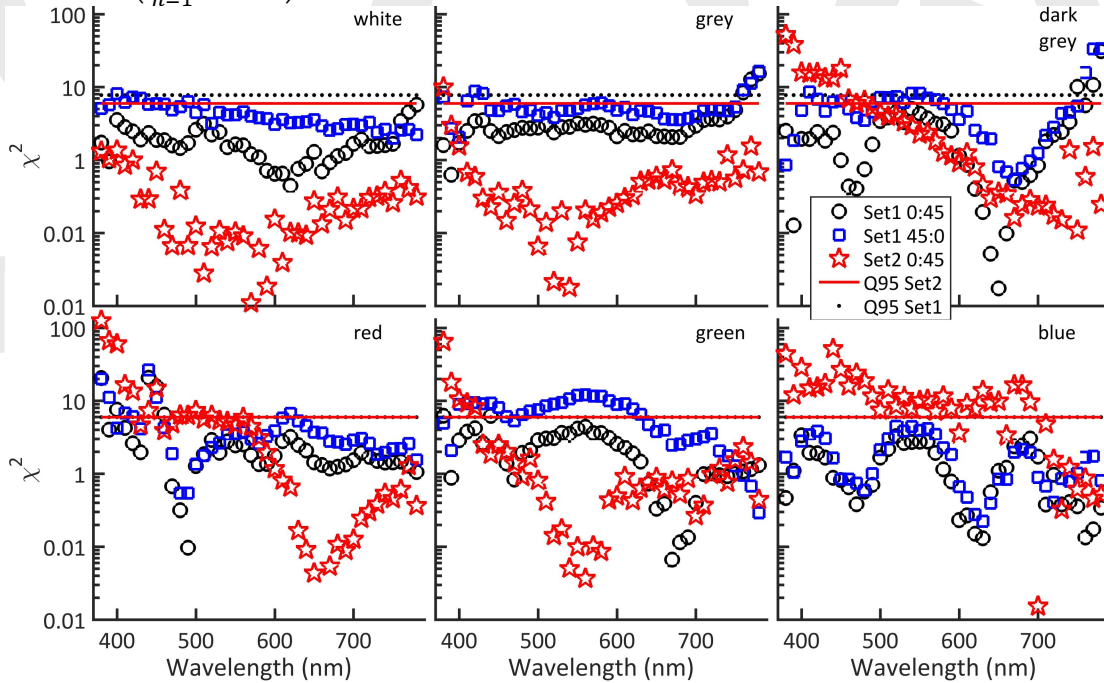


Figure 3 Overview of the χ^2 values. Q95 denotes the 95% quantile of the χ^2 distribution. This quantile is different for Set 1 and Set 2 with the grey scale samples because the degrees of freedom are different (4 vs 3 institutes, respectively).

Figure 3 depicts the χ^2 values for both sample sets and both geometries. The horizontal lines point out the value of the 95 % quantile of the χ^2 distribution with a degree of freedom that corresponds to the data. For consistent measurements an average of 95% of the χ^2 values should lie below the respective quantile limit, which equates to two data points per curve lying above it. As can be observed, the results for the green sample of set 1 in 45:0 geometry, and the blue sample of set 2 (0:45 geometry) are inconsistent over the greater part of the wavelength range. In all other cases, the consistency is either overall very good or otherwise the outliers above the 95 % quantile are not distributed spectrally in a random fashion but rather cluster in a certain wavelength range. Hence, the outliers cannot be solely attributed to the inherent statistical variation of the data but rather indicate a scale disagreement that has to be attributed to systematic effects.

For set 1, where the 0:45 geometry and its reciprocal 45:0 are measured, it can be ascertained that the χ^2 is generally smaller in the 0:45 geometry. This demonstrates that the assessment of the uncertainty contributions related to the angular positioning are not yet ideal for all participants. Otherwise, there would not be a systematic difference with respect to the measurement geometry. For instance, only PTB, MSL and CNAM take the angular positioning of the sample surface into account (cf. Table 3). It is also possible that a source of uncertainty related to the angular adjustment is formally taken into account but the magnitude of the effect may not be estimated precisely enough (cf. discussion in the next section).

The white ceramics samples exhibit the highest consistency under all measurement conditions. This test sample possesses a high spectral reflectance factor of the kind that is typical for white standards, which are probably the most often calibrated type of artefact. On the contrary, with the green sample of set 1 (45:0 geometry) and the blue sample of set 2, the measurements where the χ^2 test fails over most of the spectrum are both colored samples, which are less commonly calibrated.

Degree of equivalence

In Figure 4 – Figure 6, an overview of the degree of equivalence as the deviation from the weighted mean relative to the deviations uncertainty for all comparisons is given. The absolute uncertainties of the participants' measurements and therefore the weighting used to determine the reference values vary with reflectance, with wavelength and with sample type. An overview of the participants' standard uncertainties is given in the

appendix (Figure 10 and Figure 11). In addition, the SRF data and the corresponding standard uncertainty of every measurement can be found in the supporting information (Data File 1).

Apart from the instances in Figure 3 where the χ^2 lies above the 95% quantile and where simultaneously $|d/U| > 1$, the fact that most of the d/U values are scattered between -1 and 1 corroborates that each instrument is generally supported by a reasonable uncertainty budget, and substantiates the quality of measurements obtained by these instruments.

However, Figure 4 - Figure 6 identify areas for improvement for all NMIs. The results of set 1 in Figure 4 and Figure 5 indicate that it is meaningful for PTB and MSL to refine their uncertainty evaluation for wavelengths below 450 nm. A possible explanation for PTB's results above $|d/U| > 1$ for darker samples can be the use of a fast measurement mode of the used picoammeter, a mode which is usually not selected in standard calibrations.

For CSIC the discrepancies that occur are located between 450 nm and 650 nm and require further work. The disagreement between MIKES and PTB between 750 nm and 790 nm in four instances in Figure 4 and Figure 5 and the fact that PTB agrees well with CNAM and SP between 750 nm and 790 nm in set 2 (Figure 6) makes it plausible that MIKES needs to revise their uncertainty evaluation between 750 nm and 790 nm. Furthermore, the discrepancies in set 2 (Figure 6) suggest that CNAM has to improve their setup, particularly (but not only) for wavelengths below 450 nm and that also SP should assess the significance of missing uncertainty contributions in that spectral range.

That discrepancies can be observed in Figure 4 - Figure 6 and that we identified spectral parts where improvement is necessary is understandable considering that a look at Table 3 reveals that none of the participants uncertainty evaluations takes every contribution into account. In fact, two possible sources of uncertainty, the inhomogeneity of the sample and fluorescence are not regarded by any participant. Sample inhomogeneity becomes increasingly important the smaller the probed surface area is, so that the highest impact of this effect here can be expected for CSIC. Fluorescence is known to occur in some ceramic color standards [27], yet it was impossible to quantify the effect in this study. The (potential) effect of fluorescence is different for monochromatic illumination and monochromatic detection and can be generally avoided if monochromatic radiation is used in the illumination and detection system. For monochromatic illumination a higher signal will be measured at wavelengths where the

sample absorbs energy if the fluorescence at longer wavelengths is not filtered out before the photodetector. If the illumination is broadband the measurement signal at fluorescence wavelengths can be higher if the broadband illumination contains suitable wavelengths to excite the sample.

The effects of angular misalignments, finite solid angles of detection and observation and the variation of the solid angles across the measured sample area [9] are not always taken into account. They do require a lot of effort to characterize. For instance, to determine the effect of an angular misalignment, all possible sources of misalignment (including sample topography) need to be characterized [28] and the angular dependence of the SRF around the measurement geometry has to be determined. However, if possible, this should definitely be carried out because the angular dependence of the SRF, even on diffusely reflecting samples, can be very pronounced [29].

The polarization dependence of the detection unit and of the reflection on the sample are also not easy to account for because it generally requires a modification of the measurement setup. Though it is very meaningful to do so, this effect is currently only taken into account by two of the six participating NMIs. Of these two, PTB did not previously take this effect into account but updated the uncertainty evaluation in the course of this study.

Upon reflection the polarization state of the illumination is generally changed and unpolarized light becomes (partially) polarized and polarized light becomes (partially) unpolarized. In other words, the spectral radiance factor for s and p polarized illumination is generally different. For the white sample this effect is negligible but especially the colored ceramics reflect s and p polarized irradiation differently, leading to a degree of polarization of $\sqrt{S_1^2 + S_2^2 + S_3^2}/S_0$, where S_0, S_1, S_2, S_3 are the Stokes parameters, of up to 20 %.

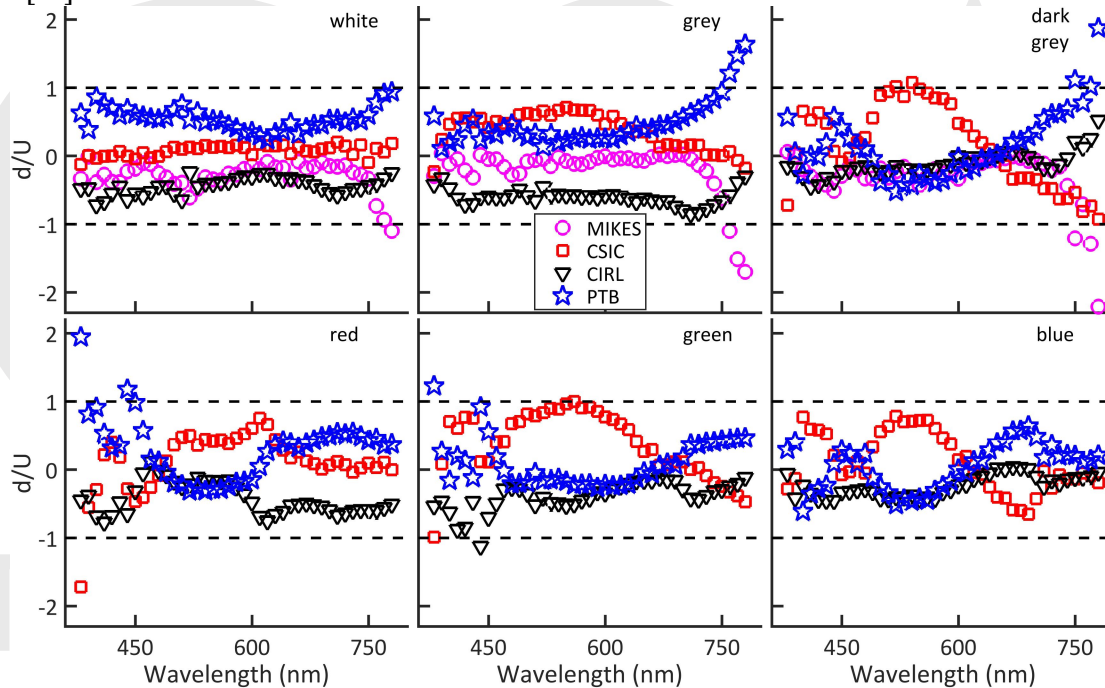


Figure 4 Overview of the degree of equivalence for the samples of set 1 in 0:45 geometry.

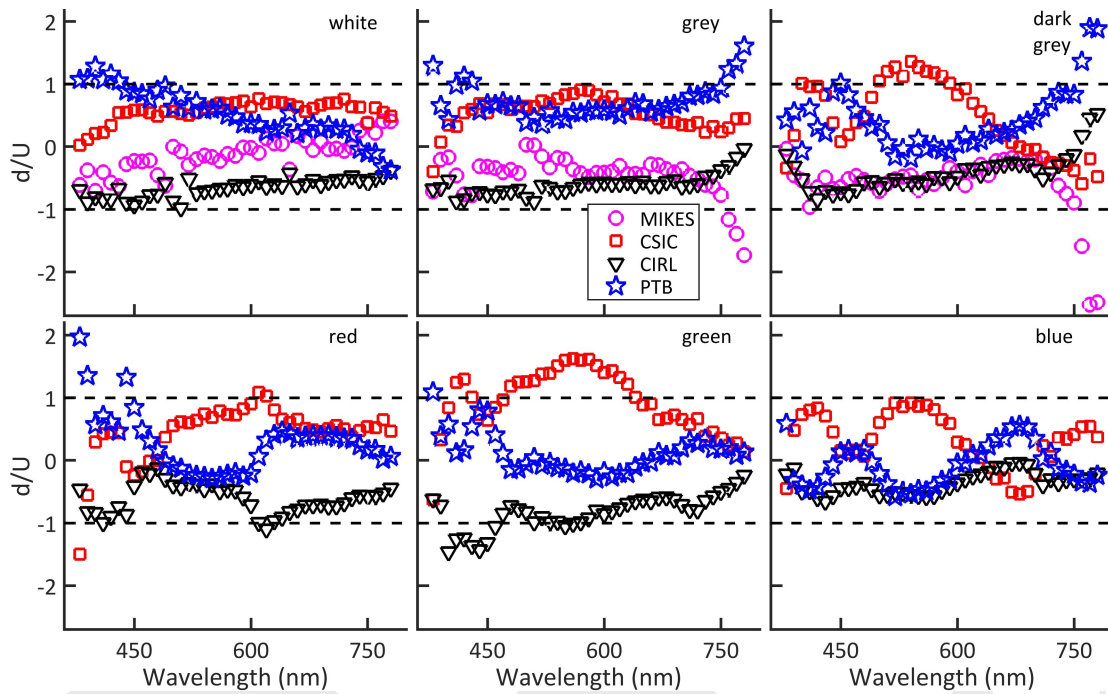


Figure 5 Overview of the degree of equivalence for the samples of set 1 in 45:0 geometry.

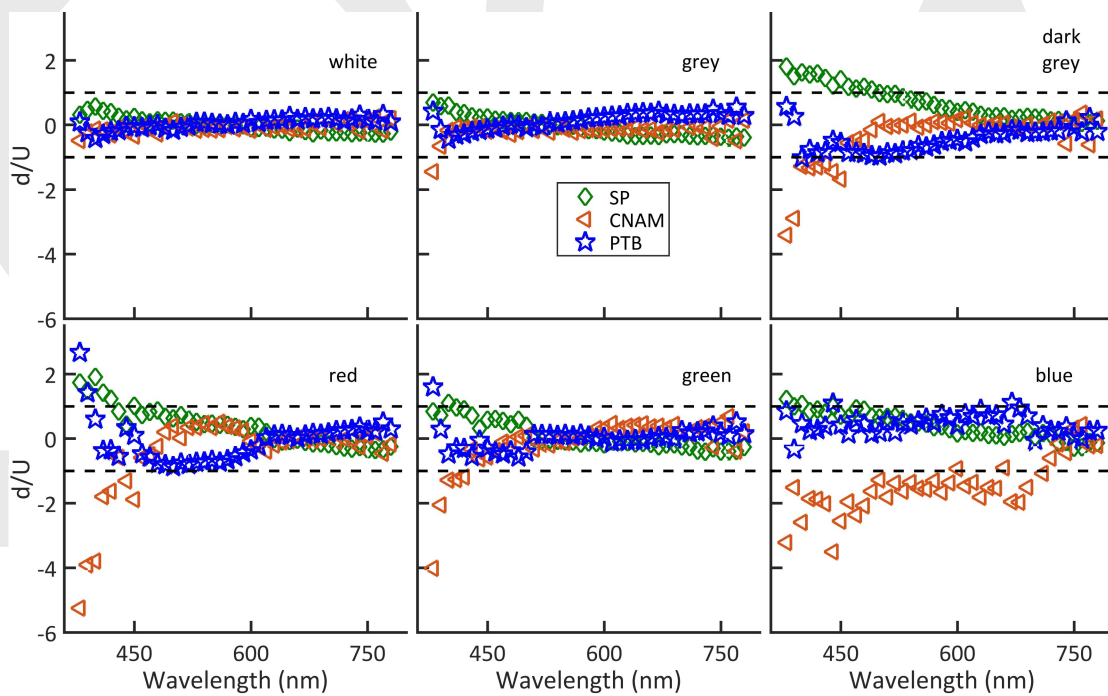


Figure 6 Overview of the degree of equivalence for the samples of set 2 in 0:45 geometry.

Non-NMI measurement results

Besides the measurements performed by instrumentation constructed by NMIs, also an investigation of instruments available at non-NMIs was considered in the study. Two widespread commercial

multi-angle spectrophotometers, namely the BYK-mac and the MA98, were provided by the University of Alicante, while KU Leuven performed the measurements with a home-built device.

The results obtained with the two commercial instruments are assessed in Figure 7 and compared to

the results of PTB. The black lines show twice of PTB's standard uncertainty, and hence, describe the 95 % confidence interval for the SRF. The results of the commercial instruments lie mostly outside this interval and the deviation is as large as 10 % in some instances. From the results it is not possible to determine which instrument performs better. First of all, both instruments seem to yield a closer correspondence for a part of the samples under test. Second, it is not possible to

generalize the results based on measurements made on just one device of each type. Apart from that, the special geometries in commercial instruments (small distance between sample and detector at fairly big measuring spots) mean that it is not obvious if the estimated measurand is the SRF or only a relative reflectance factor, affected by geometrical convolutions.

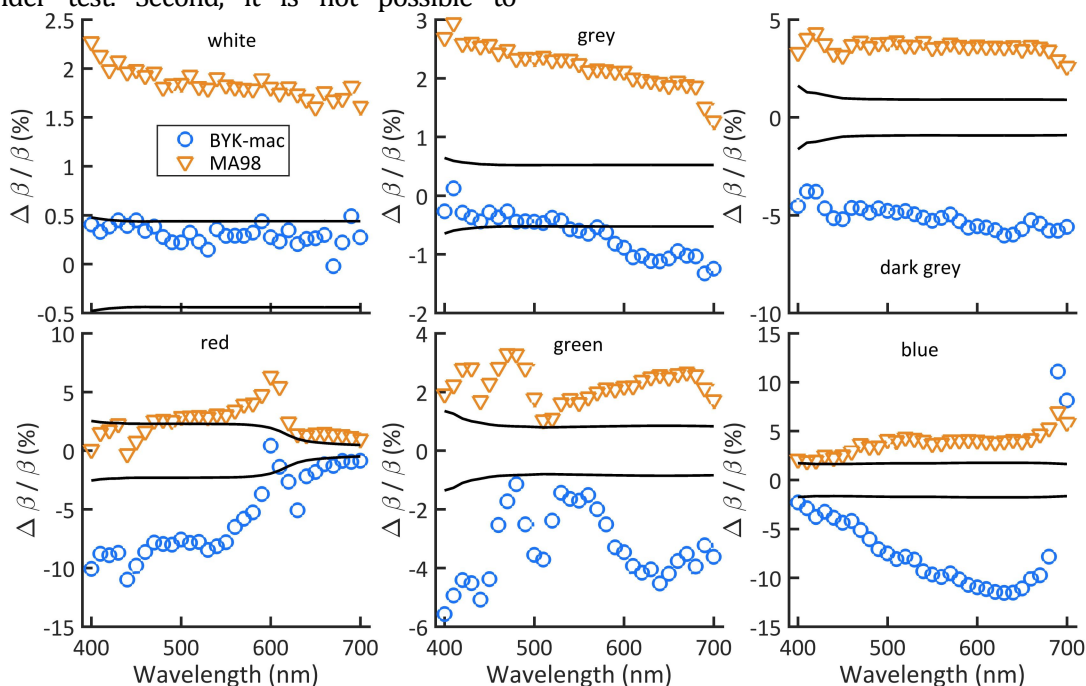


Figure 7 Relative Deviation $\Delta\beta/\beta$ of the commercial instruments from the results of PTB in 45:0 geometry. The black lines show twice of PTB's relative standard uncertainty (95 % confidence interval).

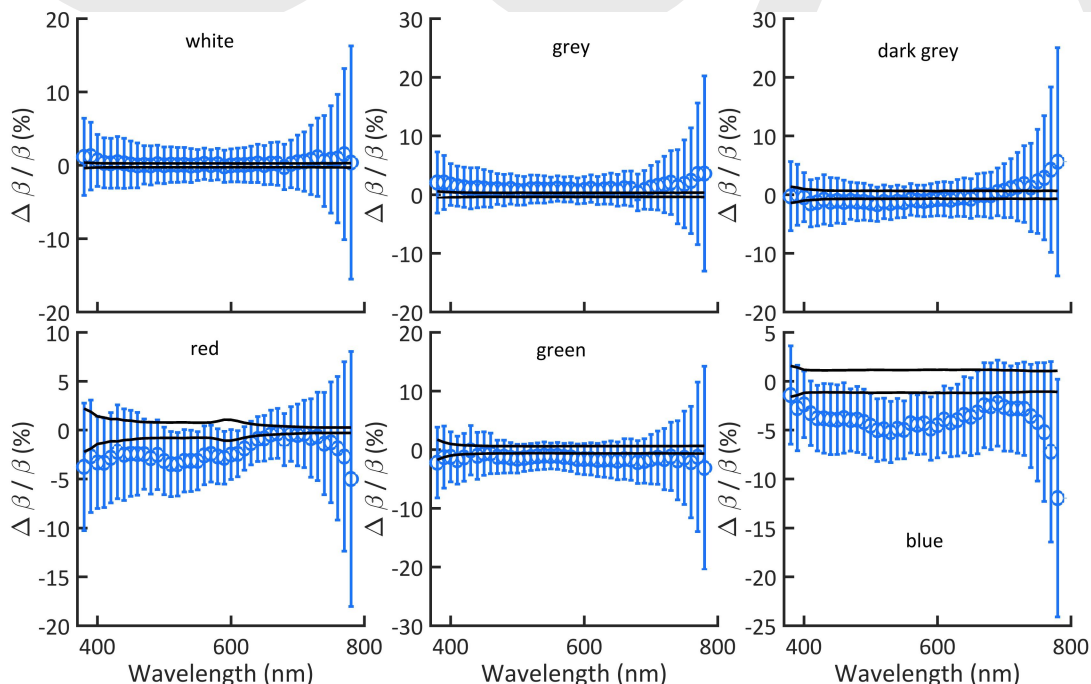


Figure 8 Relative deviation of KU Leuven's results from the weighted mean of the other participants of set 1 in 0:45 geometry. The black lines show twice the weighted mean's uncertainty and the error bars show twice of KU Leuven's relative standard uncertainty (95 % confidence intervals).

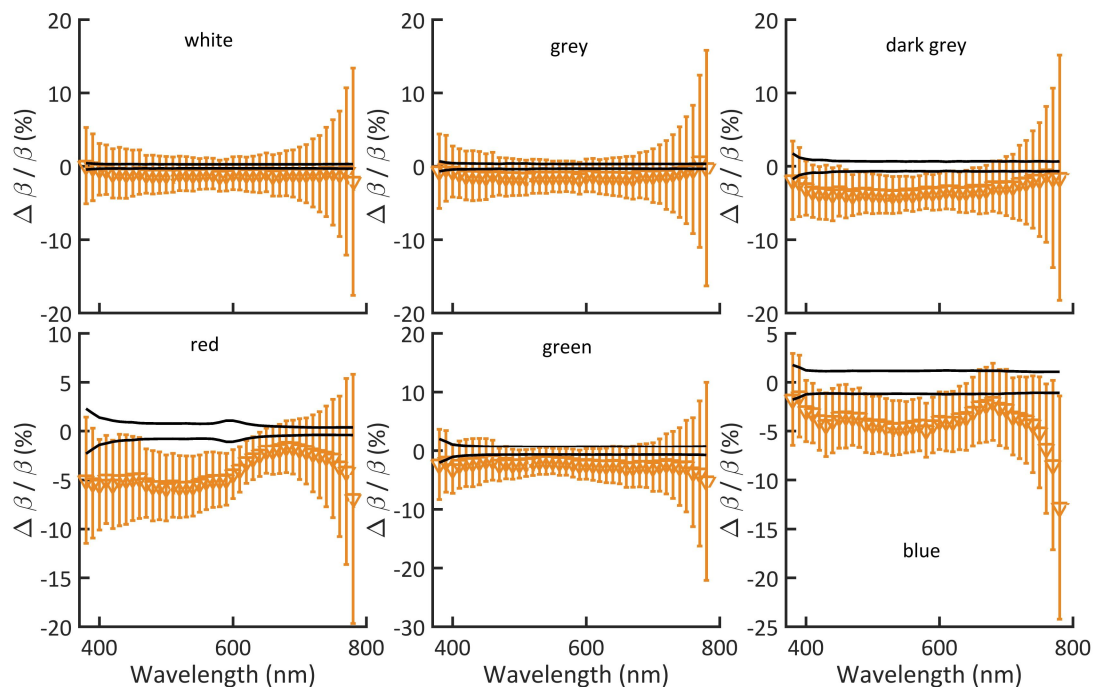


Figure 9 Relative deviation of KU Leuven's results from the weighted mean of the other participants of set1 in 45:0 geometry. The black lines show twice the weighted mean's uncertainty and the error bars show twice of KU Leuven's relative standard uncertainty (95 % confidence intervals).

The results obtained by KU Leuven are compared to the weighted mean of the NMI results for set 1 in 0:45 geometry (Figure 8) and 45:0 geometry (Figure 9)). As expected the uncertainties estimated by KU Leuven are much larger than those of the NMIs (PTB, MIKES, MSL, and CSIC). Especially for larger wavelengths the standard uncertainties increase due to a reduced signal. This uncertainty contribution could be diminished by application of a light source with enhanced power beyond 700 nm. A look at Table 3 shows that there are further areas for improvement with the uncertainty budget as 9 possible uncertainty contributions are not yet included at all. Taking into account the uncertainties, the results agree with the NMI weighted mean over the entire wavelength range in 8 out of the 12 occasions. Deviations from the weighted mean generally seem to be larger for the results of measurements performed in the 45:0 geometry. This is in accordance with the observation of a higher χ^2 in 45:0 geometry in the comparison (Figure 3). It can very likely be attributed to the missing uncertainty evaluation of angular positioning and the smaller measurement signal in 45:0 than in 0:45 geometry.

Conclusion

This multilateral scale comparison for SRF measurement facilities shows that the scales of the participating NMI's are mostly consistent and partly validates the SRF measurements of the involved NMIs at the standard geometries of 0:45 and 45:0. However, the comparison also reveals some problems that need to be addressed. Particularly at the short and the long-wavelength side of the probed spectrum from 380 nm – 780 nm scale discrepancies were identified. This result is in accordance with the fact that no participant does so far take all sources of error into account. For future work all participants of this scale comparison need to rework, respectively, include as many of the uncertainty contributions as possible that are up to now disregarded. The geometries 0:45 and 45:0 yield the same SRF according to the Helmholtz reciprocity principle. Yet, it was found that the uncertainties are better modeled in 0:45 geometry. Though the geometries in this comparison are standard geometries with probably the highest significance for industry, measurements at non-standard out-of-plane geometries (out of the incidence

plane) should be included in a future comparison, since they require the positioning of the sample with more than one degree of freedom, which introduces additional contributions to the uncertainty and thus a more difficult criterion for the comparison. Particularly interesting are geometries where detection or illumination are very slanted with respect to the sample or geometries with a high angular variation of the SRF, which for white standards are typically geometries close to reflection or retro-reflection [29]. These are best suited to reveal systematic errors of angular positioning (notice the existence of a cosine in the denominator of the measurement equation) as well as those due to the finite apertures.

In order to assess if polarization effects are properly considered in the uncertainty evaluation of a future comparison a sample with a very pronounced polarization dependence of the radiance factor should be included, though at this point it is not clear which type of non-specularly reflecting material and geometry is best suited for this task.

In addition, we considered results obtained on two commercial instruments and found deviations in the measured SRF of up to more than 10 %. A comparison of KU Leuven's result to NMI results reveals that, as expected, the universities uncertainty budget is larger than those of NMIs. A reduction of the uncertainty contributions could be envisaged, e.g. by applying a light source with enhanced signal in the longer wavelength range.

Appendix: Overview of the uncertainties

Figure 10 and Figure 11 give an overview on the standard uncertainties of the institutes with a GUM compatible uncertainty evaluation [16]. For PTB, only the resulting uncertainties from the measurements with set 1 are depicted as those for set 2 are practically the same. The complete information about the SRF data and the corresponding uncertainty can be found in the supporting information (Data File 1).

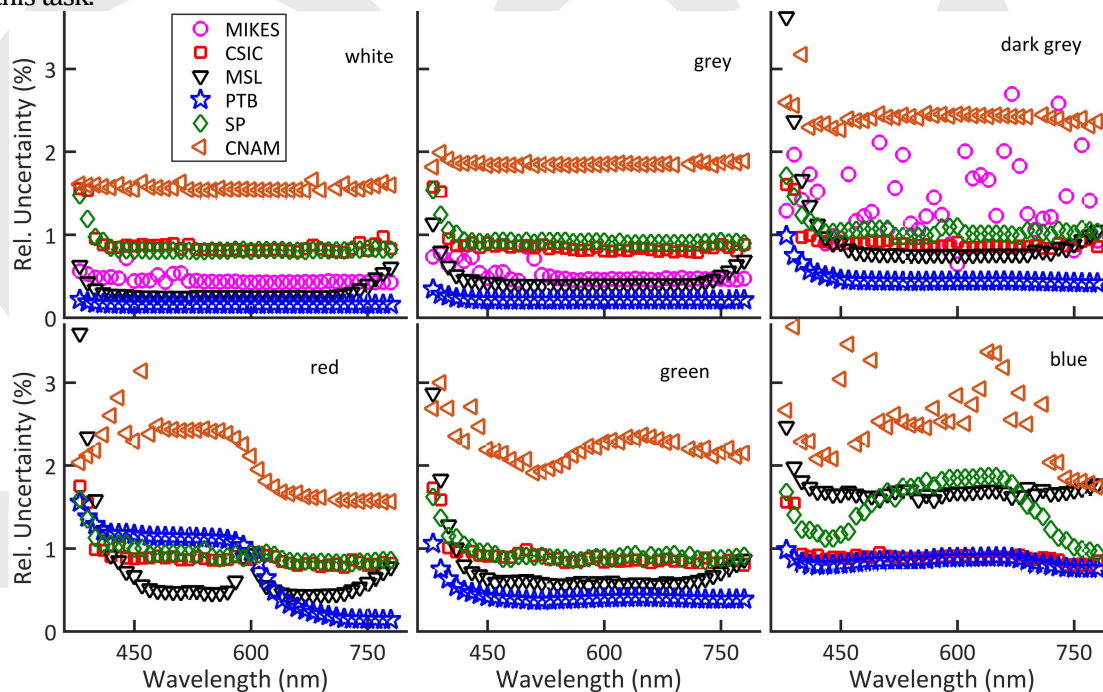


Figure 10 Uncertainties relative to the weighted mean of set 1 (MIKES, CSIC, MSL, PTB) and set 2 (SP, CNAM) in 0:45 geometry.

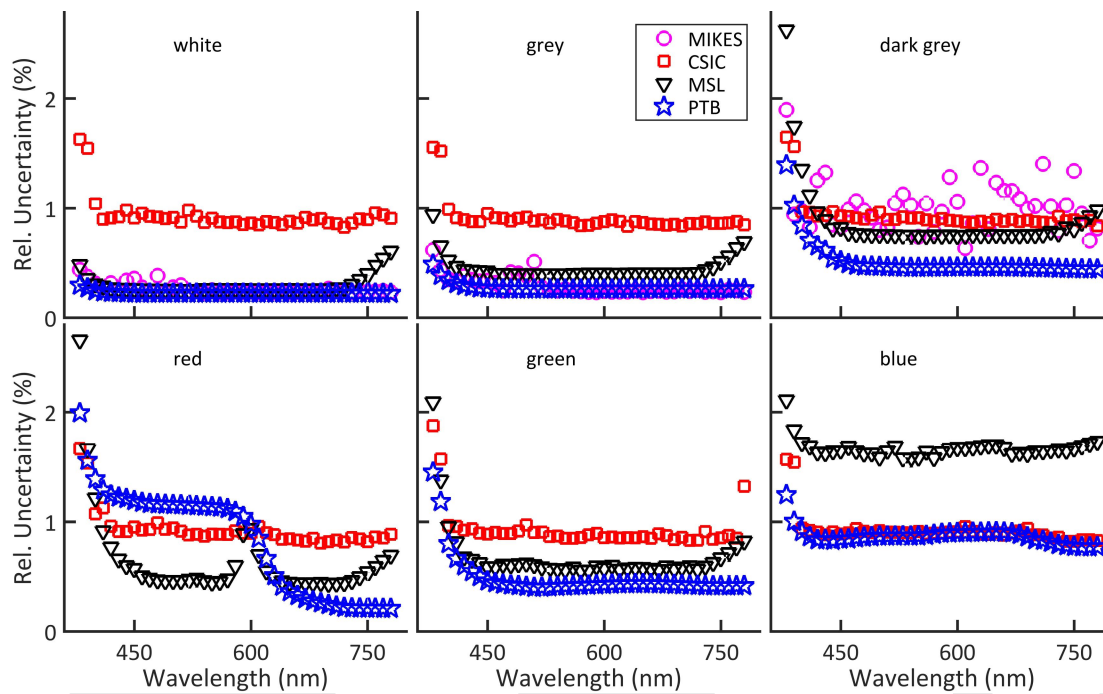


Figure 11 Uncertainties relative to the weighted mean of set 1 in 45:0 geometry.

Funding Information. This work has been realized within the European Metrology Research Program (EMRP) project JRP-IND52. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union. We thank the European commission and EURAMET e.V. for financial support.

Supporting Information CSV file (Data File 1) containing the measurement results of all participants, i.e. the spectral radiance factor and its standard uncertainty.

References

1. The spectral radiance factor is equivalent to the bidirectional distribution function, the only difference being a factor of π .
2. A. Höpe, K.-O. Hauer, P. Bergner, and T. Ziegler, "BRDF measurements at 254 nm for the LISA Pathfinder satellite mission," in Proceedings of NEWRAD, S. Park, P. Kärhä, E. Ikonen, ed. (NEWRAD, 2014), pp. 326-329.
3. W. Gurlit, H. Bösch, H. Bovensmann, J. P. Burrows, A. Butz, C. Camy-Peyret, M. Dorf, K. Gerilowski, A. Lindner, S. Noël, U. Platt, F. Weidner, and K. Pfeilsticker, "The UV-A and visible solar irradiance spectrum: inter-comparison of absolutely calibrated, spectrally medium resolution solar irradiance spectra from balloon- and satellite-borne measurements," *Atmos. Chem. Phys.* **5**, 1879 – 1890 (2005).
4. D. R. Taubert, J. Hollandt, P. Sperfeld, S. Pape, A. Höpe, K.-O. Hauer, P. Gege, T. Schwarzmaier, K. Lenhard, and A. Baumgartner, "Providing radiometric traceability for the Calibration Home Base of DLR by PTB," in AIP Conf. Proc. 1531, (AIP, 2013), pp. 376-379.
5. A. Höpe, A. Koo, F. M. Verdú, F. B. Leloup, G. Obein, G. Wübbeler, J. Campos, P. Iacomussi, P. Jaanson, S. Källberg, and M. Šmíd, "‘Multidimensional reflectometry for industry’ (xD-Reflect) a European research project," *Proc. SPIE* **9018**, 901804 (2014).
6. G. Pfaff, *Special Effect Pigments* (Vincentz Network GmbH & Co KG, 2008).
7. A. Höpe, D. Hünerhoff, and K.-O. Hauer, "Robot-based gonireflectometer," in *Industrial Robotics: Programming, Simulation, and Applications*, K. H. Low, ed. (Mammendorf: Pro Literatur-Verlag, 2007), pp. 623-632.
8. D. Hünerhoff, U. Grusemann, and A. Höpe, "New robot-based gonireflectometer for measuring spectral diffuse reflection," *Metrologia* **43**, S11 (2006).
9. G. Obein, R. Bousquet, and M. E. Nadal, "New NIST Reference Goniospectrometer," *Proc. SPIE* **5880**, 58800T (2005).
10. M. E. Nadal, E. A. Early, W. Weber, and R. Bousquet, "NIST 0:45 Reflectometer," *Color Res. Appl.* **33**, 94-99 (2008).
11. A. M. Rabal, A. Ferrero, J. Campos, J. L. Fontecha, A. Pons, A. Rubiñoo, and A. Corróns, "Automatic gonio-spectrophotometer for the absolute measurement of the spectral BRDF at in- out-of-plane and retroreflection geometries," *Metrologia* **49**, 213–223 (2012).
12. B. Bernad, A. Ferrero, A. Pons, M. L. Hernanz, and J. Campos, "Upgrade of goniospectrophotometer GEFE for near-field scattering and fluorescence radiance measurements," *Proc. SPIE* **9398**, 93980E (2015).
13. F. B. Leloup, S. Forment, P. Dutré, M. R. Pointer, and P. Hanselaer, "Design of an instrument for measuring the spectral bidirectional scatter distribution function," *Appl. Opt.* **47**, 5454-5467 (2008).
14. DIN 6175-2:2001-03, *Tolerances for automotive paints - Part 2: Goniochromatic paints*, (Beuth Verlag GmbH, 2001)
15. ASTM E2539 – 08, *Standard Practice for Multiangle Color Measurement of Interference Pigments*, (ASTM International, 2008)
16. JCGM 100:2008, "Evaluation of measurement data – Guide to the expression of uncertainty in measurement" by the International Bureau of Weights and Measures

-
17. M. Nadal, K. L. Eckerle, E. A. Early, Y. Ohno, "Final Report On The Key Comparison CCPR-K5: Spectral Diffuse Reflectance," *Metrologia* **50**, 02003 (2013)
 18. C. C. Cooksey, M. E. Nadal, D. W. Allen, K. O. Hauer, and A. Höpe, "Bidirectional reflectance scale comparison between NIST and PTB," *Appl. Opt.* **54**, 4006-4015 (2015)
 19. CIE Technical Report 152004, "Colorimetry, 3rd Edition," (International Commission on Illumination, 2004).
 20. Lucideon ceramic standards are examples of suitable products available commercially. This information is given for convenience and does not constitute an endorsement.
 21. P. Y. Barnes, E. A. Early, and A. C. Parr, "Spectral Reflectance," NIST Special Publication 250-48 (U.S. Department of Commerce, 1998).
 22. K.-O. Hauer and A. Höpe, "High-Grade Uniform Light Source for Radiometric and Photometric Applications," *MAPAN – J. Metrology Soc. India* **24**, 175 (2009).
 23. R. Le Breton, G. Ged, and G. Obein, "Out of plane BRDF measurement at LNE-Cnam using 'ConDOR' our primary goniospectrophotometer," in 28th CIE SESSION, 2015, Manchester, United Kingdom, CIE 216:2015, pp. 1401 – 1407.
 24. S. Nevas, F. Manoocheri and E. Ikonen, "Gonioreflectometer for measuring spectral diffuse reflectance," *Appl. Opt.* **43**, 6391–9 (2004).
 25. M.G. Cox, „The evaluation of key comparison data," *Metrologia* **39**, 589-595 (2002).
 26. The results of MIKES and SP are correlated because SP's reference standard was calibrated by MIKES. However, the measurements of both institutes are not directly compared.
 27. A. Koo, J. F. Clare, K. M. Nield, A. Deadman, and E. Usadi, "Fluorescence of ceramic color standards," *Appl. Opt.* **49**, 2379 – 2381 (2019).
 28. C. J. Chunnillal, F. J. J. Clarke, M. J. Shaw, "Developments in the realization of diffuse reflectance scales at NPL," *Proc. SPIE* **5880**, 58800V (2005).
 29. A. Höpe, and K.-O. Hauer, „Three-dimensional appearance characterization of diffuse reflection materials," *Metrologia* **47**, 295-304 (2010).

Published by
OSA®
The Optical Society