Leroy, Vincent; Aschan, Robin; Woolliams, Peter; Schunke, Sebastian; Manoocheri, Farshid; Govaerts, Yves

An SI-traceable protocol for the validation of radiative transfer model-based reflectance simulation

Published in:
IEEE Transactions on Geoscience and Remote Sensing

DOI:
10.1109/TGRS.2025.3547305

Published: 01/01/2025

Document Version
Publisher's PDF, also known as Version of record

An SI-Traceable Protocol for the Validation of Radiative Transfer Model-Based Reflectance Simulation

Vincent Leroy, Robin Aschan, Peter Woolliams, *Member, IEEE*, Sebastian Schunke, Farshid Manoocheri, and Yves Govaerts

***Abstract*— Because of the critical role they play in Earth observation (EO) workflows, radiative transfer models (RTM) must undergo thorough validation campaigns that will help guarantee that their output is representative of the physical reality. In this study, we test the ability of the Eradiate RTM to simulate the reflectance of a manufactured target given SI-traceable measurements of its shape and optical properties. To address issues identified in similar prior work, we select a material with reflective features that can be accurately modeled using a straightforward data-driven approach. We produce an artificial target design that is easy to manufacture with good precision. The material and artifact are characterized optically using an SI-traceable 3-D goniospectrophotometer, and the artifact is characterized geometrically in an SI-traceable facility. The geometry of the artificial target and the optical properties of the material are used to build a numerical experiment with Eradiate that simulates the optical characterization of the artificial target. The simulation includes the propagation of uncertainties on the input parameters. We compare the simulated and measured data and analyze the performance of this method using three different metrics. Our findings demonstrate consistent performance across all considered illumination and sensor pointing configurations. Simulations deviate from measurements by less than 2%, and more than 80% of measured and simulated data points agree.**

***Index Terms*— Accuracy, bidirectional reflectance factor (BRF), data-driven model, model validation, Monte Carlo (MC) methods, radiative transfer (RT), ray tracing, SI traceability, uncertainty.**

I. Introduction

REMOTE sensing is a key component of Earth observation (EO), itself instrumental in many scientific and technical activities. Images recorded by radiometric instruments mounted onboard vehicles or at fixed locations on Earth are used to understand the dynamic of the Earth system and inform policy and management decisions.

Received 10 January 2024; revised 17 June 2024 and 23 December 2024; accepted 23 January 2025. Date of publication 3 March 2025; date of current version 17 March 2025. This work was supported by the Metrology for Earth Observation and Climate [19ENV07 MetEOC-4] and has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme. *(Corresponding author: Vincent Leroy.)*

Vincent Leroy, Sebastian Schunke, and Yves Govaerts are with Rayference, 1000 Brussels, Belgium (e-mail: vincent.leroy@rayference.eu; sebastian.schunke@rayference.eu; yves.govaerts@rayference.eu).

Robin Aschan and Farshid Manoocheri are with the Metrology Research Institute, Aalto University, 02150 Espoo, Finland (e-mail: robin.aschan@aalto.fi; farshid.manoocheri@aalto.fi).

Peter Woolliams is with the National Physical Laboratory, TW11 0LW Teddington, U.K. (e-mail: peter.woolliams@npl.co.uk).

Digital Object Identifier 10.1109/TGRS.2025.3547305

The simulation of radiative transfer (RT) in the Earth system is one of the core components of the modern remote sensing workflow. RT models (RTMs) are used for a wide variety of tasks, such as mission preparation, the application of atmospheric correction to remote sensing data products, the interpretation of satellite data through climate variable retrieval, or the vicarious calibration of satellite-borne radiometric instruments. In such cases, an RTM becomes a contributor to the final remote sensing product and contributes to its uncertainty budget.

To minimize the uncertainty introduced by RT modeling in a given process, assessing how accurately the RTM represents the physical reality is critical. Validation exercises consist of comparing simulated data to reference measurements, e.g., reference satellite images. In principle, both the reference measurement and the input used to produce a simulated experiment should be SI-traceable, i.e., comparable, in an unbroken chain of calibrations, with a reference used in the practical realization of an SI unit. This ensures the repeatability of the measurements with other SI-traceable instruments: if the RTM is validated against an SI-traceable reference, its output can be compared with the output of any similar SI-traceable measurement device.

One major issue is that the desired reference is sometimes not calibrated against an SI-traceable reference. For instance, many satellites do not have an onboard post-launch calibration system and are calibrated against an RTM, as part of a vicarious calibration process. Finding appropriate validation data is therefore a major challenge, which is however hoped to be addressed in the coming years by SI-traceable missions [1]. Another major issue is that the high complexity of the experiments simulated by remote sensing-oriented RTMs makes the gathering of SI-traceable input challenging: setting up fully SI-traceable validation campaigns against actual satellite images is currently hardly possible.

RTM performance assessment therefore has to go through different routes. One of them is intercomparison, a typical example of which is the radiation transfer model intercomparison (RAMI) benchmark series [2], [3], [4], [5], launched in 1999. The RAMI exercises define a set of standard scenes that are used to intercompare RTMs and identify a consensus on what the correct output should be—for the specific scenes considered. This approach can incorporate a lot of modeling complexity and does not require reference data.

Although a fully SI-traceable validation against satellite images is currently not possible, a simplified setup can be considered to assess how accurately an RTM represents the physical reality, but also to reveal the sources of inaccuracy in the model. In that context, validating components of an RTM against measured data obtained in a controlled environment, such as a laboratory, is of particular interest. This is the approach taken by two studies on which this work builds. Govaerts and Verstraete [6] designed an experiment in which they compared the radiometric records obtained by measuring the reflectance of an artificial target in a goniospectrophotometer with the equivalent signal simulated by the Raytran RTM. This approach was then extended by Jaanson et al. [7] to make the reference measurement and model input SI-traceable, and propagate uncertainty in simulation results. In this study, we extend further this approach, to significantly improve the accuracy of the model by addressing the modeling issues identified in previous work.

The major sources of numerical modeling uncertainty identified in those works pertained to choices made to model the scattering of light by the material selected to manufacture the artificial target: a parametric Torrance–Sparrow model was used, and fit to in-lab optical measurements. Achieving a good fit was very difficult and made the resulting material model inaccurate. We, therefore, take a different approach in this work and bring the following contributions toward the goal of better-reproducing laboratory measurements with a RTM.

1) A new artificial target design focusing on manufacturing and characterization simplicity, with predictable and strong reflective features outside the retro-reflective direction.
2) An updated numerical simulation setup using an SI-traceable, data-driven material optical reflection model that limits numerical modeling uncertainty.
3) A set of complementary metrics with different features that allow for a detailed analysis of the performance of the method.

The computer graphics community has been very active in the topic of designing material reflection models based on measured data and comparing their output to laboratory measurements (e.g., [8], [9], [10], [11], [12], [13]). Compared to these studies, our work differs on the following points.

1) The goal of this study is to validate the results produced by the full RT simulation chain, which includes the material reflection model, the light source and sensor models, and the RT equation (RTE) integration algorithm.[1]
2) Our approach factors in the uncertainty on many simulation parameters and provides an estimate of the uncertainty on the simulated radiometric records.
3) The output of the RTM is compared to SI-traceable data with a detailed uncertainty budget. The metrics used for comparison account for the availability of uncertainty in both the measured and simulated data.

[1]As emphasized later, it should be noted that the setup considered in this work focuses on surface scattering modeling and therefore excludes volume scattering.

After presenting the general methodology (see Section II-A), we introduce the simulation and optical characterization tools used in this work (see Sections II-B and II-C). We review the artificial target design and manufacturing process (see Section II-D), the characterization techniques and measurement methodologies (see Sections II-E and II-F), and the numerical simulation protocol (see Sections II-G and II-H). We conclude the material and methods presentation with our comparison methodology (see Section II-I). The result analysis features a specific focus on the optical characterization of the material used to manufacture the artificial target (see Section III-A), the direct input to the data-driven material reflection model, then turns to the comparison of artificial target measurement and simulation results (see Section III-B).

II. Material and Methods

A. General Methodology

The fundamental goal of this study is to compare the reflectance of an artifact measured by a goniospectrophotometer to a simulation of that measurement by an RTM. This simulation uses as an input a digital representation of the artifacts geometric and optical properties, the light source, and the sensor. All digital representations are based on models parametrized by a number of scalar parameters.

Good metrological practices require accounting for uncertainty when comparing measurements—regardless of whether they are "actual" or simulated measurements. We therefore attach uncertainty to the measured artifact reflectance and propagate the uncertainty on parameters through the simulation to estimate the uncertainty on the simulation output.

We distinguish between 1) the simulation parameters inherent to the measurement facility (e.g., sensor field of view, light source divergence), known from the characterization of the measurement device and referred to in the following as facility parameters and 2) those related with the artifact (e.g., geometric and optical properties), measured for this study and referred to in the following as artifact parameters.

Jaanson et al. [7] primarily attributed the differences between their measured and simulated data to their material modeling choice: the Torrance–Sparrow BRDF they implemented to model material reflectance takes input parameters that are difficult to retrieve experimentally and, therefore, are obtained through a fit against experimental data. Achieving good fit quality proved to be challenging and resulted in large modeling uncertainties. To eliminate important sources of bias and uncertainty, our protocol differs from that of Jaanson et al.

1) Target shape design is part of the study, and the manufacturing process can be tailored for an accurate and easy dimensional measurement.
2) Material choice is part of the study and can be oriented to facilitate an accurate digital representation.
3) The material model consists of a data-driven BRDF model instead of a fit material model.

The comparison protocol is based on three metrics and accounts for uncertainty. Fig. 1 summarizes the general methodology of the study.

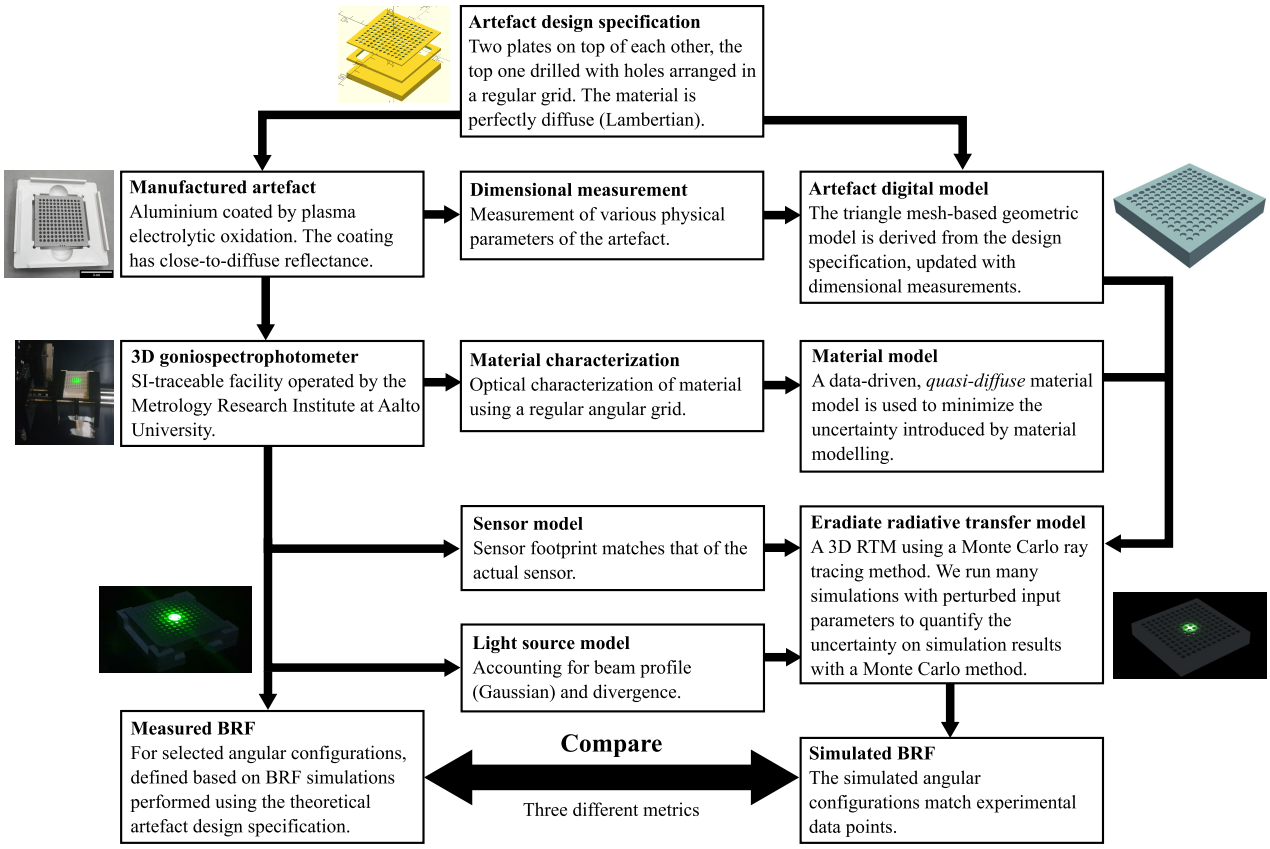


Fig. 1. General methodology applied in this study.

About Radiometric Quantities: Throughout this article, two radiometric quantities are used: a) the bidirectional reflection distribution function (BRDF) and b) the bidirectional reflectance factor (BRF).

We follow the definitions of Nicodemus et al. [14]: the BRF is the BRDF normalized by the BRDF of a diffuse (Lambertian), perfectly reflective surface. Consequently, the two are related by

$$\mathrm{BRF} = \pi \times \mathrm{BRDF}.$$

From this follows that relative uncertainties on a BRDF can be transferred to the corresponding BRF.

The BRDF is used in parts of this article that discuss optical measurements, while the BRF, more widely used in the EO community, is used to discuss simulation results. While both quantities can be qualitatively interpreted similarly, distinguishing between the two is of prime importance when performing quantitative comparisons.

In-Plane and Out-of-Plane Angular Domains: In the following, frequent references to *in-plane* and *out-of-plane* angular domains are made. *In-plane* designates the set of directions that are contained in a plane generated by the illumination direction and the local normal at the reflecting surface; *out-of-plane* designates the rest of the hemisphere.

In practice, if the illumination direction is parametrized by zenith and azimuth (θ_i, ϕ_i), the in-plane domain is $\{(\theta, \phi) \mid \phi = \phi_i \text{ or } \phi_i + \pi\}$, and the out-of-plane domain is $\{(\theta, \phi) \mid \phi \neq \phi_i \text{ nor } \phi_i + \pi\}$.

B. Eradiate RTM

Eradiate [15] is an open-source 3-D RTM designed to support remote sensing applications. It focuses on delivering highly accurate radiance estimates using Monte Carlo ray tracing (MCRT) methods. It uses a modified version of the Mitsuba 3 rendering system [16] as its radiometric kernel. MCRT methods are becoming increasingly popular because of their ability to solve efficiently highly dimensional problems such as RTE integration. Eradiate can simulate both surface and volume scattering processes. Eradiate is used in this study to design the artificial target (see Section II-D, where Eradiate is used to predict the theoretical optical behavior of candidate designs), and it is the test subject in the RTM validation process (see Section II-G).

A critical point of concern raised by previous validation attempts [6], [7] is surface material modeling: these works used a fit Torrance–Sparrow model [17], which had a limited ability to accurately represent the reflective behavior of the material the measured object was made of, and therefore introduced significant uncertainty in the simulation. To address this issue, we use a data-driven model to represent the material BRDF. The choice of a simple parametrization for this model is motivated by the characteristics of the SI-traceable goniospectrophotometer used for optical measurements and

the need to perform RT simulations without access to material data during the artificial target design activity.

A more detailed overview of Eradiate is provided in Appendix A.

C. Goniospectrophotometer

BRDF measurements are performed using a 3-D goniospectrophotometer set at Aalto University. The instrument is operated in relative mode, and it is traceable to the absolute goniospectrophotometer in Aalto [18] via a calibrated reflectance standard. The goniospectrophotometer is capable of out-of-plane measurements, and it accommodates horizontal samples. Measuring a sample horizontally is useful for volatile samples, such as sand and liquids. The instrument has been used to determine the BRDF of desert sand [19], and it has been used for determining out-of-plane BTDF of various transmissive diffuser samples [20]. The facility and its use for BRDF and BTDF measurements are described in detail in prior work [20], [21], and we provide here a brief description of the instrument for out-of-plane BRDF measurements of an artificial reflectance standard.

Setup: The setup (see Fig. 2) includes a set of three motorized arms: one arm controls the incident zenith angle of the (1) illumination, and the two remaining arms control the viewing (2) zenith and (3) azimuth angles for sampling the BRDF in any point of the hemisphere above the sample under test, except for angles shadowed by the detector $\pm 5^\circ$ around the detector head. Samples are accommodated in a horizontally aligned sample (4) holder that can be adjusted laterally by a micrometer for accommodating variable thickness samples. The beam at the sample is controlled by (5) the illumination optics, which include a collimator, a polarizer, an optical filter for laser harmonics, a beam expander, steering optics, and an iris. The scattered signal is collected by a 10.8 mm diameter off-axis (6) parabolic mirror, which sets the solid angle of the system. The detected scattering is forwarded by fiber coupling to a silicon detector housed under the instrument. The detected signal is processed and stored by a script written in LabVIEW's G dataflow language [22] and executed on a PC.

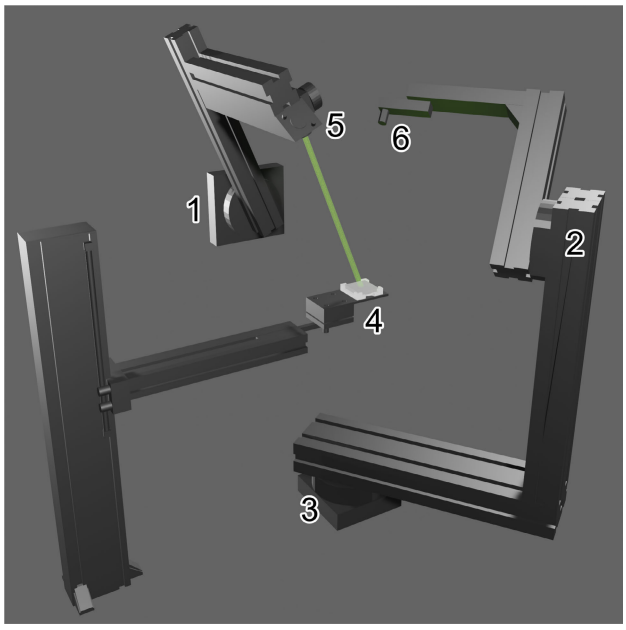


Fig. 2. Schematic of the 3-D goniospectrophotometer in its configuration for BRDF measurements. The sample (4) is illuminated by the light source (5) controlled by a robotic arm (1), and the scattered light is collected by the sensor optics (6), controlled by two arms (2, 3).

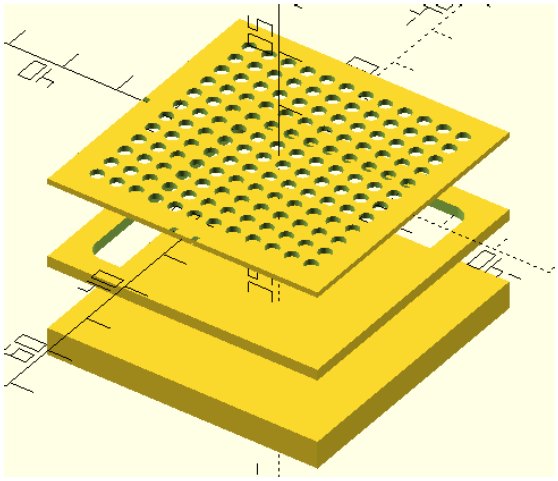


Fig. 3. Exploded view of the final design, created with the OpenSCAD modeler [23].

D. Artifact Design and Manufacturing

Following basic principles established during previous validation efforts [6], [7], the artificial target used in this study was designed to exhibit strong shadowing-masking effects that can only be simulated accurately with a 3-D RTM. We also required a distinctive reflective signature in the angular domain, with several reflective lobes accessible to the measurement facility, so that even a sparse coverage of the measurement space would allow capturing critical features that are useful for validation.

To achieve that, the final shape of the artifact was designed with the aid of iterative RT simulations: this way, we could ensure that the final design would feature a reflective pattern that would meet our requirements. At each iteration, we simulated with Eradiate (see Section II-G for details on the simulation setup) the reflective signature of a candidate shape under simplified but "reasonable" assumptions. If the reflective pattern matched the requirements, the design was accepted; otherwise, a new candidate shape was proposed and simulated.

In the previous paragraph, "reasonable" means that manufacturing an object with geometrical and optical properties close to the assumptions was achievable, and, consequently, that the manufactured artifact would exhibit a reflective pattern similar to that of the theoretical design. The final design, shown in Fig. 3, has a theoretical reflective pattern that features a main retro-reflective lobe, as well as multiple secondary peaks due to shadowing-masking effects (see Fig. 4).

The artifact was manufactured using aluminum alloy coated by plasma electrolytic oxidation (PEO) based on the final design model (see Fig. 5). Interested readers are referred to

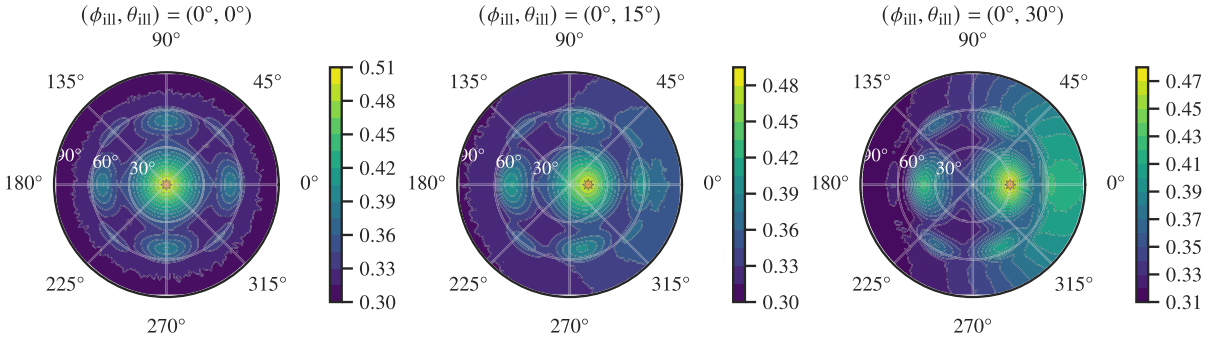
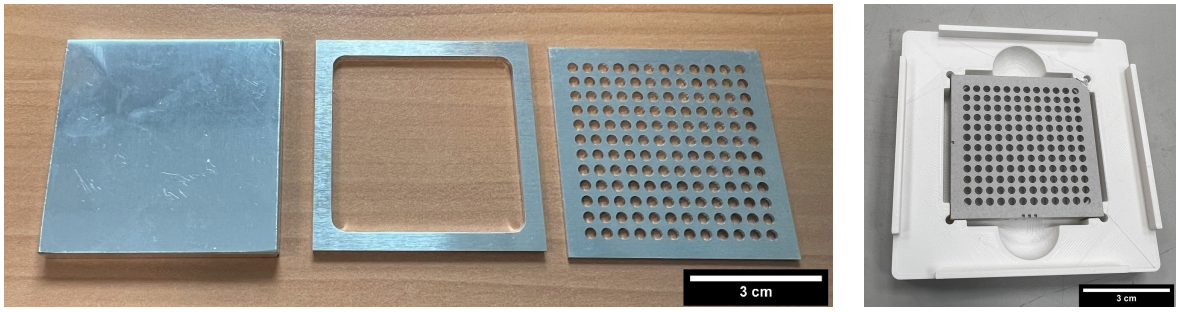


Fig. 4. Simulated BRF of the final design for various illumination configurations. In addition to the main retro-reflective lobe, visible for all configurations, a set of secondary peaks also moves with the illumination direction. The goniospectrophotometer used for this study cannot acquire measurements in the retro-reflective direction; therefore, the experimental plan focuses on capturing the secondary peaks.



(a) (b)

Fig. 5. Manufactured artifact photographs. (a) (Left) Parts prior to PEO coating. (b) (Right) Full assembly, after PEO coating, in the protective box.

Appendix B for a more detailed description of the artifact design and manufacturing processes.

E. Dimensional Measurements

Two alternatives were considered for measuring the mechanical dimensions of the samples.

1) A contact coordinate measuring machine (CMM), where a physical tip is used to determine the position of the surfaces at predefined points. The results produced with this method can be affected by surface asperities; also it is not suited to measuring samples that have a lot of features, such as those in this work, so would take a very long time to undertake the measurements required.
2) A vision CMM, which uses an optical sensor and detects the optical surface, is the same process as when measured by the goniospectrophotometer. This provides a more accurate result and is better suited to samples with many features.

Considering the trade-offs associated with each alternative, a vision CMM was used (Mitutoyo Quick Vision Hyper CNC Vision), calibrated using an SI-traceable calibration chart. The parts to be measured were placed on the work table of the instrument. All dimensional measurements were made with respect to an identified corner of the sample.

The following measurements were made on the individual parts, after coating.

1) Mean diameter (with a metric for their circularity) and location of all the holes with respect to a specified corner of the sample.
2) Flatness and thickness variation of the top hole plate.
3) Flatness variation of both surfaces of the 5 mm thick block.

The samples were then assembled using a 3-D-printed jig to ensure that the components were correctly aligned during assembly and to avoid anything touching the main faces. A low-viscosity two-part epoxy adhesive (Loctite HY4070) was used to secure the parts together, a weight was applied to minimize the glue thickness and any angular errors between the surfaces. Post-assembly, the following dimensional measurements were then made using the same measurement system as before.

1) Spacing between the top of the top plate and the top of the 5 mm block, measured through the holes.

The actual hole positions, sizes, and dimensions were used to build an accurate 3-D model of the target, so that the actual sample geometry can be simulated, giving the best-modeled response for the sample. The measurements are summarized in Table I. The estimated instrumental measurement uncertainty in the diameters and distances was typically less than 7 μm.[2]

A witness sample of coated aluminum was mechanically sectioned and optically polished so that a measurement of the coating thickness could be obtained, this was found to be 55 ± 5 μm.

The samples were shipped in a custom box that supported the sample only by the extreme edges of the main face and

[2]The uncertainties reported in Table I are a statistical variation in the values measured which aggregate all sources of uncertainty, not only the instrumental contribution estimated at 7 μm.

TABLE I
DIMENSIONAL MEASUREMENT SUMMARY. UNCERTAINTY ($k = 1$) IS REPORTED

Quantity	Value	Units
Mean hole diameter	2.55 ± 0.02	mm
Mean hole centre spacing	3.600 ± 0.012	mm
Mean hole plate thickness	1.130 ± 0.044	mm
Mean spacing of top of hole plate to top of 5 mm block	3.26 ± 0.02	mm
Hole plate top surface flatness (peak-to-valley)	25	µm
5 mm block surface flatness (peak-to-valley)	17	µm

clamped it in place; in this way, the main surfaces would not be contacted and possibly contaminated during storage/transit and handling.

F. Optical Measurements

For measurement, the sample under test was mounted in a custom jig that supported the sample by the extreme edges of the rear main face, allowing the edges of the top face to be pressed against the reference plane of the measurement goniospectrophotometer. The jig enabled the sample to be clamped from behind, without the back of the sample being directly contacted. The coating BRDF was measured from the back of the sample and the BRDF of the pattern from the front. The measurement beam was centered on the sample in both cases.

A routine measurement of a test sample begins by aligning its illumination side surface to the optical center of the instrument by adjusting the sample holder micrometer. The correct level is found by aligning the surface with a self-leveling auxiliary laser that is pointed at the optical center with a horizontal beam. For aligning the instrument, mirrors are used to check the level of parallelism of back reflection at relevant instrument components. Each measurement is preceded by a reference material measurement to calibrate the BRDF scale. The reference material is measured for its in-plane BRDF at an incident zenith angle of 0° and a viewing zenith angle of 15°. Several viewing azimuth angle values are also evaluated for averaging purposes.

The illuminating motorized arm can move the incident zenith angle from −90° to 90° in-plane, above the sample. The incident azimuth angle is set by a manual rotational stage. The illumination is shadowed in a ±5° cone around the detector arm: therefore, the instrument cannot measure retro-reflection. Illumination is produced by a combination of a supercontinuum laser and a laser-line tunable filter, which provides a variable bandwidth to the illumination optics. During the measurements, a 4 nm bandwidth was used. Illumination power can be varied in the supercontinuum laser source to suit the reflectance of various types of samples, with an average power of 300 mW in the visible wavelength range. All measurements presented hereafter were performed at 500 nm. The illuminating beam size is set in the iris, which had a 10.8 mm diameter for the measurements, covering a total of 12 holes during artifact measurement.

The BRDF is sampled by the viewing zenith angle arm, in combination with the viewing azimuth angle arm, which samples the entire hemisphere by scanning samples in a viewing zenith angle range of −90° to 90°, and in a viewing azimuth angle range of 0° to 180°. The detector arm is equipped with a 10.8 mm diameter off-axis parabolic mirror which translates to a solid angle of ∼0.0022 sr using a distance of 204.7 mm. This corresponds to an angular resolution of 2.89°. Finally, the reflected signal is provided to a silicon detector by fiber coupling and amplified by a transimpedance amplifier in a dynamic range from 10^2 to 10^9.

The LabVIEW script controlling BRDF sampling has a scan speed of roughly 2.7 points per minute (i.e., 162/h). The scan speed is dominated by the averaging process that takes nine samples at each geometry. The averaging process reduces measurement noise from various sources, such as noise in the supercontinuum laser stability and fluctuations in the electronics as a function of time. The BRDF sampling script uses a state machine that programmatically sets the wavelength, incident zenith angle, viewing azimuth angle, and viewing zenith angle, in order: it samples 2-D slices of the hemisphere when varying the viewing azimuth angle while maintaining a static wavelength and incident angle. In the following, for a given sensor, azimuth value ϕ_{sen}, positive zenith values map to ϕ_{sen}, and negative values map to $\phi_{\text{sen}} + 180°$.

1) Uncertainty Budget: Lanevski et al. [21] have estimated the uncertainty budget for the measurement of the reflectance of a flat sample with the 3-D goniospectrophotometer. We derive here from that work a new uncertainty budget specific to the tested artifact (see Table II). We report the uncertainty of various sources, as well as the resulting uncertainty on the measured reflectance. Standard uncertainties that directly contribute to BRDF uncertainties, such as measurement noise, have been combined quadratically with uncertainties related to the reference sample's BRDF. As a result, BRDF uncertainties for the artifact feature additional components in comparison to nonrelative measurement instruments.

The most significant contributor to the uncertainty budget arises from instrument stability. We evaluated it by calculating the average standard deviation in BRDF values for the artifact across various geometries and wavelengths. BRDF values were taken on different days, with the sample removed and reinserted in the instrument. The evaluated instrument stability had an uncertainty of 0.82%. Enhancements in instrument stability can potentially be achieved through improvements in the consistency of sample placement within the sample holder. Ongoing efforts are underway to develop techniques for more precise control of the sample's incident azimuth angle.

The uncertainty in BRDF due to wavelength was determined by multiplying the standard uncertainty in wavelength by the maximum observed slope in BRDF as a function of wavelength at each geometry. The artifact shows small variations in BRDF with respect to wavelength, which can be attributed to its spectrally flat material properties. Consequently, the artifact had a relative uncertainty of 0.02% in BRDF due to wavelength.

TABLE II
UNCERTAINTY BUDGET FOR ARTIFACT REFLECTANCE MEASUREMENTS

Source of uncertainty	Standard uncertainty	Uncertainty in BRDF [%]
Measurement noise[a,b]	0.42 % to 0.57 %	0.51 to 0.70
Instrument stability	0.82 %	0.82
Wavelength[c]	0.15 nm	0.02
Straylight (isochromatic)	< 0.01 %	< 0.01
Detector linearity[b]	0.05 % to 0.09 %	0.10 to 0.18
Beam size and sample placement[c]	0.20 %	0.20
Incident zenith angle[c]	0.25°	0.21 to 0.38
Viewing zenith angle[c]	0.25°	0.10 to 0.52
Viewing azimuth angle[a,c]	0.25°	0.20 to 0.58
Sample surface level	0.1 mm	0.14
Polarization[c]	0.10 %	0.10
Reference material	0.5 %	0.50
Combined standard uncertainty ($k = 1$)		1.16 to 1.50

[a] This uncertainty depends on the viewing zenith angle
[b] This uncertainty depends on the detector type and selected gain of amplifiers
[c] This uncertainty depends on the sample

To evaluate the effect of sample placement on the BRDF, considering both the incident beam size and the detector field of view, a 14.4 mm beam spot was scanned across the artifact surface in 1.2 mm increments over seven iterations. The standard deviation of BRDF values at different surface positions was calculated for a specific viewing zenith angle. The resulting BRDF variations, as a function of sample surface position, ranged from 0.05% to 0.46%. The uncertainty due to beam size and sample placement was determined as the average variation in BRDF across the sample surface, resulting in a relative uncertainty of 0.20 %.

The standard uncertainty associated with setting the incident and viewing angles was previously determined by Lanevski et al. [21] and is primarily based on the alignment of the optical axes. It is crucial that all optical axes converge at a single point in space, which should align with the incident surface of the sample. To obtain the relative uncertainty in BRDF as a function of incident and viewing angles, we multiplied the standard uncertainty in the respective angles by the maximum observed slope in BRDF corresponding to the incident zenith angle, viewing zenith angle, and viewing azimuth angle. Due to the array of holes in the artifact surface, the reflective characteristics of the sample exhibit a strong dependence on the geometric parameters. Therefore, the relative uncertainty in BRDF as a function of angles had a major contribution to the uncertainty table, with uncertainties ranging from 0.21% to 0.38% for the incident zenith angle, 0.10% to 0.52% for the viewing zenith angle, and 0.20% to 0.58% for the viewing azimuth angle.

2) Material Measurements: The material BRDF measurement was performed using a regularly gridded measurement space. The illumination azimuth was restricted to 0° based on the assumption that the material is isotropic (see Section II-D and Appendix B-B). The other angular dimensions were sampled to balance between dataset density and the feasibility of the measurement campaign, also accounting for the fact that the material's reflectivity is assumed to be close to Lambertian (see Table III). The uncertainty budget associated with this measurement campaign is described by Lanevski et al. [21].

TABLE III
MATERIAL OPTICAL MEASUREMENT SPACE

Quantity	Symbol	Values
Wavelength	λ	500 nm
Illumination azimuth	ϕ_{ill}	0°
Illumination zenith	θ_{ill}	$(0, 15, 30, 45, 60)°$
Sensor azimuth	ϕ_{sen}	$(0, 45, 90, 135)°$
Sensor zenith	θ_{sen}	$(80, 75, \ldots, -80)°$

TABLE IV
ARTIFACT OPTICAL MEASUREMENT SPACE. SEE ALSO FIG. 6

Quantity	Symbol	Values
Wavelength	λ	500 nm
Illumination azimuth	ϕ_{ill}	0°
Illumination zenith	θ_{ill}	$(0, 15, 30)°$
Sensor azimuth	ϕ_{sen}	*see below*
Sensor zenith	θ_{sen}	$(80, 75, \ldots, -80)°$
θ_{ill} [°]	ϕ_{sen} [°]	
0	0, 45, 90, 135	
15	0, 50, 75, 105, 130	
30	0, 60, 120	

3) Artifact Measurements: The artifact reflectance measurement space was defined to capture as well as possible the reflective features detected during the design phase. Since the reflective pattern varies against the illumination configuration (see Fig. 4), a different set of sensor angles is required for each illumination. In practice, the sensor zenith coverage remains unchanged ($\theta_{\text{sen}} \in (-80, -75, \ldots, 80)°$), and the azimuth ϕ_{sen} values are adapted to follow the most prominent reflective peaks (see Table IV and Fig. 6). The uncertainty budget associated with this measurement campaign is summarized in Table II.

G. Numerical Simulations

All simulations done during the design and characterization phases were performed using the Eradiate RTM, with a path-tracing algorithm.

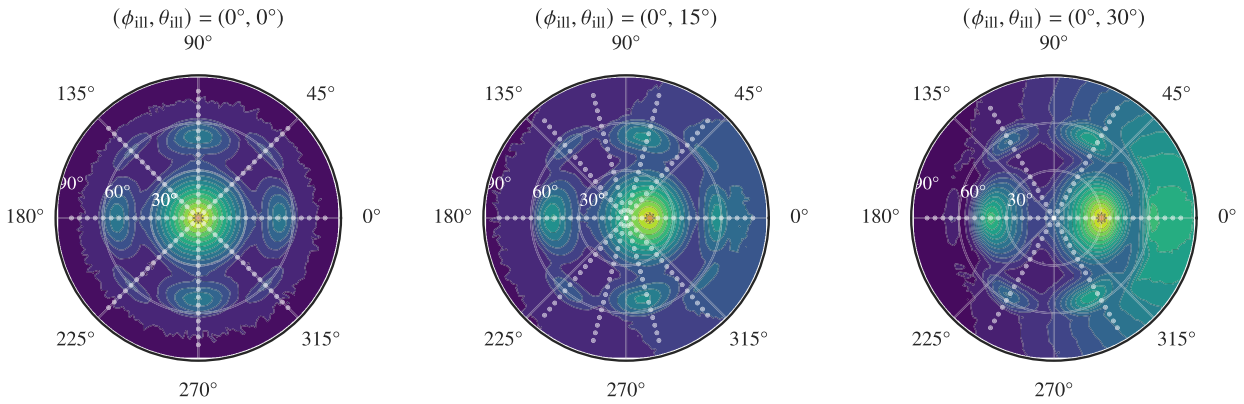


Fig. 6. Artifact measurement space. Contours are the same as on Fig. 4. Each white marker corresponds to a measurement data point. Data points are selected to capture characteristic features in each configuration, under the constraint of iso-azimuth plane sweeps.

During the design phase, the goal was to compute the BRF of the artifact model. Therefore, ideal, perfectly directional illumination and sensors were used. The sensor's field of view was restricted to a footprint similar to that of the goniospectrophotometer's sensor. The material was assumed Lambertian and spectrally uniform, with a reflectance equal to 0.5.

For the characterization phase, a more realistic setup was used.

1) The illumination was simulated using a divergent beam with textured emission to account for the Gaussian beam profile of the actual illuminator.
2) The sensor was simulated by a directional sensor using a cross section target (diameter 10.8 mm) reproducing the field of view of the actual radiometer mounted on the measurement facility.
3) The material was modeled using a *quasi-diffuse* data-driven reflection model.

Our quasi-diffuse BRDF model represents a uniform and isotropic[3] Material whose BRDF depends on the incoming and outgoing zenith angles θ_i and θ_o, and the azimuth angle difference $\phi_d = \phi_o - \phi_i$ between the incoming and outgoing directions. The implementation relies on a table holding BRDF values tabulated against the cosine incoming (resp. outgoing) zenith angle μ_i (resp. μ_o) and ϕ_d. This parametrization optimizes lookup performance. Sampling relies on a basic cosine-hemisphere distribution, which makes it inefficient for materials featuring strong reflective lobes.[4]

Each reflectance simulation run follows the steps outlined below, similar to typical laboratory and field reflectance measurement protocols, for each illumination direction $\boldsymbol{\omega}_{ill}$ and sensor direction $\boldsymbol{\omega}_{sen}$.

1) Generate an artifact model.
2) Simulate the radiance $L_{art}(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen})$ reflected by the artifact model and recorded by the sensor.
3) Simulate a reference radiance record $L_{ref}(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen})$ using a Lambertian surface with reflectance $\rho_{ref} = 1$.

The reflectance is then given by

$$R(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen}) = \frac{L_{art}(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen})}{L_{ref}(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen})}. \tag{1}$$

The quantity R defined in (1) is not identified as one of the reflectance quantities defined by Nicodemus et al. [14] on purpose.

In an ideal setup with directional illumination and sensor, which corresponds to the numerical experiments performed during the design step, the radiance reflected by the nonabsorbing Lambertian surface is a perfect proxy of the incident radiance, i.e., the radiance L_e emitted by the illuminant

$$\begin{aligned} L_{ref}(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen}) &= \frac{\rho_{ref}}{\pi} \int_{2\pi} L_i(\boldsymbol{\omega}_i)\, d\Omega_i \\ &= \frac{1}{\pi} \int_{2\pi} L_e\, \delta(\boldsymbol{\omega}_i - \boldsymbol{\omega}_{ill})\, d\Omega_i \\ &= \frac{L_e \cos\theta_{ill}}{\pi} \end{aligned}$$

where δ is Dirac's delta distribution, and $d\Omega_i = \cos\theta_i\, d\boldsymbol{\omega}_i$.[5] In that case, the reflectance is therefore [coming back to (1)]

$$R_{BRF}(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen}) = \pi \frac{L_{art}(\boldsymbol{\omega}_{ill}, \boldsymbol{\omega}_{sen})}{L_e \cos\theta_{ill}}$$

and is the BRF as defined by Nicodemus et al. [14].

In the more realistic setup used for the optical characterization of the artificial target and the corresponding simulations, (1) yields a hemispherical–conical reflectance factor (HCRF) as defined by Nicodemus et al. [14]. The closer the illumination and sensor get to the ideal directional case, the better the HCRF approximates the BRF.

It is worth noting that although this HCRF is certainly a good approximation of the actual BRF thanks to the low illumination and sensor beam divergence, it is not required that they match perfectly for the purpose of fulfilling our objective:

[3] *Uniform* is to be understood as having optical properties that are invariant against spatial coordinates. *Isotropic* is to be understood as having optical properties invariant by rotation of the material around the local normal direction.

[4] Hence the name *quasi-diffuse*.

[5] It should be noted that we distinguish the *incoming* and *outgoing* directions (ω_i, ω_o), local to the scattering frame and used to describe the scattering properties of surfaces, and the *illumination* and *sensor* directions (ω_{ill}, ω_{sen}), which describe the physical setup of the actual or simulated experimental facility (also sometimes called *scene parameters*).

what really matters is that the experimental methodology is reproduced as accurately as possible numerically.

All artifact models are generated using the volume Boolean operation features of the PyVista [24] mesh processing library. Its Python interface allows for seamless integration in an Eradiate-centric workflow.

H. Uncertainty Propagation

In the characterization step, the uncertainty on input parameters is propagated to allow for metrologically correct validation of simulation results. For that purpose, the simulation workbench wraps scene assembly and Eradiate runs in a workflow based on the punpy library [25]. Punpy propagates uncertainties on input quantities through any Python function, evaluating the uncertainty on the output. It implements a law of propagation of uncertainties-based method, as well as Monte Carlo (MC) sampling. The highly dimensional input space we are dealing with makes an MC method the natural choice: evaluating the Jacobian of the measurement function would be very costly.

The measurement function takes as the input a state vector containing all simulation parameters. This includes.

1) Artifact model generation parameters, namely hole positions and diameters (432 parameters).
2) Artifact positioning (three parameters).
3) Material BRDF data points (132 parameters).
4) Illumination beam divergence (one parameter) and angular positioning (two parameters).
5) Sensor angular positioning (two parameters) and beamwidth (one parameter).

In total, the state vector has 573 dimensions.

The measurement function aggregates the artifact model generation and RT computation steps. The path tracing algorithm used implements next event estimation, a common variance reduction technique consisting of accumulating emitter contributions at each node of the path generated during the radiative random walk. For a given illumination configuration and target model, all sensor configurations are evaluated at once thanks to Eradiate's automated multisensor sequencing: the measurement function, therefore, returns all requested sensor configurations at once.

Fig. 7 provides a summarized overview of the simulation chain. The simulation chain outputs the variance of the generated dataset as the uncertainty ($k = 1$). This is a notable difference from previous work [7], where the uncertainty ($k = 2$) was used: in principle, comparing the present uncertainty estimates with prior values is not possible. For comparability, we estimate the uncertainty ($k = 2$) by doubling the standard uncertainty. This estimate, although only valid for normally distributed random variables, provides acceptable orders of magnitude for comparison with prior work.

I. Comparison Method

The main metric monitored to check for agreement between the simulations and measurements is the relative difference $\Delta R^\star$ between the simulated and measured reflectance values

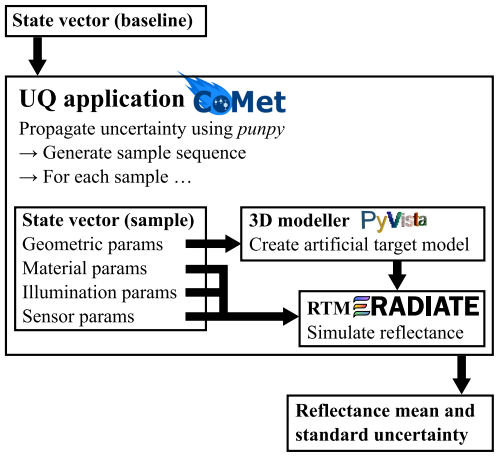


Fig. 7. Schematic of the uncertainty quantification application used in this study. The entire toolchain is written in Python.

R_{sim} and R_{mes}, using the measurement as the reference

$$\Delta R^\star = \frac{R_{\text{sim}} - R_{\text{mes}}}{R_{\text{mes}}} \tag{2}$$

where the dependencies to the illumination and sensor angles are omitted for brevity. A common practice in the remote sensing community is to categorize data points based on the value $\Delta R^\star$: for instance, a simulated data point will be said to "match well" the corresponding experimental data point if $|\Delta R^\star| < 2\,\%$. This straightforward comparison method has the disadvantage of defining a criterion independent of the uncertainty or the order of magnitude of the compared quantities, yielding false positives or negatives when the uncertainties on the simulated and measured data are different or when R_{mes} and R_{sim} have a low order of magnitude.

To address these issues, we suggest a comparison method based on the t-score

$$t' = \frac{R_{\text{sim}} - R_{\text{mes}}}{\sqrt{\sigma_{\text{sim}}^2 + \sigma_{\text{mes}}^2}} \tag{3}$$

where σ_{sim} (resp. σ_{mes}) is the uncertainty ($k = 1$) associated with the simulated (resp. measured) data. The cut-off criterion is defined by

$$\left|t'\right| < \alpha \tag{4}$$

i.e.,

$$\left|\Delta R^\star\right| = \left|\frac{R_{\text{sim}} - R_{\text{mes}}}{R_{\text{mes}}}\right| < \alpha \frac{\sqrt{\sigma_{\text{sim}}^2 + \sigma_{\text{mes}}^2}}{R_{\text{mes}}}. \tag{5}$$

For simplicity, we will denote σ_c the combined simulation and measurement uncertainty

$$\sigma_c = \sqrt{\sigma_{\text{sim}}^2 + \sigma_{\text{mes}}^2}. \tag{6}$$

In practice, this metric compares the distance between the mean values of the measured and simulated data samples with the combined uncertainties. The α factor defines the coverage factor associated with the considered uncertainties: with $\alpha = 1$, (5) becomes

$$\left|\Delta R^\star\right| < \frac{\sqrt{\sigma_{\text{sim}}^2 + \sigma_{\text{mes}}^2}}{R_{\text{mes}}} = \frac{\sigma_c}{R_{\text{mes}}} \tag{7}$$

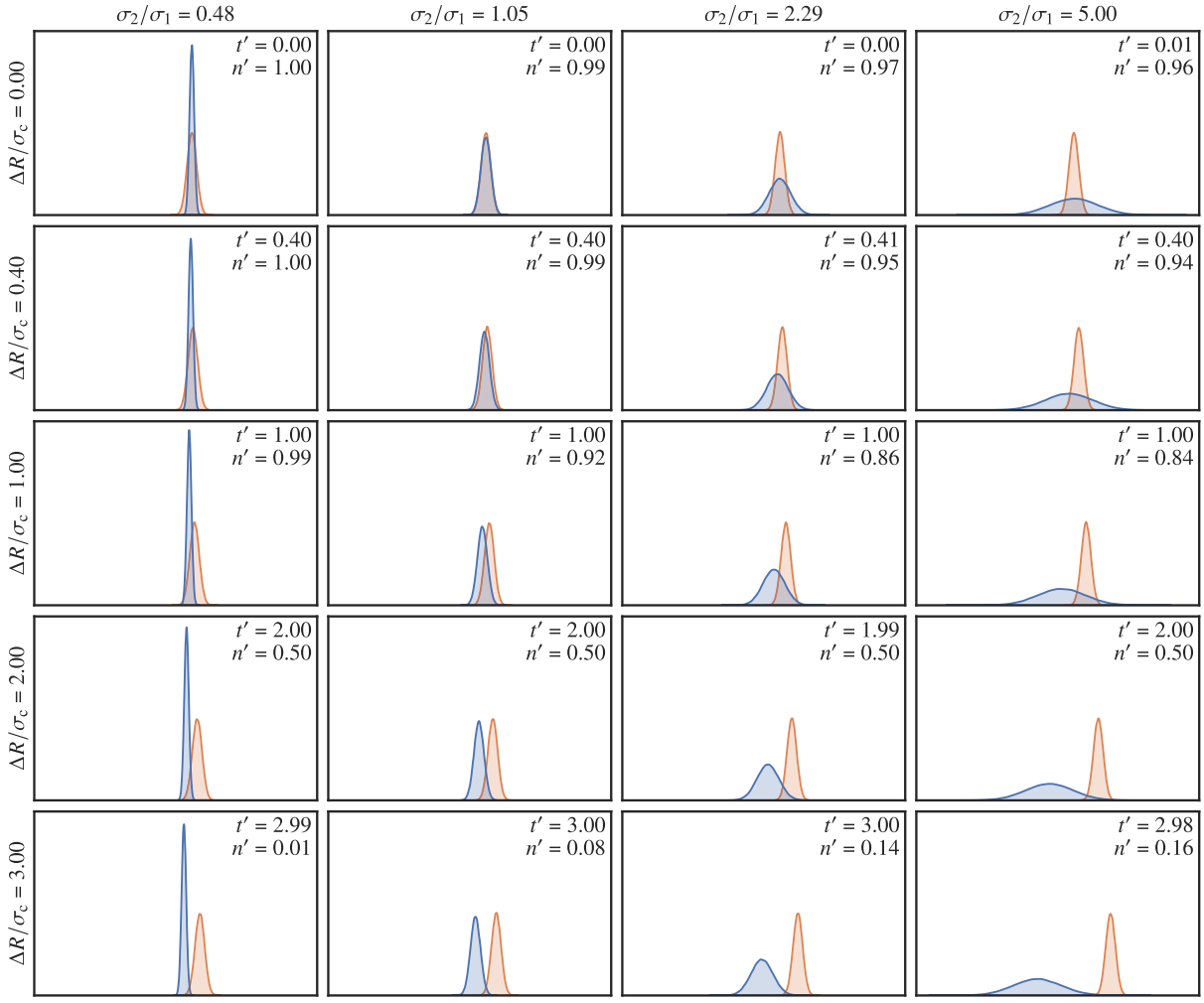


Fig. 8. Samples (population of 10^5) of two normally distributed random variables (reference, orange, index 1; and test, blue, index 2) for various standard deviation and mean values, and corresponding t-score t' and level of agreement n'. The t-score reflects the distance between sample means relative to the combined uncertainties of both samples, while the level of agreement reflects the amount of overlap between the test sample distribution and the reference within combined uncertainties.

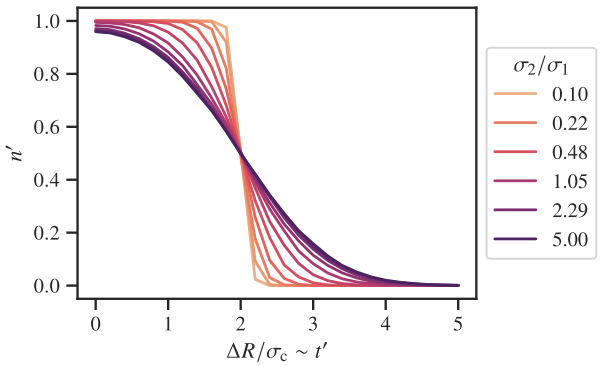


Fig. 9. Level of agreement against the distance between mean values and standard deviation ratios for two normally distributed random variables. Depending on the standard deviation ratio, the statistical similarity quantified by t' results in different values of n'. The slope of the n' versus t' relationship in the $t' \in [1, 3]$ range is strongly dependent on the standard deviation ratio, which, in practice, results in the level of agreement metric being less subtle than the t-score in that range.

and uses the standard uncertainty; with $\alpha = 2$, it becomes

$$\left|\Delta R^{\star}\right| < \frac{\sqrt{(2\sigma_{\text{sim}})^2 + (2\sigma_{\text{mes}})^2}}{R_{\text{mes}}} = \frac{2\sigma_c}{R_{\text{mes}}} \tag{8}$$

effectively doubling the standard uncertainty, which is equivalent to a coverage of $k = 2$, assuming that the uncertainty follows a normal distribution. For the purpose of making our analysis comparable to the results of Jaanson et al. [7], we set $\alpha = 2$ and consider the double-standard uncertainty $2\sigma_{\text{sim}}$ and $2\sigma_{\text{mes}}$.

It should be noted that t' does not appear as a metric in Jaanson et al.'s work; instead, a level of agreement (denoted n' in the following) is used, and defined as the fraction of simulated data points that agree with their corresponding measured data points within the combined uncertainty of simulated and measured data ($k = 2$). In other words

$$n' = \text{fraction of samples for which}$$
$$\left|R_{\text{sim},i} - R_{\text{mes}}\right| < \underbrace{\sqrt{(2\sigma_{\text{sim}})^2 + (2\sigma_{\text{mes}})^2}}_{=2\sigma_c}$$

where $R_{\text{sim},i}$ is the simulated reflectance for sample i.

Since our simulation workbench only outputs the mean and standard uncertainty of the MC uncertainty propagation process, we cannot compute n' directly. However, assuming that all variables follow a normal distribution, we can estimate

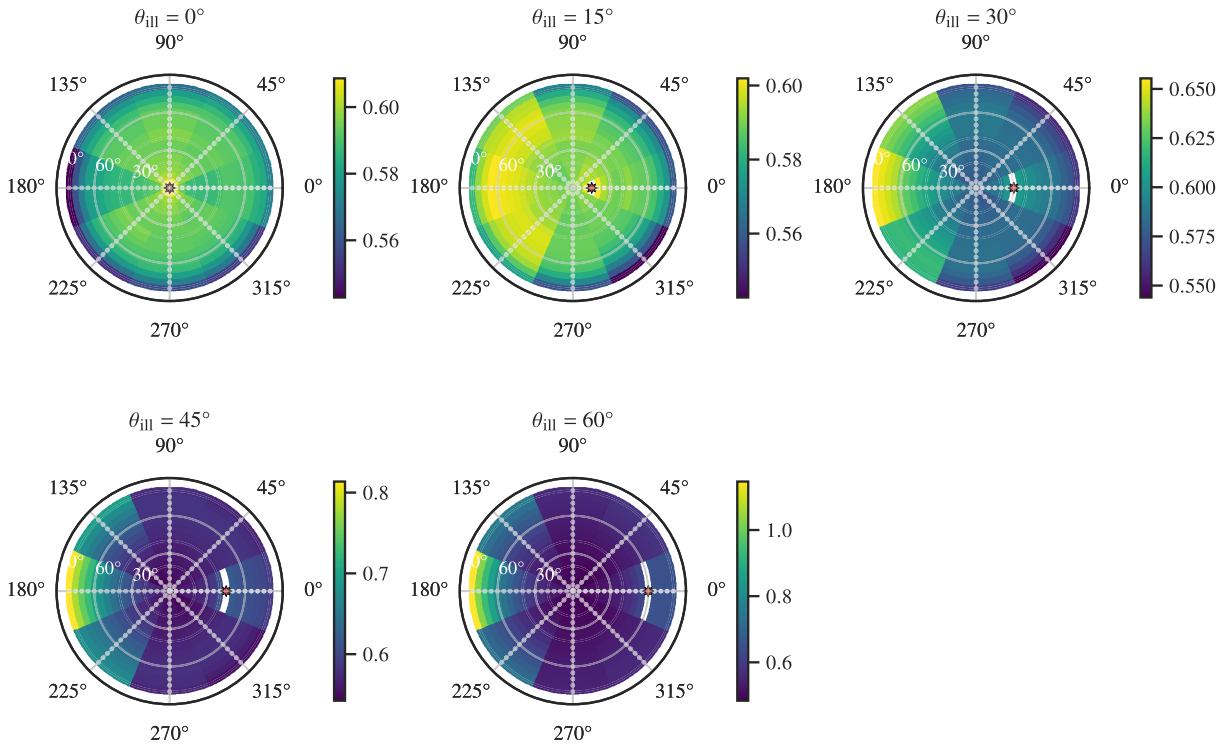


Fig. 10. Material BRF measurement dataset ($\lambda = 500\,\text{nm}$) faceted against the illumination zenith dimension. The white dots represent the measurement space. A data point is missing from all measurements in the retro-reflective direction, which cannot be visited by the goniospectrophotometer.

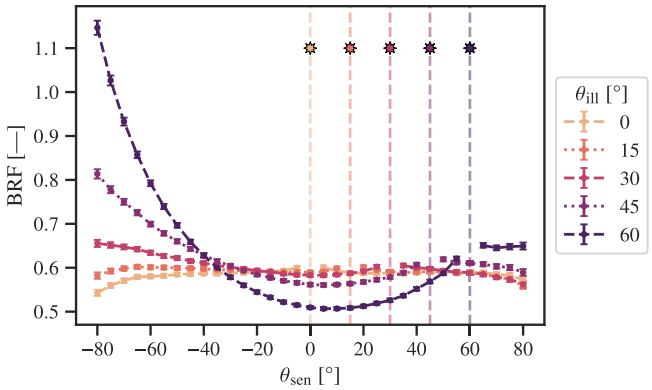


Fig. 11. Material BRF measurement dataset ($\lambda = 500\,\text{nm}$), principal plane view. By convention, the 180° half-hemisphere is remapped to the 0° half-hemisphere with negative zenith angle values. Error bars report the standard uncertainty ($k = 1$).

n' from the t-score t' and the standard deviation ratio $\sigma_{sim}/\sigma_{mes}$ using a simple MC process. We do this using two normally distributed random variables (see Fig. 8). Although the extreme trends for t' and n' are similar, the level of agreement metric might fail to report a visible disagreement between the two data series [e.g., Fig. 8 for $(\Delta R/\sigma_c, \sigma_2/\sigma_1) = (1, 0.48)$]. More generally, the t-score provides a more progressive assessment of the statistical similarity of both data series in the range $t' \in [0, 2]$ than the level of agreement, regardless of the relative values of the uncertainties on both data series (see Fig. 9). For this reason, we advocate the use of the t-score to assess the performance of the method. Nevertheless, we tabulate the relationship $n'(\Delta R/\sigma_c, \sigma_{sim}/\sigma_{mes})$ to estimate the level of agreement between our simulated and experimental datasets. This ensures comparability with prior work.

We therefore monitor three quantities in the analysis.

1) The relative difference $\Delta R^\star$.
2) The t-score t'.
3) The level of agreement n'.

These three metrics provide several views on the agreement between simulated and measured data, each achieving different trade-offs.

Finally, a global performance indicator relevant for comparison with prior work is the relative root-mean-square error (RMSE), defined as

$$\epsilon = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}\left(\frac{R_{sim,i} - R_{mes}}{R_{mes}}\right)^2}. \tag{9}$$

III. Analysis

A. Material Measurements

Material measurements are performed using the 3-D goniospectrophotometer, following the aforementioned material measurement plan (see Section II-F). All measured data are associated with a corresponding standard uncertainty ($k = 1$) based on the material reflectance uncertainty budget (see Section II-F). This uncertainty is used as an input of the uncertainty propagation step. In addition to global views on the dataset in polar coordinates and the principal plane (see Figs. 10 and 11), we provide the full dataset in a dedicated appendix (see Appendix C).

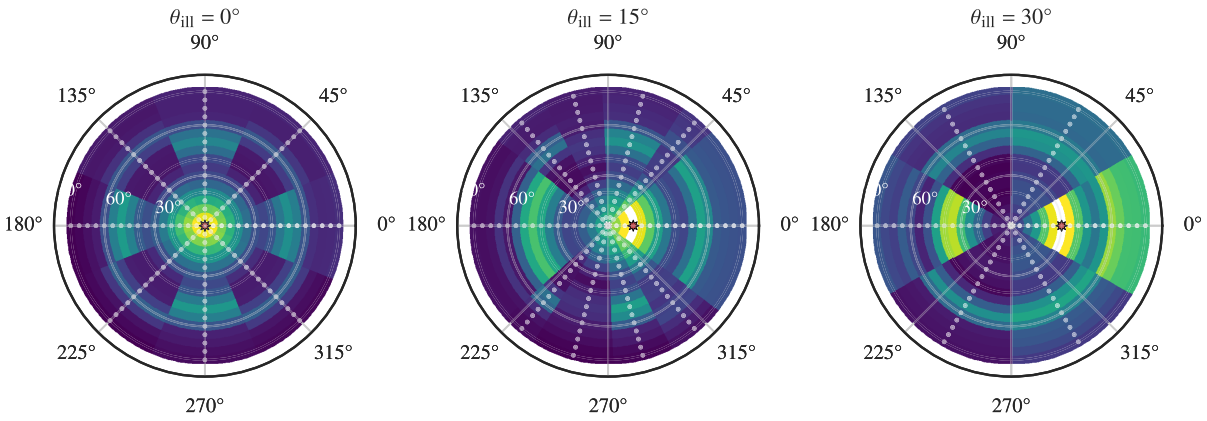


Fig. 12. Artifact BRF measurement dataset ($\lambda = 500\,\text{nm}$) faceted against the illumination zenith dimension. The white dots represent the measurement space. As in Fig. 10, the retro-reflective direction cannot be visited. Comparison with Fig. 6 reveals that the reflective features are positioned where expected.

In nadir and close-to-nadir illumination pointing configurations ($\theta_{ill} \in \{0, 15\}°$), the material has near-Lambertian reflectivity with a reflectance of approximately 0.6; at higher illumination zenith angle values, it significantly departs from the Lambertian behavior and exhibits a forward reflective behavior. In all configurations, a retro-reflective peak (see Fig. 11), inaccessible to the goniospectrophotometer, is also present; however, it cannot be measured, due to aforementioned instrument limitations (see Section II-F).

The data-driven BRDF model requires full coverage of the angular space ($(\theta_i, \phi_d, \theta_o) \in [0, \pi/2] \times [0, \pi] \times [0, \pi/2]$) with dense data (no missing data are allowed). We therefore have to assign data values at locations where data are missing. For $\theta_{sen} > 80°$ and $\theta_{ill} > 60°$, the data are extrapolated by assigning the value of the nearest neighboring point. For the retro-reflective point, linear interpolation is performed, leading to underestimating the retro-reflective BRDF value by an undefined amount. Although actively mitigated by the material selection, the impact of those missing data on simulation results is certainly tangible; however, the large number of degrees of freedom and undefined value bounds make quantifying the added uncertainty difficult.

B. Artifact Optical Measurements and Simulation Results

Artifact BRF measurements follow the aforementioned experimental plan (see Section II-F). In addition to a polar view on the experimental data (see Fig. 12) and azimuth plane slice views for comparison of the simulated and measured data (see Fig. 13), we provide the full datasets in a dedicated appendix (see Appendix C). A set of azimuth plane slice views on the relative difference $\Delta R^\star$ (see Fig. 14) also shows how the t-score metric behaves for this dataset.

The simulated and measured reflectance datasets exhibit features similar to those predicted during the design process (see Section II-D) and show consistent qualitative agreement among all sensor azimuth angle values. We observe that the t-score often degrades at high sensor zenith angle values, i.e., for $\theta_{sen} > 60°$ (e.g., Figs. 13 and 14 for $(\theta_{ill}, \phi_{sen}) \in \{(0°, 0°), (15°, 105°), (30°, 0°), (30°, 120°), \ldots\}$).

TABLE V
DATA CATEGORIZATION DEPENDING ON CONSIDERED METRIC

Metric	Good	Acceptable	Poor
$\lvert\Delta R^\star\rvert$ [%]	0 – 2	2 – 4	4+
$\lvert t'\rvert$	0 – 1	1 – 2	2+
n' [%]	95+	95 – 50	50 – 0

The simulation tends to slightly overestimate the reflectance at high negative zenith angle values and to underestimate it at high positive zenith angle values, including in configurations where symmetry is expected. This asymmetry reflects that of the measured data (see Appendix, Fig. 19), amplified at high sensor zenith angles. In other words, the t-score increases because the simulated data, which has the expected symmetries, departs from the measured data, in which the symmetry is broken. This is possibly attributable to alignment issues [26]. These systematic deviations at grazing angles are not included in the uncertainty budget, which is likely undervalued.

To facilitate the analysis, we categorize data points in three categories, depending on metric values (see summary in Table V):

1) *"Good" Agreement:* for $\Delta R^\star$, the threshold is set to 2 %, a typical requirement for accurate radiometric sensors in the visible range at the time of writing; in the context of this study, this order of magnitude translates to a threshold of 1 for t' with the illumination and sensor pointing around the nadir; similar to Jaanson et al. [7], the target value defined for the level of agreement is 95 %.
2) *"Acceptable" Agreement:* for $\Delta R^\star$, the threshold for a "good" agreement is arbitrarily doubled and set to 4 %; we associate a threshold of 2 for t', which corresponds exactly to a threshold of 50 % for n' (see Figs. 8 and 9).
3) *"Poor" Agreement:* this corresponds to $|\Delta R^\star| > 4\,\%$, $|t'| > 2$ or $n' < 50\,\%$.

The global polar view on all metrics (see Fig. 15) shows that the relative deviation metric tends to reject slightly

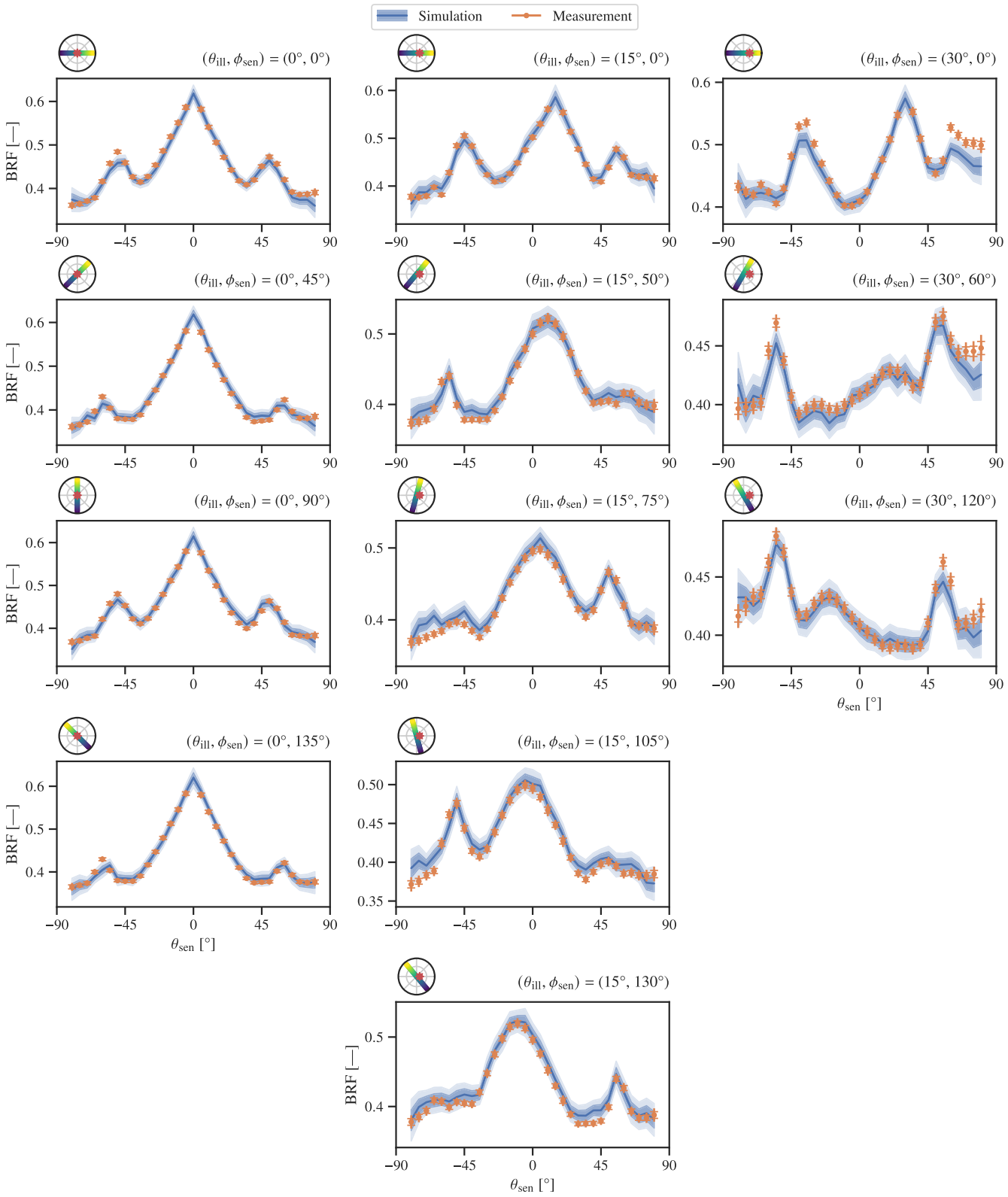


Fig. 13. Comparison of BRF simulation results against the measured dataset. Each plot has a thumbnail attached that specifies the illumination geometry and sensor azimuth angle. The simulated (resp. measured) data are attached to its standard deviation (resp. uncertainty), represented by the lighter shading (resp. horizontal error bars), corresponding to coverage $k = 1$, as well as the double, represented by the darker shading (resp. error bar tips), corresponding to coverage $k = 2$. Qualitative agreement is considered "good" if $k = 1$ error markers overlap, and "acceptable" if $k = 2$ error markers overlap.

more easily data points at high zenith angles than the others because it disregards the increased uncertainty budget in these configurations, but otherwise scatters data points between the "good" and "acceptable" categories

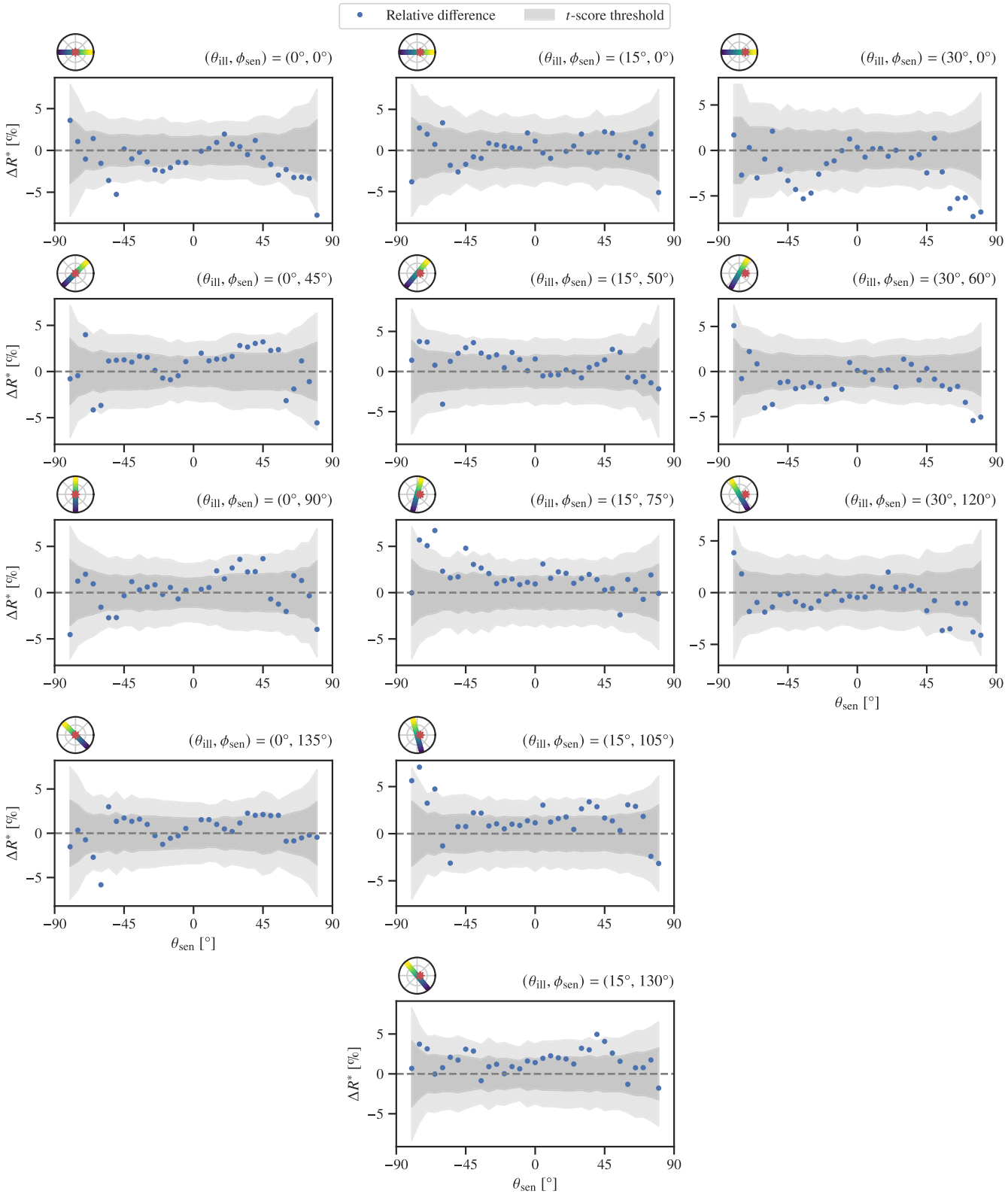


Fig. 14. t-score analysis for all considered geometries (in-plane view). Points display the relative difference between simulated and measured data, and shades show the t-score threshold for $\alpha = 1$ (darker, "good" agreement) and $\alpha = 2$ (lighter, "acceptable" agreement).

similar to the t-score. On the other hand, the level of agreement metric rejects the same data points as the t-score but assigns less easily a data point to the "good" category than the t-score.

The impression of a good qualitative agreement between simulated and experimental data are confirmed by the global per-category aggregates (see Fig. 16): on average, less than 10 % data points show "poor" agreement, and they are

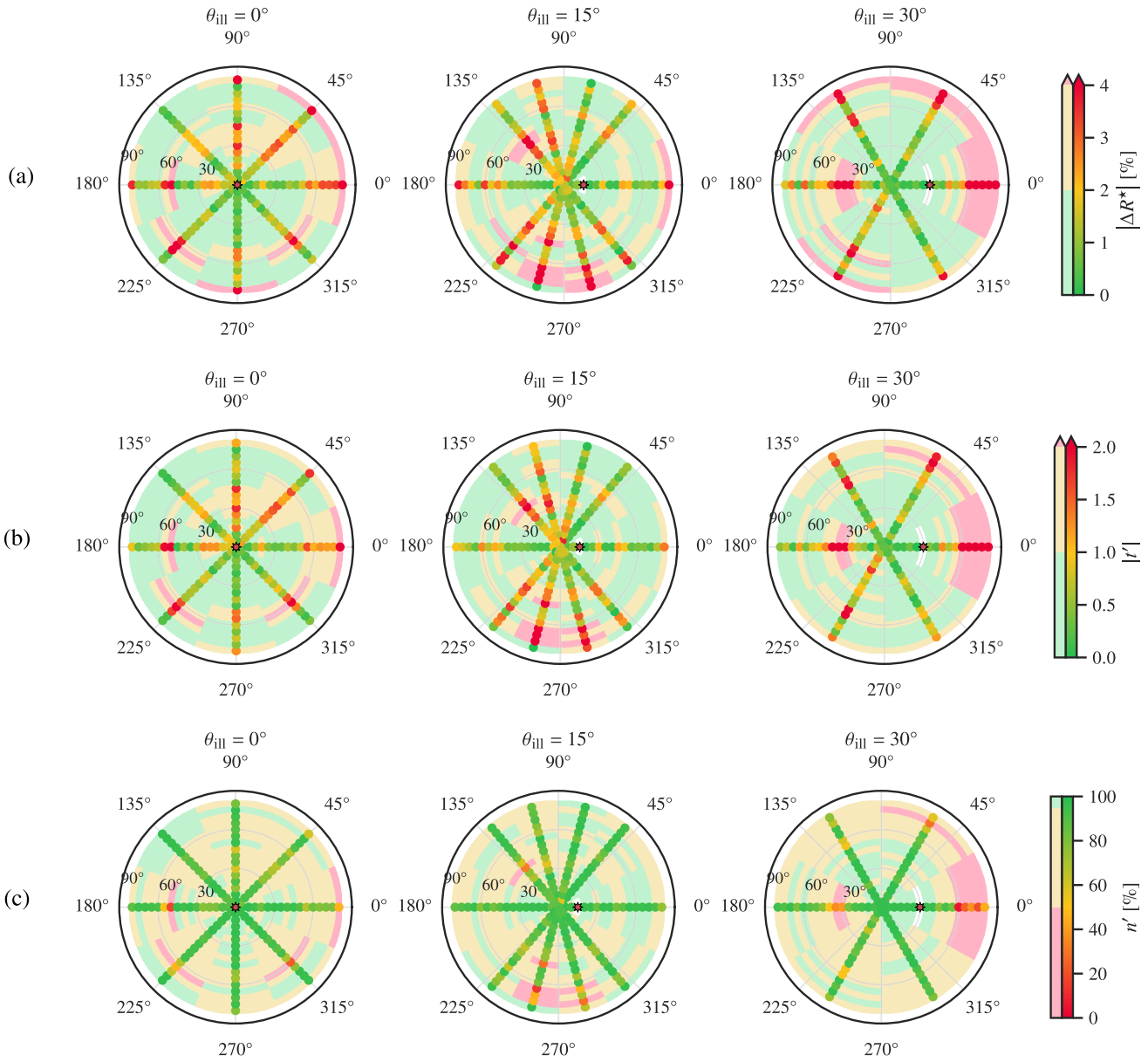


Fig. 15. Global polar view on method metrics. Markers are colored based on the metric evaluated for the corresponding geometric setup. The shading corresponds to the category to which the data point is assigned: "good" (green), "acceptable" (yellow), and "poor" (red). (a) (Top) Relative difference. (b) (Middle) t-score. (c) (Bottom) Level of agreement.

generally associated with configurations with high viewing zenith angle (see above for an interpretation). It is worth noting that improving the uncertainty budget to account for that systematic drift would improve the method performance with respect to the level of agreement and t-score metrics, but not with respect to the relative difference.

Finally, global statistical aggregates on metrics (see Fig. 17) allow us to quantify the method performance. The mean relative deviation is between 1.30 % and 2.42 % on average for each azimuth plane, with a global mean below 2 %. The level of agreement is between 76 % and 90 % on average for each azimuth plane, with a consistent behavior in-plane and out-of-plane.

TABLE VI
COMPARISON OF PERFORMANCE METRICS WITH PRIOR WORK

Metric	This study	Jaanson et al. [7]
Level of agreement (in-plane)	0.84	0.72
Level of agreement (out-of-plane)	0.86	0.45
Level of agreement (global)	0.86	0.60
Relative RMSE (in-plane)	0.026	0.19
Relative RMSE (out-of-plane)	0.022	0.64
Relative RMSE (global)	0.023	0.45

We can compare the global level of agreement and relative RMSE to prior work [7] (see Table VI): both metrics are significantly improved for in-plane and out-of-plane geometries, and also globally.

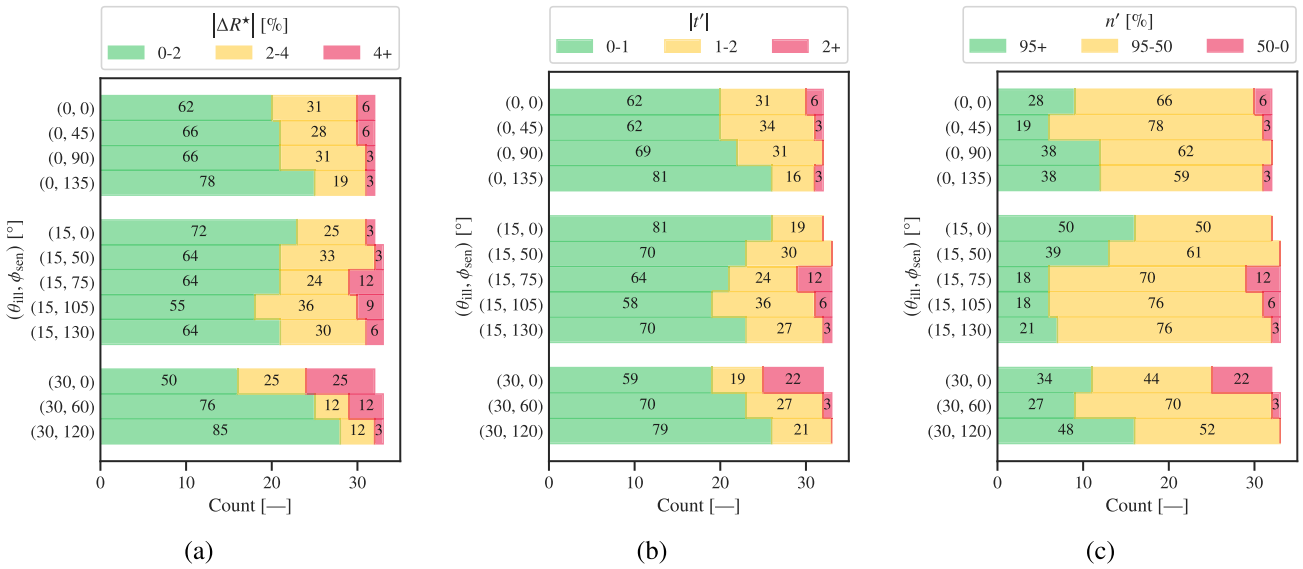


Fig. 16. Global metric reports. For each sensor azimuth and illumination zenith angle (ordinate axis), bars display the number of points associated with each category: "good" (green), "acceptable" (yellow) or "poor" (red). Bar labels indicate the corresponding fraction, in percent, of the total number of data points. (a) (Left): Relative difference. (b) (Middle): t-score. (c) (Right): Level of agreement.

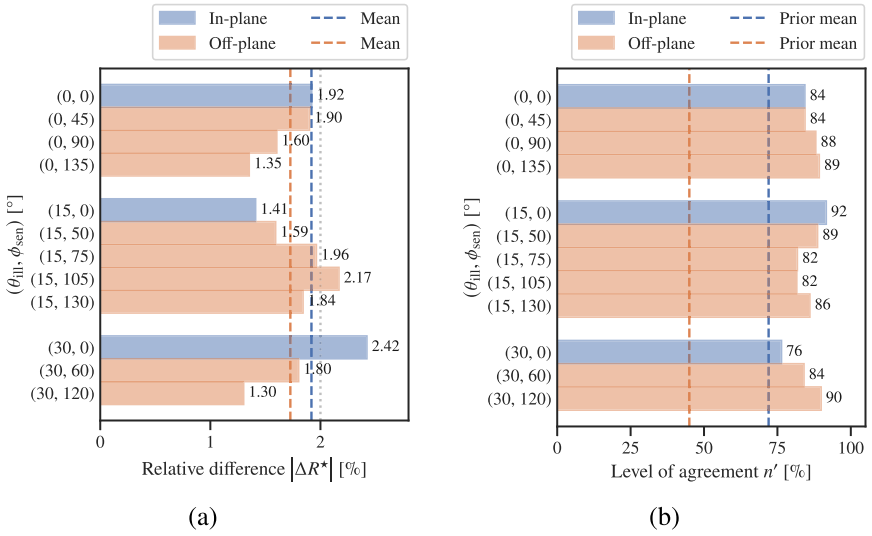


Fig. 17. Mean metric performance for each sensor azimuth and illumination zenith angle. Bar labels indicate the corresponding value. (a) (Left): Relative difference. Dashed lines show the mean value for all in-plane and out-of-plane configurations. (b) (Right): Level of agreement. As a reference, the global value reported by Jaanson et al. [7] is displayed for both the in-plane and out-of-plane configurations.

IV. Conclusion and Outlook

Implementing an unbiased MCRT algorithm is nowadays a routine task: techniques for producing accurate representations of light propagation in natural environments are now easily accessible scientific knowledge. The theoretically perfect accuracy of unbiased MCRT methods dismisses the MC-induced variance in simulation results as a contribution to simulation uncertainty because converging to the correct result only requires a higher sample count. Consequently, the major source of bias and uncertainty one should be concerned with lies in surface and volume scattering modeling—do our models faithfully represent how light is reflected on surfaces or scattered by volumes?—as well as in the quality of the input data used for scene setup.

Our proposed validation approach builds on prior work by Govaerts and Verstraete [6] and Jaanson et al. [7], acknowledging the fact that the most important issue to address is the quality of surface modeling. With the goal of demonstrating the possibility of reproducing laboratory reflectance measurements using an RTM with high accuracy, based on SI-traceable data, we updated the prior methodology to address identified issues. Thanks to our data-driven reflection model, as well as careful material selection and target design, the agreement between simulated and measured data is significantly improved, with a global level of agreement of 86 % (versus 60 % in prior work) and a global relative RMSE of 0.023 (versus 0.45 in prior work), and our method exhibits consistent performance independently of the viewing geometry. The t-score-based metric we suggest also provides an improved

measure of the statistical similarity between simulated and experimental data and better detects systematic modeling bias. The SI-traceability of the measured data makes it possible to link simulation results to a standard reference and ensures comparability with other validation campaigns.

In addition, this study shows that the Eradiate RTM can, for the simulated setup, produce simulated reflectance values that match SI-traceable laboratory measurements with a relative bias of 2 % or lower, provided that the uncertainty on input parameters and measurements are low enough.

Improving further this accuracy requires reducing the uncertainty on both the input data and the reference results, notably by improving the alignment procedure.

The surface scattering model can also be improved: the data-driven model used in this study relies on a linear interpolator and hemisphere-cosine sampling routine and could be improved by making reasonable assumptions on the material (e.g., extrapolating missing data based on a carefully chosen microfacet model). It could also be replaced with a more advanced data-driven model optimizing both modeling accuracy and database size. Among the many approaches that have been proposed, we found Dupuy and Jakob [27] data-driven model to be an interesting alternative because 1) its adaptive parametrization results in excellent accuracy over the full incoming and outgoing hemispheres at a very low storage cost, and 2) it can be sampled efficiently in the context of MCRT. Additionally, it can model anisotropic materials. This approach, however, requires a specific experimental protocol, and only a few SI-traceable facilities (e.g., IO-CSIC's GEFE facility [28], [29]) are equipped to implement it.

Finally, and importantly, the protocol we propose validates only the components of the RTM used in that specific context (here, the specific sensor, illumination, and material models, as well as the path tracing algorithm). In particular, only the surface scattering part of the RT simulation problem is addressed here: while this is of great value, it only addresses a part of the challenges encountered in RT modeling in the context of EO, where atmospheric modeling is of equal importance. Different validation exercises are required to validate other components of RTMs, some of which can be derived from this protocol, and others have to be entirely different.

Appendix A
Eradiate RTM

This appendix provides a summary of essential facts about the Eradiate RTM that are relevant to this work and scattered in this article, but also more general information about the model, more relevant to its general range of applications.

Eradiate [15] is a 3-D RTM designed to support calibration/validation (cal/val) activities in the field of EO. The primary objective underlying its development is to match the radiometric accuracy levels foreseen for the most ambitious satellite missions foreseen in the next decade (e.g., TRUTHS [30]). Eradiate is also an open-source project, distributed under the GNU Lesser General Public License.

Internally, Eradiate is a two-layer system consisting of pre- and post-processing facilities, written in the Python programming language, and a radiometric kernel based on the Mitsuba 3 rendering system [16] that provides cutting-edge technology for the implementation of MCRT methods. The MC method used in this study is a classic path tracer that performs a random walk in the scene starting from the sensor. It implements well-known variance reduction techniques: multiple importance sampling, Russian roulette, and next-event estimation (see e.g., [31] for details). Eradiate also provides an interface to Mitsuba's volumetric path tracer, which supports the simulation of volumetric scattering in participating media and is essential to the inclusion of atmospheric effects in EO-related simulations.

Taking advantage of the plugin architecture of Mitsuba, Eradiate offers various theoretical and realistic sensors and emitters that can be easily swapped depending on the application. In the design phase of this study, perfectly directional light sources and sensors are used to produce an idealized version of the experimental setup, allowing to effectively simulate the actual BRDF of the artifact over the entire outgoing hemisphere. In the experiment simulation phase, a more realistic light source is used to account for the beam profile, and sensors are placed to match the nominal position of experimental radiometric records.

Eradiate also provides interfaces for a variety of geometric primitives, both analytical and data-driven. In this work, the geometry of the target is incorporated into the simulation as a triangulated mesh. Eradiate provides an interface for the OBJ and PLY formats, the latter being recommended for its better specification and more compact binary data storage.

Eradiate offers a range of EO-oriented and more generic surface scattering models. This work uses two of them: a *diffuse*, i.e., Lambertian model, for the design phase; and a data-driven *quasi-diffuse* model. The diffuse model represents a surface that scatters light isotopically, independently of the incoming direction. This model is only theoretical and natural or artificial surfaces generally do not exhibit such reflective behavior, although some can come close. The sampling routine for this model is a cosine-hemisphere.

The quasi-diffuse model provides an interface to a BRDF value table indexed by the incoming and outgoing zenith angles, as well as the relative azimuth angle between the incoming and outgoing directions. It therefore assumes that the material is isotropic, i.e., that its reflective properties have a rotational invariance with respect to the local surface normal. This model uses the same cosine-hemisphere sampling method as the diffuse model, which makes it inefficient if the material has strong reflective lobes. The zenith angle coordinates are spaced evenly in the cosine space, which improves lookup performance. Between table values, the model uses linear interpolation. The accuracy of this reflection model is therefore excellent for data points that require no interpolation but will degrade between table values if the BRDF has steep variations with insufficient angular coverage. For this reason, material selection was done with the strong requirement of being as close as possible to Lambertian, to minimize the inaccuracy introduced using the quasi-diffuse reflection model.

Eradiate can operate in two modes: a *monochromatic* mode, and a *CKD* mode. In the monochromatic mode, used in this work, the RTE is solved for a single wavelength.

Computational results are stored and aggregated along the spectral dimension without modifications. The CKD mode implements a correlated k-distribution method [32] that trades accuracy for performance, effectively allowing to solving the RTE over spectral bands with a sequence of monochromatic simulations combined using a spectral quadrature rule. The CKD mode is used when atmospheric molecular absorption has to be taken into account.

Appendix B
Artifact Design and Manufacturing

This section introduces the artifact design and manufacturing process in depth and expands the summary provided in Section II-D.

A. Artifact Design

Previous validation efforts [6], [7] used an artificial target made of sanded aluminum with a lattice of cubic features, exhibiting strong shadowing-masking effects, which require the ability to account for 3-D radiative effects. In the present case, the design process focused on producing an object featuring a reflective signature with multiple lobes in various directions. In particular, reflective lobes located outside the retro-reflective direction were required: while surface retro-reflection is typical of many natural surfaces, it was not accessible to the measurement facility used in this study.

The following main constraints were applied during shape design.

1) *Object Size:* The sample is to be mounted in the aforementioned optical measurement facility (see Section II-C). This results in a size constraint of $50 \times 50 \times 10$ mm.
2) *Geometric Feature Size:* Geometric features must be large enough to allow for accurate and cost-efficient manufacturing and dimensional measurement, and small enough to ensure representativity given the facility's beam size and sensor field of view so that pointing accuracy does not have a critical impact on measurement uncertainty.
3) *Reference Measurements Plane:* The sample front face is mounted against restraints that define the measurements plane and coincide with the goniospectrophotometer rotation axes. The top surface of the artifact has to be flat, so it has a consistent distance to the measuring instrument.

The artificial target design process was iterative and based on RT simulations performed with Eradiate. At each iteration, a digital model was built, and its BRDF was computed using simplified modeling assumptions.

1) Lambertian material with (spatially and spectrally) uniform reflectance equal to 0.5.
2) Opaque material.
3) Uniform, perfectly directional lighting.
4) Perfectly directional sensor with a footprint matching exactly the upper surface of the digital model.

A set of plates pierced with circular holes organized in a square lattice emerged as a valid design fulfilling the requirements and satisfying the constraints. The final design consists of a $50 \times 50 \times 1$ mm top plate, spaced nominally 2 mm above a solid and flat 5 mm-thick block. The top plate is pierced with a 12×12 array of circular holes (nominal diameter 2.6 mm) with nominal center-to-center spacing 3.6 mm (see Fig. 3). The corresponding BRDF features a main retro-reflective lobe, as well as multiple secondary peaks due to shadowing-masking effects (see Fig. 4).

B. Artifact Manufacturing

The anodized aluminum that the grooved target used in previous work [6], [7] was made of featured highly directional, hard-to-characterize reflective features. This introduced significant modeling uncertainty; while this could be addressed with advanced material characterization and modeling techniques, e.g., by clustering measurement points around reflective lobes using an adaptive sampling method [27], the SI-traceable measurement facility available for this study did not implement such approaches. This is the main reason why target design was performed under a Lambertian material assumption.

To get as close as possible to design assumptions, the following constraints were applied for material selection.

1) *Opacity:* The selected material shall have minimal internal light diffusion so that it can be considered opaque at the length scales relevant to the study. A semi-transparent material would be more complex to model and introduce additional uncertainty in simulation results.
2) *Lambertian Reflectance:* The selected material shall have a reflectance as close as possible to Lambertian, i.e., with 4-D invariances and depolarizing properties.
3) *Gray Spectrum:* The selected material shall have a spectrally flat response, in the mid-gray tone range, close to the 0.5 value used for the design process.
4) *Easy Machining:* The selected material shall allow for cost-effective and precise fabrication of the final design.

Conventional matt white reflectance standard materials include pressed PTFE and ceramic tiles, the former is known to exhibit significant translucence [33] and the latter only comes in fixed sizes and is hard to machine.

Two materials were assessed for their possible use.

1) Pyroceram 9606, a polycrystalline glass ceramic material from Corning Inc., used for thermal conductivity and diffusivity reference materials [34], was considered. It is composed of mixed phases of oxides of silicon, aluminum, magnesium, and titanium, with some residual glassy phase, increasing opacity. It is a hard, refractory material, which can be machined to complex geometries.
2) Building on the idea of a processed aluminum substrate, a PEO coating on aluminum was tested. This process is becoming more widely used for producing very hard-wearing, corrosion-resistant coatings on light metal substrates, similar to anodizing, but using higher voltages and current densities to generate discharges at the surface that build up a denser oxide coating. A number

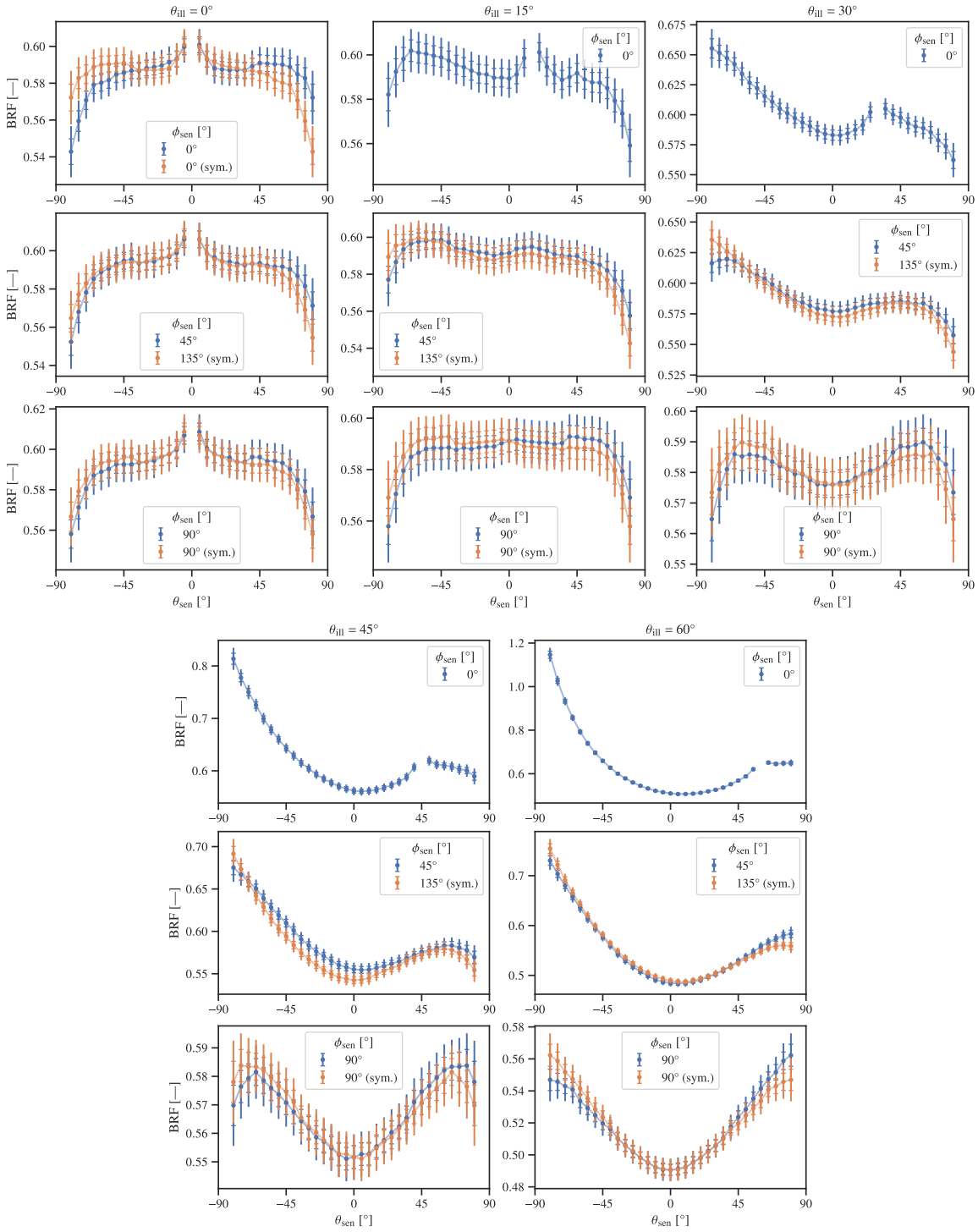


Fig. 18. Full material optical measurement dataset ($\lambda = 500\,\text{nm}$) faceted against the illumination zenith dimension. This figure complements Figs. 10 and 11. Each plot shows BRF records for a sensor azimuth value and, when relevant, the corresponding symmetrical configuration, assuming that the material is isotropic. Horizontal error bars report the standard uncertainty ($k = 1$); the vertical tips report the double (equivalent to $k = 2$). Configurations that are expected to be symmetrical agree qualitatively within the uncertainty ($k = 2$) in general, with a few exceptions. A systematic drift against the sensor azimuth leads the baseline (blue) to be below the symmetric (orange) for negative zenith values, and above for positive values. The sensitivity of this drift to the zenith angle makes alignment issues a plausible explanation. Another possible explanation is that the material would not be perfectly isotropic.

of trial samples were obtained from Manchester University using slightly different process conditions [35], and their spectral BRDF was assessed for their suitability, with the best response found for a coating that was around 50 ± 10 μm thick.

The Pyroceram was found to have some anisotropic reflectance, which might be due to the cutting process, resulting in fine directional surface texture or from inherent anisotropic microstructure, whereas once the PEO coating was thick enough (approx. 50 μm), the response was found to be

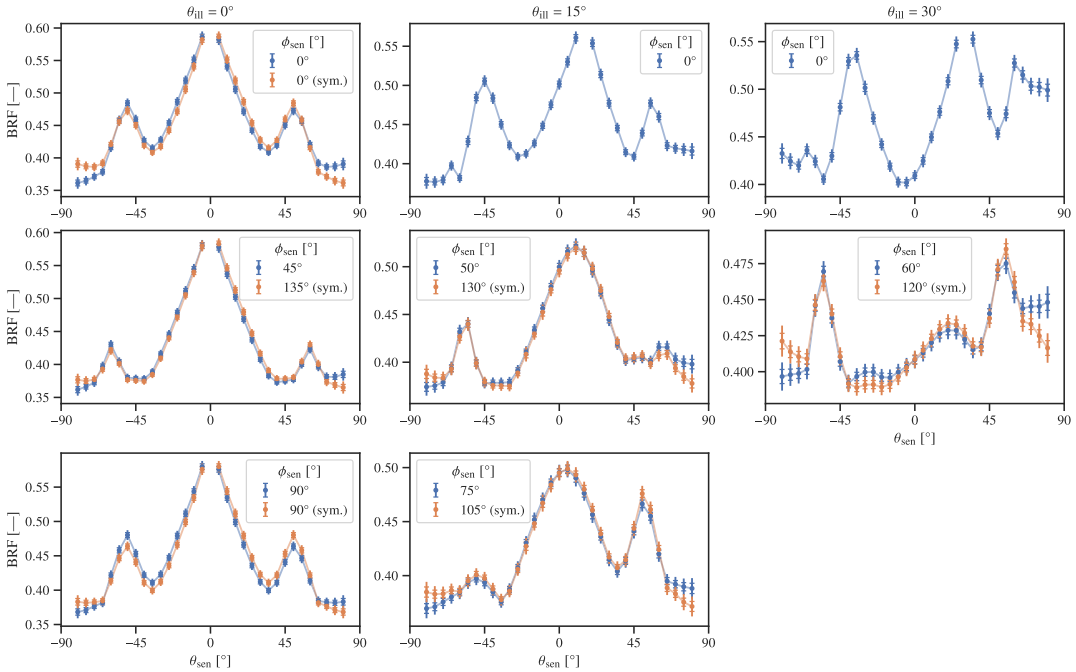


Fig. 19. Full artificial target optical measurement dataset ($\lambda = 500\,\text{nm}$) faceted against the illumination zenith dimension. This figure complements Figs. 12 and 13. Each plot shows BRF records for a sensor azimuth value and, when relevant, the corresponding symmetrical configuration based on symmetries identified in the design phase (see Fig. 4). Horizontal error bars report the standard uncertainty ($k = 1$); the vertical tips report the double (equivalent to $k = 2$). Configurations that are expected to be symmetrical agree qualitatively within the uncertainty ($k = 2$) in general in the $\theta_{\text{sen}} \in [-60°, 60°]$ range, with a few exceptions. A systematic drift against the sensor azimuth leads the baseline (blue) to be above the symmetric (orange) for negative zenith values, and below for positive values. The drift is amplified at high zenith angles ($\theta_{\text{sen}} > 60°$) and leads to high differences. As for material measurements, a plausible explanation for this sensor zenith-dependent drift is instrument alignment issues.

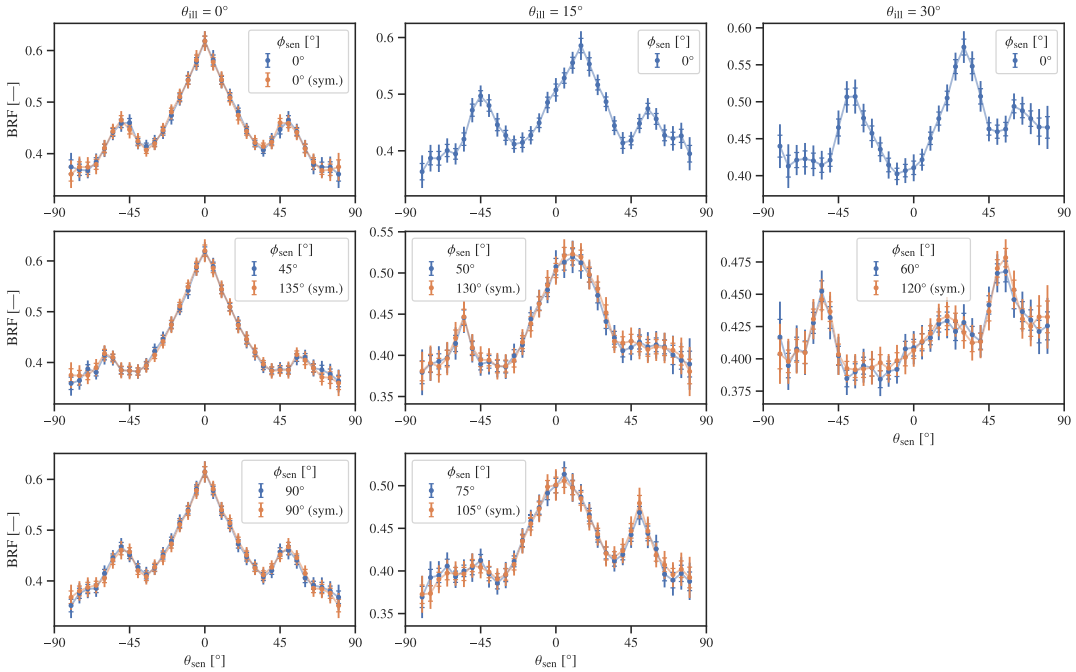


Fig. 20. Full numerical simulation dataset ($\lambda = 500\,\text{nm}$) faceted against the illumination zenith dimension. This figure complements Fig. 13. Each plot shows BRF records for a sensor azimuth value and, when relevant, the corresponding symmetrical configuration based on symmetries identified in the design phase (see Fig. 4). Horizontal error bars report the standard uncertainty ($k = 1$); the vertical tips report the double (equivalent to $k = 2$). Configurations that are expected to be symmetrical all agree within the uncertainty ($k = 2$).

quite close to Lambertian and isotropic. Precision machining of aluminum being a lot simpler than ceramic enabled greater freedom in the design parameters (e.g., hole size and density).

The test artifacts were made from 6086 aluminum alloy in three parts: a 1 mm thick top plate featuring a 12 × 12 array of nominally 2.4 mm diameter holes, nominally spaced 3.6 mm

on a square grid, and a 2 mm thick spacer that had a 4 mm edge width and a 5 mm thick parallel sided block [see Fig. 5(a)]. All the samples had their main faces precision ground to minimize flatness and thickness variations. Witness coupons were also provided for coating evaluation.

The test samples were PEO coated by Manchester University in small batches, with all the parts for each final artifact being coated together, to ensure the same coating performance across all the surfaces [see Fig. 5(b)]. As with conventional diffuse reflectance standards and especially as the surface is slightly rough the samples must be handled with gloves; only a light compressed air dusting was undertaken, as removing any surface contamination could be hard and may affect the optical properties.

Appendix C
Complete Datasets

For interested readers, we provide a global view of the following datasets.

1) Material optical measurements (see Fig. 18).
2) Artificial target optical measurements (see Fig. 19).
3) Numerical simulations (see Fig. 20).

For reproducibility purposes, material and artifact measurement data are also available online[6] [36].

Acknowledgment

The authors would like to thank Dr. A. Yerokhin and Dr. A. Rogov of Manchester University, Manchester, U.K., for the supply of the plasma electrolytic oxidation (PEO) coatings used in this article.

References

[1] H. Boesch et al., "Si-traceable space-based climate observation system: A CEOS and GSICS workshop," Nat. Phys. Lab., London, U.K., Tech. Rep. 9319, 2022.

[2] B. Pinty et al., "Radiation transfer model intercomparison (RAMI) exercise," *J. Geophys. Res., Atmos.*, vol. 106, no. 11, pp. 11937–11956, Jun. 2001.

[3] B. Pinty et al., "Radiation transfer model intercomparison (RAMI) exercise: Results from the second phase," *J. Geophys. Res., Atmos.*, vol. 109, no. D6, Mar. 2004, Art. no. D06210. [Online]. Available: https://discovery.ucl.ac.uk/65145/1/2003JD004252.pdf

[4] J.-L. Widlowski et al., "Third radiation transfer model intercomparison (RAMI) exercise: Documenting progress in canopy reflectance models," *J. Geophys. Res., Atmos.*, vol. 112, no. 9, May 2007, Art. no. D09111.

[5] J.-L. Widlowski et al., "The fourth phase of the radiative transfer model intercomparison (RAMI) exercise: Actual canopy scenarios and conformity testing," *Remote Sens. Environ.*, vol. 169, pp. 418–437, Nov. 2015.

[6] Y. M. Govaerts and M. M. Verstraete, "Raytran: A Monte Carlo ray-tracing model to compute light scattering in three-dimensional heterogeneous media," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 2, pp. 493–505, Mar. 1998.

[7] P. Jaanson et al., "Toward Si traceability of a Monte Carlo radiative transfer model in the visible range," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1360–1373, Mar. 2018.

[8] M. M. Bagher, C. Soler, and N. Holzschuch, "Accurate fitting of measured reflectances using a shifted gamma micro-facet distribution," *Comput. Graph. Forum*, vol. 31, no. 4, pp. 1509–1518, Jun. 2012. [Online]. Available: https://inria.hal.science/hal-00702304/document

[6]Access the Zenodo record at https://zenodo.org/doi/10.5281/zenodo.11637163

[9] M. M. Bagher, J. Snyder, and D. Nowrouzezahrai, "A non-parametric factor microfacet model for isotropic BRDFs," *ACM Trans. Graph.*, vol. 35, no. 5, pp. 1–16, Sep. 2016, doi: 10.1145/2907941.

[10] N. Holzschuch and R. Pacanowski, "A two-scale microfacet reflectance model combining reflection and diffraction," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–12, Aug. 2017. [Online]. Available: https://inria.hal.science/hal-01515948/document

[11] J. Löw, J. Kronander, A. Ynnerman, and J. Unger, "BRDF models for accurate and efficient rendering of glossy surfaces," *ACM Trans. Graph.*, vol. 31, no. 1, pp. 1–14, Jan. 2012, doi: 10.1145/2077341.2077350.

[12] A. Ngan, F. Durand, and W. Matusik, "Experimental analysis of BRDF models," in *Proc. Eurographics Symp. Rendering*, Jun. 2005, pp. 117–126.

[13] A. Sztrajman, G. Rainer, T. Ritschel, and T. Weyrich, "Neural BRDF representation and importance sampling," *Comput. Graph. Forum*, vol. 40, no. 6, pp. 332–346, Sep. 2021. [Online]. Available: https://discovery.ucl.ac.uk/id/eprint/10131304/1/cgf.14335.pdf

[14] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis, *Geometrical Considerations and Nomenclature for Reflectance*, vol. 160. Washington, DC, USA: U.S. Department of Commerce, National Bureau of Standards, 1977.

[15] V. Leroy, Y. Nollet, S. Schunke, N. Misk, and Y. Govaerts. (Aug. 2023). *Eradiate Radiative Transfer Model*. [Online]. Available: https://github.com/eradiate/eradiate

[16] W. Jakob et al. (2022). *Mitsuba 3 Renderer*. [Online]. Available: https://mitsuba-renderer.org

[17] K. E. Torrance and E. K. Sparrow, "Theory for off-specular reflection from roughened surfaces," *J. Opt. Soc. Amer.*, vol. 57, no. 9, pp. 1105–1114, 1967.

[18] S. Nevas, F. Manoocheri, and E. Ikonen, "Gonioreflectometer for measuring spectral diffuse reflectance," *Appl. Opt.*, vol. 43, no. 35, p. 6391, 2004.

[19] D. Lanevski, A. Bialek, E. Woolliams, F. Manoocheri, N. Fox, and E. Ikonen, "Gonioreflectometric properties of the sand from RadCalNet Gobabeb test site," in *Proc. Int. Conf. New Develop. Appl. Opt. Radiometry*, 2021, pp. 1–2.

[20] J. Fu et al., "Development of primary reference facilities and measurement comparison of standard artifacts for the bidirectional transmittance distribution function," *Rev. Scientific Instrum.*, vol. 94, no. 8, Aug. 2023, Art. no. 083102. [Online]. Available: https://pubs.aip.org/aip/rsi/article-pdf/doi/10.1063/5.0151803/18079089/083102_1_5.0151803.pdf

[21] D. Lanevski, F. Manoocheri, and E. Ikonen, "Gonioreflectometer for measuring 3D spectral BRDF of horizontally aligned samples with traceability to Si," *Metrologia*, vol. 59, no. 2, Apr. 2022, Art. no. 025006. [Online]. Available: https://iopscience.iop.org/article/10.1088/1681-7575/ac55a7/pdf

[22] H. Andrade and S. Kovner, "Software synthesis from dataflow models for G and LabVIEW™," in *Proc. Conf. Rec. 32nd Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2002, pp. 1705–1709.

[23] M. Kintel and C. Wolf. (2014). *OpenSCAD, the Programmers' Solid 3D CAD Modeller*. [Online]. Available: https://openscad.org

[24] C. Sullivan and A. Kaszynski, "PyVista: 3D plotting and mesh analysis through a streamlined interface for the visualization toolkit (VTK)," *J. Open Source Softw.*, vol. 4, no. 37, p. 1450, May 2019. [Online]. Available: https://joss.theoj.org/papers/10.21105/joss.01450.pdf

[25] P. D. Vis. (2023). *Punpy*. [Online]. Available: https://github.com/comet-toolkit/punpy

[26] P. Jaanson, "Angle-resolved measurements and modelling of diffuse reflectance and luminescence," Ph.D. dissertation, Dept. Signal Process. Acoust., Aalto Univ., Espoo, Finland, 2017.

[27] J. Dupuy and W. Jakob, "An adaptive parameterization for efficient material acquisition and rendering," *ACM Trans. Graph.*, vol. 37, no. 6, pp. 1–14, Dec. 2018.

[28] A. M. Rabal et al., "Automatic gonio-spectrophotometer for the absolute measurement of the spectral BRDF at in- out-of-plane and retroreflection geometries," *Metrologia*, vol. 49, no. 3, pp. 213–223, Jun. 2012.

[29] B. Bernad, A. Ferrero, A. P. Aglio, M. L. Hernanz, and J. C. Acosta, "Upgrade of goniospectrophtometer GEFE for near-field scattering and fluorescence radiance measurements," *Proc. SPIE*, vol. 9398, Mar. 2015, Art. no. 93980E.

[30] N. Fox and P. Green, "Traceable radiometry underpinning terrestrial- and helio-studies (TRUTHS): An element of a space-based climate and calibration observatory," *Remote Sens.*, vol. 12, no. 15, p. 2400, Jul. 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/15/2400/pdf?version=1595847322

[31] E. Veach, "Robust Monte Carlo methods for light transport simulation," Ph.D. dissertation, Dept. Comput. Sci., Stanford Univ., Stanford, Stanford, CA, USA, 1997. [Online]. Available: https://graphics.stanford.edu/papers/veach_thesis/thesis-bw.pdf

[32] A. A. Lacis and V. Oinas, "A description of the correlated K distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres," *J. Geophys. Res., Atmos.*, vol. 96, no. 5, pp. 9027–9063, May 1991.

[33] L. Gevaux, D. Saha, and G. Obein, "Investigating the optical translucency of spectralon using BSSRDF measurements," *Appl. Opt.*, vol. 62, no. 18, p. 5003, 2023.

[34] D. R. Salmon and R. P. Tye, "Pyroceram 9606, a certified ceramic reference material for high-temperature thermal transport properties: Part 1—Material selection and characterization," *Int. J. Thermophys.*, vol. 31, no. 2, pp. 338–354, Feb. 2010.

[35] A. B. Rogov, A. Nemcova, T. Hashimoto, A. Matthews, and A. Yerokhin, "Analysis of electrical response, gas evolution and coating morphology during transition to soft sparking PEO of Al," *Surf. Coatings Technol.*, vol. 442, Jul. 2022, Art. no. 128142.

[36] L. Vincent, S. Schunke, Y. Govaerts, R. Aschan, F. Manoocheri, and P. Woolliams, "An SI-traceable protocol for the validation of radiative transfer model-based reflectance simulation: Datasets (version 3)," Zenodo, Tech. Rep., Jun. 2024, doi: 10.5281/zenodo.11637163.

Vincent Leroy received the Ph.D. degree in thermal engineering, specializing in the modeling and simulation of multiscale coupled heat transfer in porous media at high temperatures from École Centrale Paris (now CentraleSupélec), Gif-sur-Yvette, France, in 2013.

He was a Post-Doctoral Researcher, he applied has expertise to protective materials for the atmospheric entry of spacecraft at the Institute of Condensed Matter Chemistry of Bordeaux, Pessac, France, and the von Karman Institute for Fluid Dynamics, Sint-Genesius-Rode, Belgium. He joined Rayference, Brussels, Belgium, in 2017, transitioning to the field of remote sensing, and a Lead Developer of the Eradiate Radiative Transfer Model. Since 2023, he has been the Technical Director with Rayference.

Robin Aschan received the B.Sc. and M.Sc. degrees in electrical engineering from Aalto University, Espoo, Finland, in 2018 and 2019, respectively, with a focus on micro and nanoscience, where he is currently pursuing the Ph.D. degree in electrical engineering with the Metrological Research Institute.

His research interests include the design and operation of a facility for bidirectional reflectance/transmittance distribution function (BRDF/BTDF) measurements and optical metrology of various materials, including reflective and transmissive diffusers, and materials with smooth, roughened, and structured surfaces.

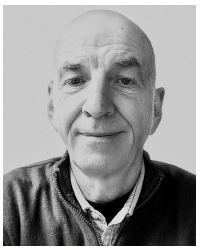
Peter Woolliams (Member, IEEE) received the M.Sc. degree in physics from the Imperial College, London, U.K., in 1998.

He has worked at the U.K. National Measurement Institute, National Physical Laboratory, Teddington, U.K., in several groups providing new measurement solutions in areas such as nonlinear fiber optics, diffuse reflectance, biophotonics, piezo-ceramics and most recently additive manufacturing (3-D-printing). He is currently a Senior Scientist with the Advanced Engineering Materials team supporting companies with novel materials measurement solutions.

Sebastian Schunke received the B.Sc. and M.Sc. degrees in physics from Ludwig-Maximilians Universität Munich, Munich, Germany, in 2015.

He was a Test Engineer at Bertrand Ingenieurbüro, Ehningen, Germany, for three years, he joined Rayference, Brussels, Belgium, in 2019, as a Research Engineer. He is part of the core development team of the Eradiate radiative transfer model. His research interests include radiative transfer simulations for Earth observation applications, with a specific emphasis on the modeling of 3-D surfaces.

Farshid Manoocheri, photograph and biography not available at the time of publication.



Yves Govaerts received the Ph.D. degree in physics from the Joint Research Centre, Ispra, Italy, in 1995, specializing in 3-D radiative transfer modeling in plant canopies.

He was an ESA Post-Doctoral Researcher, he contributed significantly to the development of a novel vegetation index tailored for the MERIS Instrument. Following this, he worked for 12 years for EUMETSAT, Darmstadt, Germany, where he developed a coupled surface-atmosphere radiative transfer model used for radiometric vicarious calibration. He also worked on the development of algorithms for the joint retrieval of surface reflectance and aerosol from space observations. In 2012, he founded Rayference, Brussels, Belgium, an SME, Brussels, providing expertise in radiative transfer modeling applied to Earth Observation.