Uncertainty Budget for a Traceable Operational Radiometric Calibration of Field Spectroradiometers, Calibration of the Heliosphere

Mike Werfeli, Andreas Hueni, *Member, IEEE*, Dessislava Ganeva, Giulia Ghielmetti, and Laura Mihai, *Member, IEEE*

***Abstract*—To measure the distinct interaction of the Earth's materials with solar electromagnetic radiation, field spectroradiometers are commonly utilized. These are used to validate spectroradiometers deployed on various platforms through comparison exercises. Following metrology standards, the inclusion of uncertainties is required. Thus, field spectroradiometers need to be calibrated regularly against traceable radiance sources. In this article, we present a laboratory radiometric calibration protocol for the calibration of a heliosphere integrating sphere to make it traceable to the International System of Units as well as to establish an uncertainty budget. We adopted a transfer radiometer approach including four spectroradiometers that were calibrated at the Deutsches Zentrum für Luft und Raumfahrt Radiometric Standard facility before transferring that calibration to the heliosphere. After considering various sources of uncertainty by employing an uncertainty tree diagram approach, we arrive at an overall propagated uncertainty of approximately 1.5%. In future publications, we will present how to extend the traceability to other attenuations provided by the heliosphere. Its application to the calibration of a field spectroradiometer will be the focus of a future publication.**

***Index Terms*—Calibration, radiometry, spectroscopy, uncertainty.**

Manuscript received 21 August 2023; revised 20 October 2023 and 31 October 2023; accepted 31 October 2023. Date of publication 3 November 2023; date of current version 23 November 2023. The work of D. Ganeva was supported by the Bulgarian Ministry of Education and Science under the National Research Program "Young Scientists and Postdoctoral Students-2" approved by DCM 206/07.04.2022. The work of L. Mihai was supported through a grant from the Romanian Ministry of Research, Innovation, and Digitization, CCCDI–UEFISCDI, under Project PN-III-P2-2.1-PED-2019-1729, within PNCDI III. The work was supported in part by the EMPIR Program by the Participating States and the European Union's Horizon 2020 Research and Innovation Program under Project 19ENV07 MetEOC-4, and in part by the COST Action CA17134 "Optical for spatiotemporal SENsing of Scalable ECOphysiological traits" (SENSECO) for networking and dissemination. *(Mike Werfeli and Andreas Hueni contributed equally to this work.) (Corresponding author: Mike Werfeli.)*

Mike Werfeli, Andreas Hueni, and Giulia Ghielmetti are with the Remote Sensing Laboratories, University of Zurich, 8057 Zurich, Switzerland (e-mail: mike.werfeli@geo.uzh.ch; andreas.hueni@geo.uzh.ch; giulia.ghielmetti@geo.uzh.ch).

Dessislava Ganeva is with the Space Research and Technology Institute, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria (e-mail: dganeva@space.bas.bg).

Laura Mihai is with the CETAL, Photonic Investigations Laboratory, National Institute for Laser, Plasma and Radiation Physics, 77125 Măgurele, Romania (e-mail: laura.mihai@inflpr.ro).

Digital Object Identifier 10.1109/JSTARS.2023.3329769

I. Introduction

THE Earth's surface and its materials interact with incident solar electromagnetic radiation in unique ways that allow for the analysis of biological, chemical, and physical properties of the Earth's surface and thus its environment [1], [2], [3]. The reflectance and absorption features of the electromagnetic spectrum enable the identification and characterization of different materials by measuring the reflected solar radiation [4], [5]. Spectral radiance data are collected by imaging or point spectroradiometers at different scales, using space-based and airborne platforms or in situ appliances [6], [7].

Field spectroradiometers (FSs) are used to measure single features with a very high spectral resolution [in both full-width half max and spectral sampling interval (SSI)] and are thus used to validate and calibrate space-based and airborne data through comparison exercises [8]. Sound metrological methodology requires the inclusion of uncertainty budgets as an integral part of such comparisons. Therefore, it is of utmost importance that the FSs are calibrated on a regular basis and that the uncertainty budget is estimated for every calibration. This further enables the provision of uncertainty budgets for every field spectroscopy measurement and thus follows established practices in metrology. As part of establishing the uncertainty budget, the different sources of uncertainty are identified and quantified, either analytically or statistically, assuming all distributions approach a normal distribution. The quantified single uncertainties are then propagated to the final uncertainty by adding the weighted uncertainties in quadrature [9]. In essence, radiometric measurements must include a propagated uncertainty budget traceable to the International System of Units (SI) [10]. This forms the basis to later estimate the uncertainties of reflectance factors calculated from in situ radiance measurements. The characterization, calibration, and associated uncertainty estimation for spectroradiometric measurements are mostly done in a laboratory.

FS laboratory calibration consists of two parts: 1) a spectral calibration, followed by 2) a radiometric calibration. The spectral calibration establishes the center wavelength per detector element as a minimum but should also provide the spectral sensitivity per element in the form of a spectral response function. Recent algorithms allow the retrieval of such response functions beyond their traditional Gaussian parameterization [11]. The subsequent radiometric calibration is used to establish the relationship between the at-sensor radiance (L) and the

corresponding digital numbers (DN) recorded by the instrument. The radiometric calibration is, thus, encapsulating the FS sensor model, which describes the transformation from DN to L considering the optical transmittance, the detector quantum efficiency (QE), the electronic noise during read out, the amplification gain, and the integration time, assuming a linear sensor response to radiance change [12]. Radiometric calibration coefficients and sensor linearity are commonly determined with laboratory experiments by pointing the spectroradiometer at a source of known radiance, such as a calibrated integrating sphere [12]. In addition, the uncertainties of the FS calibration can be estimated, assuming linearity and uniformity of the radiation source [13].

For in-house FS calibration, a well-calibrated and characterized integrating sphere or a lamp and reference panel setup is required. At the Remote Sensing Laboratories (RSL), Department of Geography, University of Zurich, an integrating sphere (Labsphere, Inc., S/N 0815196734) with a quartz tungsten halogen (QTH) lamp and an integrated diaphragm for radiance intensity manipulation are available (from here on we refer to the integrating sphere by its product name "heliosphere"). However, this heliosphere needs to be calibrated, traceable to SI, and characterized to estimate the uncertainties in radiance introduced by the calibration. For this purpose, two different calibration approaches exist: 1) a laboratory calibration [14] or 2) a transfer radiometer (XFR) calibration [15], [16], [17]. With the second approach, an XFR is calibrated against a light source calibrated to a national standard [e.g., National Institute of Standards and Technology or Physikalisch – Technische Bundesanstalt (PTB)] and is subsequently used to calibrate the targeted integrating sphere [15].

In this article, we present a method for an in-house heliosphere calibration following the best practice of metrology as described in the Guide to the Expression of Uncertainty in Measurement (GUM) [9]. More specifically, we aim to contribute to the radiometric calibration of FS through the estimation of the propagated uncertainty budget related to the traceable Heliosphere radiance.

II. Methods

A. Transfer Radiometer Calibration

To calibrate and characterize the heliosphere, an XFR approach (see Fig. 1) was chosen. Four different FS [one spectral evolution (SE), one spectra vista company (SVC), and two analytical spectral devices (ASD I (FS4) and ASD II (FS3))] were calibrated against the Deutsches Zentrum für Luft und Raumfahrt (DLR) RAdiometric STAndard (RASTA) [18] at the DLR Calibration Home Base (CHB) facility. RASTA is a calibration system, which consists of a halogen lamp, illuminating an isotropic–diffuse reflectance panel at a set distance, with the entire system being calibrated against PTB standards and thus being traceable to SI. The heliosphere was calibrated at a radiance intensity matching the RASTA radiance and requires, therefore, the development of an attenuation transfer function (gray branch in Fig. 1) presented in a future publication. All measurements were carried out in laboratories under controlled climatic conditions (20 °C with air conditioning). The recorded data were processed and stored in the spectral information system SPECCHIO to enable a database-centric processing approach [19].

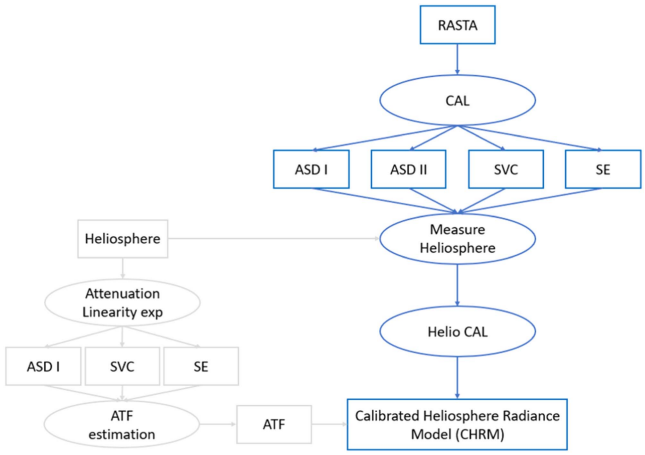

Fig. 1. Schematic overview of the XFR approach applied to calibrate the heliosphere radiance against the PTB-calibrated RASTA radiance source. The RASTA radiance (blue branch) is used to calibrate four spectroradiometers and use them as XFRs to transfer the PTB radiance standard to the Heliosphere. In this article, we are focusing on the blue branch.

The RASTA radiance was measured at the same time with: 1) the four XFRs and 2) five calibrated filter radiometers that are an integral part of RASTA to monitor its stability. The RASTA radiance data provided by PTB were interpolated with a spline fit to an SSI matching the SSI of the ASDs and SE. In addition, a radiometric calibration factor had to be applied to the RASTA data to compensate for the average radiometric drift of the RASTA lamp, which was observed with the RASTA filter radiometers. Given that the SVC data are saved with different units and a different SSI, these data had to be transformed to the common units $(\mathrm{W}/(\mathrm{m}^2 \cdot \mathrm{nm} \cdot \mathrm{sr}))$ and interpolated to match the SSI of the ASDs with the same spline method applied to interpolate the RASTA to the ASD SSI.

The XFRs were calibrated against RASTA by computing a per wavelength radiometric calibration factor (c_x) for each XFR

$$c_x = \frac{L_{\mathrm{RASTA}_{ip}} \cdot c_{\mathrm{drift}}}{L_{x_{\mathrm{RASTA}}}} \tag{1}$$

where $L_{\mathrm{Rasta}_{ip}}$ is the RASTA radiance interpolated to XFR center wavelengths, c_{drift} is a RASTA radiometry drift correction to compensate changes since its last calibration at PTB, and $L_{x_{\mathrm{Rasta}}}$ is the mean of N RASTA radiance measurements, measured with one individual XFR x (e.g., ASD I). From here onwards, we use x to denote that variables were acquired with or applied to one of the four XFRs. The drift correction factor c_{drift} is set to a value of 1 as no actual correction model is yet available. The drift observed by the RASTA filter radiometers is being used as an uncertainty of cdrift instead, as will be shown later.

B. Heliosphere Radiation Calibration

Transferring the RASTA radiance calibration to the heliosphere via the XFRs involves a number of assumptions, such as the temporal stability of the XFRs and their radiometric

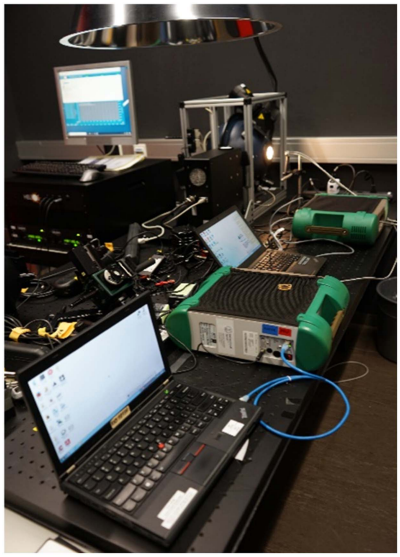

Fig. 2. Experimental setup while calibrating the heliosphere with ASD I and ASD II. The two ASDs were applied to transfer the calibrated DLR RASTA radiance to the heliosphere and thus make its radiance traceable to SI.

linearity. To minimize the impacts of the linearity assumption, the diaphragm was set to an attenuation of 76% to achieve a radiance intensity as close as possible to the RASTA radiance. Once the heliosphere attenuation was defined, 30 measurements were taken (see Fig. 2) with each XFR and averaged to obtain $L_{x_{\text{Helio}}}$. At this point, we introduced two correction factors used to capture the change of measured radiance: c_R represents the attenuation repeatability of the heliosphere and c_{TS} the temporal stability of the XFRs.

$$L_{x_{\text{Helio drift cor}}} = L_{x_{\text{Helio}}} \cdot c_R \cdot c_{\text{TS}}. \tag{2}$$

The two factors c_R and c_{TS} are assigned a value of 1 as no models exist to obtain actual correction factors. They are, however, important for the uncertainty propagation, as will be shown later.

The quantity $L_{x_{\text{Helio drift cor}}}$ was then radiometrically corrected with the respective RASTA radiometric calibration factor c_x to obtain $L_{\text{Helio}_{c(x)}}$

$$L_{\text{Helio}_{c(x)}} = L_{x_{\text{Helio drift cor}}} \cdot c_x. \tag{3}$$

Finally, an ensemble radiance $L_{\text{Helio}_{\text{ENS}}}$ was calculated from the mean of all four calibrated XFR radiances

$$L_{\text{Helio}_{\text{ENS}}} = \frac{1}{4} \sum_{x=1}^{4} L_{\text{Helio}_{c(x)}}. \tag{4}$$

C. Uncertainty Estimation and Propagation

This section represents the core of this study by detailing the establishment of a propagated uncertainty of the traceable heliosphere radiance with the four XFRs. We follow the steps and methods to establish uncertainty budgets as promoted by the MetEOC and FIDUCEO projects [18], [19]. Central to this approach is the creation of an uncertainty tree diagram [20], [21].

The following sections detail the uncertainty estimation and propagation of these two main blocks. Also, they show the uncertainties that apply to the central equation as part of a plus zero term when computing the combined standard uncertainty.

D. RASTA CAL Uncertainty of the Transfer Radiometer Calibration Factor

The uncertainties related to the RASTA CAL are shown as three branches (brown, red, and purple in Fig. 3). The radiometric calibration factors c_x are calculated according to (1). This calculation includes the following sources of uncertainty.

1) $u(L_{\text{Rasta}})$: The uncertainty associated with the RASTA radiance as established by PTB [22].
2) $u(L_{\text{Rasta}_{ip}})$: The uncertainty associated with the interpolation of the RASTA radiance vector to match the center wavelengths of the XFRs. The interpolation propagates $u(L_{\text{Rasta}})$ to $u(L_{\text{Rasta}_{ip}})$, where we chose a Monte Carlo (MC) approach to estimate the uncertainty $u(\text{spline interp.})$ added by the interpolation process. As an attribution of $u(L_{\text{Rasta}})$ to systematic and random effects was not available, an assumption was made that 20% of $u(L_{\text{Rasta}})$ are caused by noise. Hence, we added 20% of $u(L_{\text{Rasta}})$ as noise to the RASTA radiance vector in the MC process and carried out 1000 spline interpolations. The per wavelength standard deviation of the spline fit realizations was then taken as the uncertainty of the interpolations. $u(L_{\text{Rasta}_{ip}})$ was obtained from the combination of $u(L_{\text{Rasta}})$ with the estimated spline interpolation uncertainty $u(\text{spline interp.})$. Hence, the total uncertainty of the brown branch, propagating to $u(L_{\text{Rasta}_{ip}})$ [in (1)] is estimated by adding their relative uncertainties (in the brown branch) in quadrature: $u(L_{\text{Rasta}_{ip}})^2 = u(L_{\text{Rasta}})^2 + u(\text{spline interp.})^2$, i.e., the estimated uncertainty of the interpolation is added to the original RASTA radiometric uncertainty.
3) $u(c_{\text{drift}})$: The uncertainty related to the drift of the RASTA radiance. c_{drift} is assumed to have a value of one and its uncertainty is estimated from the RASTA long-term stability monitored by the filter radiometers.
4) $u(L_{x_{\text{Rasta}}})$: The system noise associated with each FS measurement, which is assumed to be normally distributed, following the central limit theorem (CLT) [23]. It is estimated by computing the standard deviation of N radiance measurements and dividing by the square root of N as measurements are averaged for further computation.

These sources of uncertainty related to the XFR calibration are propagated according to the GUM by first calculating their sensitivity coefficients via the partial derivatives of (1) and estimating the uncertainties, where required, before adding the components in quadrature to compute the combined uncertainty $u(c_x)$

$$u(c_x)^2 = \left(\frac{c_{\text{drift}}}{L_{x_{\text{Rasta}}}}\right)^2 \cdot u\left(L_{\text{Rasta}_{ip}}\right)^2 + \left(\frac{L_{\text{Rasta}_{ip}}}{L_{x_{\text{Rasta}}}}\right)^2 \cdot u(c_{\text{drift}})^2$$

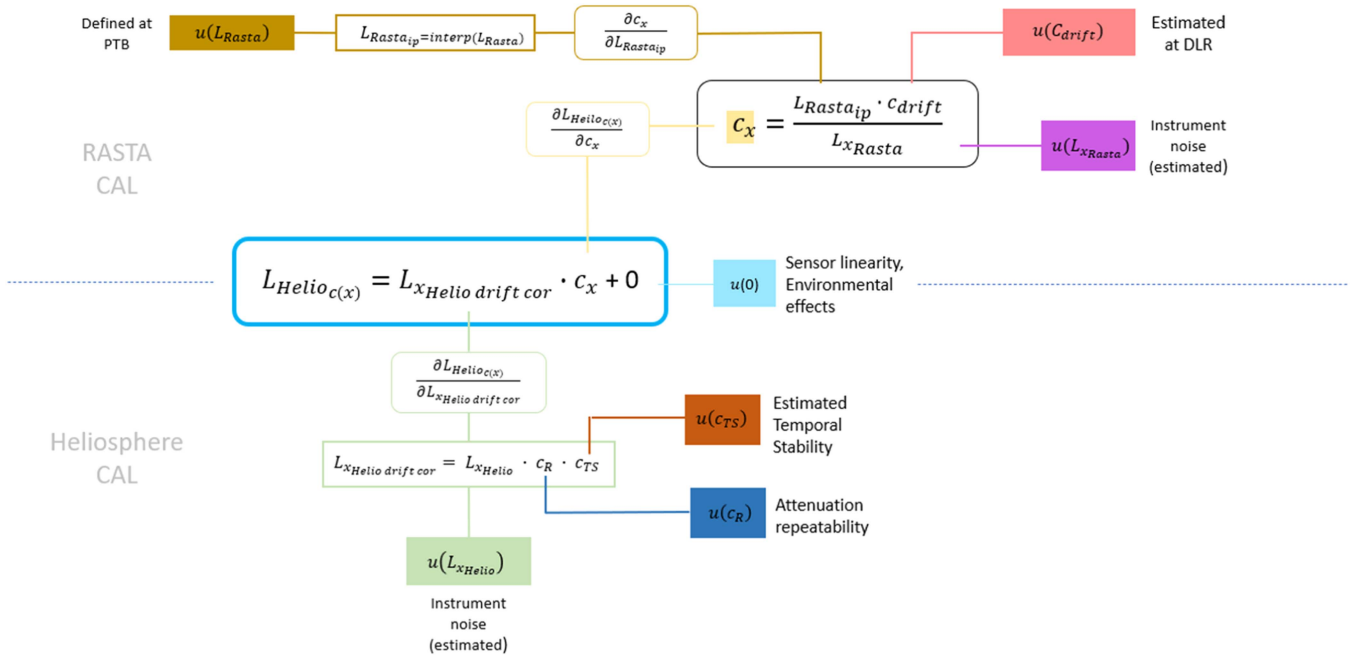

Fig. 3. Uncertainty tree diagram illustrating the different factors influencing the calibration uncertainty of the heliosphere radiance, originating in the two major steps of calibration: 1) XFR calibration at RASTA (RASTA CAL) and 2) heliosphere calibration with the XFRs (heliosphere CAL). In the RASTA CAL step (top half), the uncertainties are traceable to PTB and are propagated through to the calibration factor c_x applied to each XFR, respectively. In the heliosphere CAL step (bottom half), the heliosphere radiance is measured with the XFRs, which flows into the final equation (blue square) and is corrected with c_x to get the calibrated heliosphere radiance.

$$+\left(\frac{-c_{\text{drift}} \times L_{\text{Rasta}_{ip}}}{L^2_{x_{\text{Rasta}}}}\right)^2 \cdot u(L_{x_{\text{Rasta}}})^2. \tag{5}$$

Combining the uncertainties at this stage could be done, but the final goal of computing the uncertainty of the ensemble heliosphere radiance requires that $u(c_x)$ is separated into random and systematic uncertainty components. Only the random component $u(c_{x_{\text{random}}})$ will benefit from the ensemble approach, which allows to reduce the uncertainty, while the systematic component $u(c_{x_{\text{systematic}}})$ fully contributes to the ensemble uncertainty.

The systematic component includes the radiometric calibration of RASTA and the drift of RASTA over time monitored by the filter radiometers (6). Note that we choose at this stage to assume $u(L_{\text{Rasta}_{ip}})$ to be fully systematic.

$$u\left(c_{x_{\text{systematic}}}\right)^2 = \left(\frac{c_{\text{drift}}}{L_{x_{\text{Rasta}}}}\right)^2 \cdot u\left(L_{\text{Rasta}_{ip}}\right)^2 + \left(\frac{L_{\text{Rasta}_{ip}}}{L_{x_{\text{Rasta}}}}\right)^2 \cdot u(c_{\text{drift}})^2. \tag{6}$$

The random component is equivalent to the uncertainty added due to measurement noise while calibrating the XFRs against RASTA

$$u(c_{x_{\text{random}}})^2 = \left(\frac{-c_{\text{drift}} \cdot L_{\text{Rasta}_{ip}}}{L^2_{x_{\text{Rasta}}}}\right)^2 \cdot u(L_{x_{\text{Rasta}}})^2. \tag{7}$$

Hence, $u(c_{x_{\text{systematic}}})$ and $u(c_{x_{\text{random}}})$ are computed for each XFR separately to be combined later in the uncertainty of the ensemble Heliosphere calibration (9).

E. Heliosphere CAL Uncertainty of the Measurement of the Heliosphere With a Transfer Radiometer

Calibrating the heliosphere with the XFR introduces the following uncertainties.

1) $u(L_{x_{\text{Helio}}})$: System noise associated with the FS measurements of the heliosphere, which is assumed to be normally distributed according to the CLT and reduced by the square root of *N* due to using the average of 30 measurements.
2) $u(c_R)$: The uncertainty introduced by the mechanically based attenuation control. This attenuation repeatability was established experimentally by measuring the radiance of four different diaphragm settings (0%, 25%, 50%, and 75%) in 20 consecutive measurement rounds with 15 readings taken at each step. An ASD FieldSpec4 instrument was used to acquire these measurements over a time span of about 5 h in a laboratory with a nominal air temperature of 20 °C. This allowed us to partition the observed differences into instrument noise and attenuation repeatability. First, the mean values per measurement round per attenuation setting were computed. The standard deviation of these 20 per-round means was then taken to be the repeatability uncertainty. These per attenuation uncertainties were averaged across the shortwave infrared (SWIR I) channel, which was deemed the most stable, and applied to all wavelengths to avoid including instrument-related artifacts. For this article, only the 75% attenuation closely matching the calibration attenuation was used, but we report the results for the other attenuations as well due to the interesting conclusions that may be drawn from it.

3) $u(c_{\mathrm{TS}})$: The estimated temporal stability of the XFRs between the RASTA calibration and the heliosphere calibration. No experimental data to estimate this quantity were available. We, thus, assume an uncertainty associated with temporal stability of 1% over the course of several days, including the transport between facilities.

The uncertainty of the heliosphere radiometric measurement per XFR was calculated, based on (2), by analytical propagation

$$u(L_{x_{\mathrm{Helio\ drift\ cor}}})^2 = (c_R \cdot c_{\mathrm{TS}})^2 \cdot u(L_{x_{\mathrm{Helio}}})^2 + (L_{x_{\mathrm{Helio}}} \cdot c_{\mathrm{TS}})^2 \cdot u(c_R)^2 + (L_{x_{\mathrm{Helio}}} \cdot c_R)^2 \cdot u(c_{\mathrm{TS}})^2. \tag{8}$$

The correction factors c_R and c_{TS} are both set to a value of one while their uncertainties are nonzero. Hence, these factors are introduced with the sole purpose of enabling the propagation of their uncertainties as no data exist to provide actual values.

F. Combined Standard Uncertainty

The radiance of the calibrated heliosphere has a combined standard uncertainty $u(L_{\mathrm{Helio_{ENS}}})$, which is estimated by propagating and combining the four uncertainties associated with the single XFR measurements $u(L_{\mathrm{Helio}_{c(x)}})$. These instrument-specific XFR uncertainties are estimated by propagating $u(c_x)$ and $u(L_{x_{\mathrm{Helio\ drift\ cor}}})$ analytically, following the GUM [17]. As $u(c_x)$ includes both random and systematic components, we arrive at

$$u\left(L_{\mathrm{Helio}_{c(x)}}\right)^2 = c_x^2 \cdot u(L_{x_{\mathrm{Helio\ drift\ cor}}})^2 + (L_{x_{\mathrm{Helio\ drift\ cor}}})^2 \cdot u\left(c_{x_{\mathrm{systematic}}}\right)^2 + (L_{x_{\mathrm{Helio\ drift\ cor}}})^2 \cdot u(c_{x_{\mathrm{random}}})^2. \tag{9}$$

Further uncertainties related to XFR linearity and environmental factors (see Fig. 3) are captured by the zero term and currently assumed to be zero. To retain the partitioning into systematic and random effects, we split $u(L_{\mathrm{Helio}_{c(x)}})$ as follows: the combination of the four single XFR measurements into an ensemble heliosphere radiance results in the combination of their respective random and systematic uncertainties. Equation (9) explicitly labels the uncertainty components per XFR. The random components are weighted by sensitivity coefficients of $\frac{1}{4}$, while the systematic uncertainties are averaged between the four XFRs (10).

$$u(L_{\mathrm{Helio_{ENS}}})^2 = \left(\frac{1}{4}\right)^2 \cdot u\left(L_{\mathrm{Helio}_{c(\mathrm{ASD\ I})\ \mathrm{random}}}\right)^2 + \left(\frac{1}{4}\right)^2 \cdot u\left(L_{\mathrm{Helio}_{c(\mathrm{ASD\ II})\ \mathrm{random}}}\right)^2 + \left(\frac{1}{4}\right)^2 \cdot u\left(L_{\mathrm{Helio}_{c(\mathrm{SVC})\ \mathrm{random}}}\right)^2 + \left(\frac{1}{4}\right)^2 \cdot u\left(L_{\mathrm{Helio}_{c(\mathrm{SE})\ \mathrm{random}}}\right)^2 + \left(\frac{1}{4}\left(u\left(L_{\mathrm{Helio}_{c(\mathrm{ASD\ I})\ \mathrm{systematic}}}\right) + u\left(L_{\mathrm{Helio}_{c(\mathrm{ASD\ II})\ \mathrm{systematic}}}\right) + u\left(L_{\mathrm{Helio}_{c(\mathrm{SVC})\ \mathrm{systematic}}}\right) + u\left(L_{\mathrm{Helio}_{c(\mathrm{SE})\ \mathrm{systematic}}}\right)\right)\right)^2. \tag{10}$$

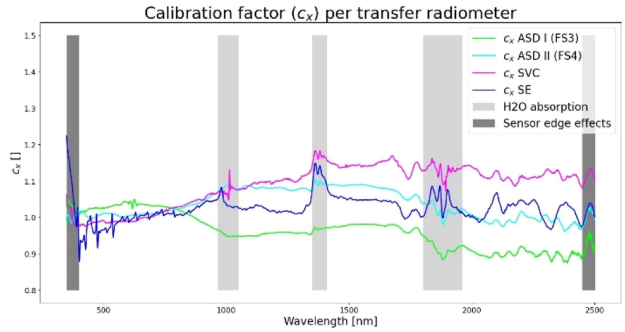

Fig. 4. Wavelength-dependent calibration factors of all XFRs. The light gray areas are the regions where water absorption features are found, whereas the dark gray areas delineate the sensor edges, where higher levels of noise are found.

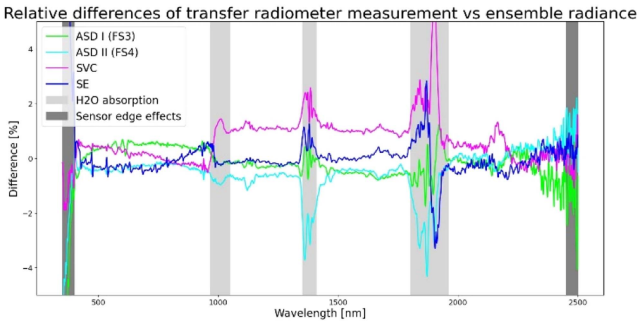

Fig. 5. Relative differences between XFR measurements and the ensemble heliosphere radiance.

III. Results

A. Radiometric Calibration Factors of Transfer Radiometers

The radiometric calibration factors established during the calibration of the XFRs against RASTA are shown in Fig. 4 and range roughly between 0.85 and 1.25. Three of the four XFRs (ASD FS4, SE, and SVC) increase toward the SWIR, whereas ASD FS3 shows an inverse trend, starting at its highest in the visible (VIS) and decreasing toward the SWIR. A higher divergence between factors is found in the SWIR. A high existing miscalibration is also evident in the UV bands of the SE.

B. Heliosphere Calibration

The RASTA radiometric calibration is transferred to the heliosphere by correcting the XFR Heliosphere radiance measurements with the respective c_x calibration factors before calculating the ensemble radiance. Fig. 5 shows the difference between the heliosphere radiance measured with the four XFR and the heliosphere ensemble radiance at calibration attenuation (76%). The greatest differences between XFR and ensemble radiance are found in the UV (350–400 nm) where the SE and the two ASD spectroradiometers are diametrically opposed. The positive delta calculated for the SE points to an overcorrection by the calibration factor c_x. This may in turn point to a radiometric instability in the UV between the measurements at RASTA and at RSL. Medium differences are apparent in the NIR (950–1000 nm) and in the SWIR (1400–1500 nm (SWIR I)

Fig. 6. Repeatability data for four attenuation settings, shown for selected wavelengths as time series with error bars indicating the measurement noise. Data have been normalized per measurement round.

and 1900–2000 nm (SWIR II)), which correspond to the water absorption bands in the region of 975, 1450, and 1950 nm and at the end of SWIR ($>$ 2400 nm) where the signal is noise limited.

The following descriptions ignore the above regions of higher deltas, i.e., they focus on the nonshaded areas of Fig. 5. In the VIS region, the four XFR measurements show a good agreement with the calibrated ensemble heliosphere radiance, with all sensors obtaining mean relative differences around $\pm 0.5\%$. In the SWIR I region, the SVC shows the greatest mean relative difference (1.14%), whereas the SE has the lowest mean relative difference of 0.02%. The mean differences of the ASDs are both smaller than -1% but differ from one another, with the FS3 being smaller than 0.5% and the FS4 greater than 0.5%.

In the SWIR II region, the differences are generally a bit smaller, with all spectra not being more than 0.5% different from the calibrated heliosphere ensemble radiance. In this range, the SE has the lowest mean relative difference (-0.12%), whereas the SVC has the highest (0.71%). Overall, the ASD FS3 is the closest to the calibrated heliosphere radiance with differences of less than 0.5% throughout the entire measured spectrum.

C. Attenuation Repeatability

The design of the repeatability experiment aimed at discriminating sensor measurement noise and attenuation repeatability, which defines the uncertainty in radiometry caused by the diaphragm component. We were only partly successful in achieving this goal, as our analysis shows that the presumed changes due to attenuation repeatability include sensor instabilities that occurred over the 5 h of measurement time. Fig. 6 shows the radiance time series grouped per attenuation level, where each mean per measurement round was normalized by the mean of all rounds. These time series are plotted for strategically selected wavelengths, taking into account our a priori knowledge about the behavior of ASD full-range instruments [24], [25]:

1) 350 nm is the start of the VNIR and typically has a low signal-to-noise ratio (SNR) due to lower QE combined with low radiances provided by QTH lamps.
2) 1000 nm represents the end of the VNIR channel and is characterized by lower SNR due to low QE compared to the middle of the VNIR detector and in addition displays the highest response to changes in environmental temperatures.

3) 1001 nm is the first band provided by the SWIR I detector and has been found to be radiometrically stable during environmental temperature tests.
4) 1801 nm is the first band of the SWIR II detector and is the most sensitive band of this channel to environmental temperature changes.
5) 2500 nm is the last band of the SWIR II channel and has typically low SNR values, both in field and laboratory illumination conditions.

Several effects can be seen in these time series, which are not governed by short-term system noise as can be observed by the small error bars representing the standard deviation of 15 readings per measurement round.

1) Bands at 1000 and 1801 nm reveal a radiometric drift over time. As has been shown in earlier studies, the radiometric response due to thermal changes is opposed in VNIR and SWIR II channels [24]. These drifts are visually not obvious for the 75% attenuation as they are overshadowed by midterm random radiometric instabilities in the 350 and 2500 nm bands.
2) Bands at 350 and 2500 nm are characterized by a midterm random noise that shows between consecutive measurement rounds while the short-term system noise is very low. This is most apparent at attenuations of 75% and 50%.
3) Starting from 50% attenuation, the range of normalized data values is defined by the minima and maxima of the 350 and 1801 nm bands. These ranges are virtually identical between attenuations of 0%, 25%, and 75%. This, in turn, means that the observed changes over time are governed by radiometric drifts and that the diaphragm repeatability must be established from bands that are neither limited by instabilities over time (350 and 2500 nm) nor from bands that show a clear indication of sensitivity to thermal changes (1000 and 1801 nm).

In the next step, the relative uncertainties per attenuation were computed based on the standard deviation of the radiance means per measurement round *R*

$$u\ (\text{attenuation repeatability}) = \frac{\text{std}\,(\text{mean}\,(L_R))}{\text{mean}\,(\text{mean}\,(L_R))} \cdot 100. \tag{11}$$

Fig. 7 spectrally illustrates the effects that were discussed above for selected bands shown as time series.

The VNIR channel shows higher uncertainties below 430 nm, which are linked to midterm sensor instabilities, and above 760 nm, where temperature-related radiometric changes were shown to start [24]. The SWIR I channel displays a steady uncertainty apart from the water vapor absorption region around 1380 nm. The SWIR II channel is characterized by higher uncertainties at the start of the detector (1801–1935 nm) while bands above 2250 nm gradually rise to their maximum uncertainty at 2500 nm, caused by SNR values dropping toward the end of the detector.

To constrain attenuation repeatability uncertainty to effects that are not instrument related, we select the mean of a spectral subset of the SWIR I channel (1001–1300 nm) to be representative for all spectral bands.

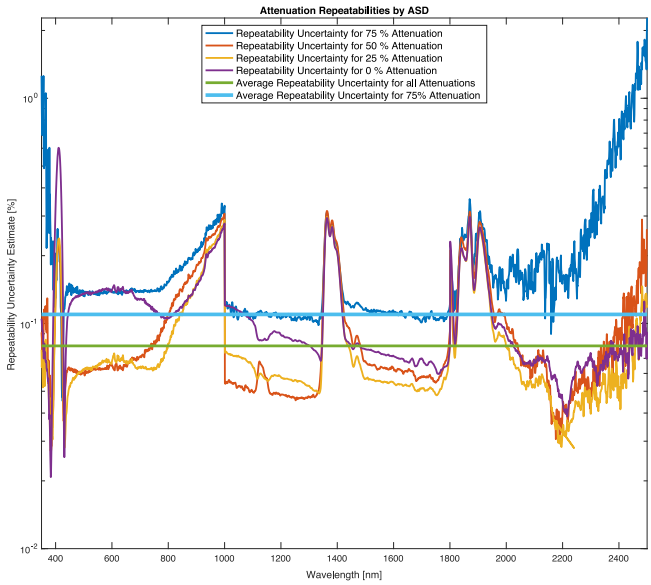

Fig. 7. Relative uncertainties per attenuation and the estimated mean uncertainties for all attenuations and attenuation at 75%, assuming a stable radiometric response of the SWIR I channel over the course of the repeatability experiment.

The resulting mean uncertainty over all attenuations assumes a value of 0.08% while the mean uncertainty for 75% attenuation is estimated at 0.1%.

D. Uncertainty Estimation and Propagation for Transfer Radiometer

The uncertainty was estimated for each XFR, following the uncertainty diagram, as given in Fig. 3. In this section, we describe the estimated and propagated uncertainties of one XFR (ASD II, FS4). In Fig. 8, the combined uncertainty $u(c_x)$ and its two components $u(c_{x_{\text{random}}})$ and $u(c_{x_{\text{systematic}}})$ are shown. The uncertainty due to $u(L_{x_{\text{Rasta}}})$ weighted by its sensitivity coefficients is identical to $u(c_{x_{\text{random}}})$ [see (7)] and describes the uncertainty caused by the XFR readout noise (yellow line). The peaks where the uncertainty estimate is the greatest are the two detector jumps at 1000 and 1800 nm as well as the sensor edges. The next component adding to $u(c_x)$ is the uncertainty associated with the radiance measurement of RASTA. After propagation through the radiation interpolation, the uncertainty $u(L_{\text{Rasta}_{ip}})$ accumulates to a value just above 0.5% in the VNIR and 1.3% in the SWIR (blue line in Fig. 8). What increases the RASTA uncertainty, however, is the drift of the RASTA lamp $u(c_{\text{drift}})$, which contributes especially in the VNIR and introduces an uncertainty of 1.2% (red line in Fig. 8). Finally, the systematic component (purple line in Fig. 8) is combining $u(L_{\text{Rasta}_{ip}})$ and $u(c_{\text{drift}})$, resulting in an almost similar uncertainty as $u(c_x)$. The combination of $u(c_{x_{\text{random}}})$ and $u(c_{x_{\text{systematic}}})$ results in the total calibration factor uncertainty $u(c_x)$ (green line in Fig. 8).

Fig. 9 shows the different components of the uncertainty propagating to the final uncertainty $u(L_{x_{\text{Helio drift cor}}})$, estimated for each XFR respectively. The smallest uncertainty is $u(L_{x_{\text{Helio}}})$ (blue line in Fig. 9), which is associated with the sensor read out noise and stays below 0.5% for the most part.

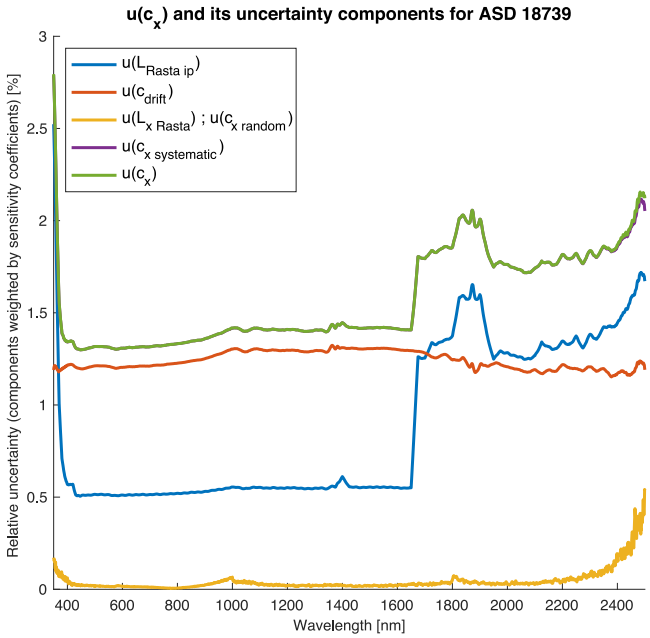

Fig. 8. Uncertainties related to the XFR calibration at RASTA, which are propagated and combined to the uncertainty of the wavelength-dependent and sensor-specific calibration factor c_x. This graphic shows the effects of the individual uncertainties propagated to a combined uncertainty.

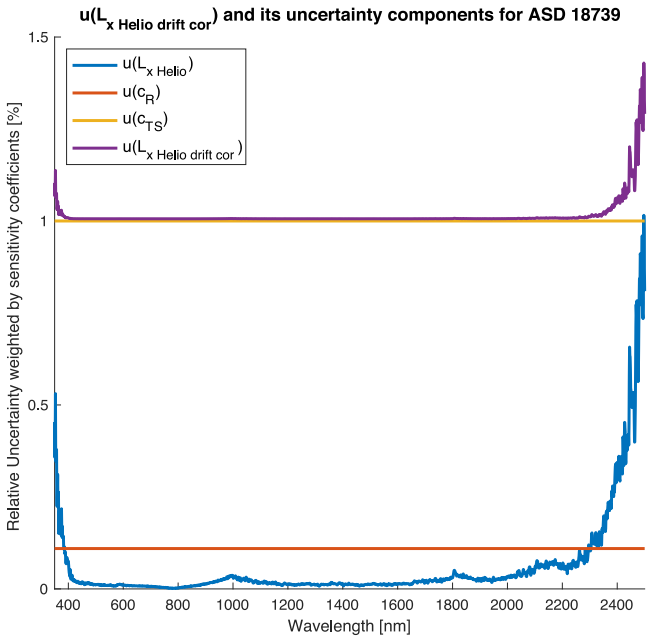

Fig. 9. Uncertainties of the heliosphere radiance measurement with the XFR ASD II. This graph shows the effects of the single uncertainties propagated to the final (purple) uncertainty of the heliosphere radiance measurement. Most evidently, is the contribution.

Higher uncertainties are found near the detector edges around 1000 and 1800 nm and, in particular, in the UV and SWIR 2 regions where the noise increases until it maximizes at the detector edges (close to 350 and 2500 nm). The uncertainty $u(c_R)$

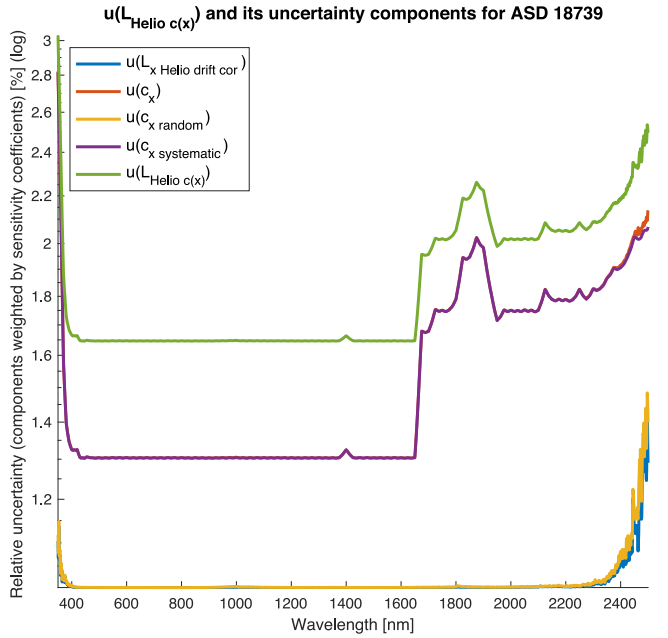

Fig. 10. Uncertainties contributing to the final calibrated heliosphere measurement for one specific XFR. This figure shows the contribution of uncertainties originating from the XFR calibration at RASTA (see Fig. 8) and the uncertainties related to the heliosphere measurement (see Fig. 9).

associated with the repeatability of the diaphragm experiment (red line in Fig. 9) is estimated to be close to 0.1%. In addition, a greater uncertainty is found in the data recorded by the SWIR II detector (0.7%). The last uncertainty added to $u(L_{x_{\text{Helio}}})$ is the uncertainty caused by the temporal stability of the XFR (orange line in Fig. 9), which in this case is assumed to be one. This uncertainty is at the same time the greatest contributor to the final combined standard uncertainty of the Heliosphere measurement.

The fully propagated uncertainty $u(L_{\text{Helio}_{c(x)}})$ for XFR ASD II and the components contributing to the sensor-related uncertainty are shown in Fig. 10. The two main uncertainty sources (RASTA calibration and Heliosphere measurement) are represented by the propagated uncertainty $u(c_x)$ (red line in Fig. 10), which is the combination of $u(c_{x_{\text{random}}})$ and $u(c_{x_{\text{systematic}}})$ (yellow and purple lines in Fig. 10) whereas, the uncertainties related to the Heliosphere measurement are represented by $u(L_{x_{\text{Helio drift cor}}})$ (blue line in Fig. 10).

The uncertainty $u(L_{x_{\text{Helio drift cor}}})$ represents the sensor characteristics and the heliosphere repeatability, contributing approximately 1% of uncertainty. The rest of the uncertainty is introduced by the calibration of the XFR at RASTA. The smallest contribution is added by $u(c_{x_{\text{random}}})$, which introduces the uncertainty driven by sensor noise during calibration (yellow line in Fig. 10). The systematic uncertainty $u(c_{x_{\text{systematic}}})$, representing the drift of the RASTA QTH lamp, is adding 1.3% to the uncertainty budget. These two components combine in the uncertainty of the calibration factor $u(c_x)$ (red line in Fig. 10). The combined uncertainty $u(L_{\text{Helio}_{c(x)}})$ (green line in Fig. 10) results in total uncertainty of the calibrated heliosphere measurement with ASD

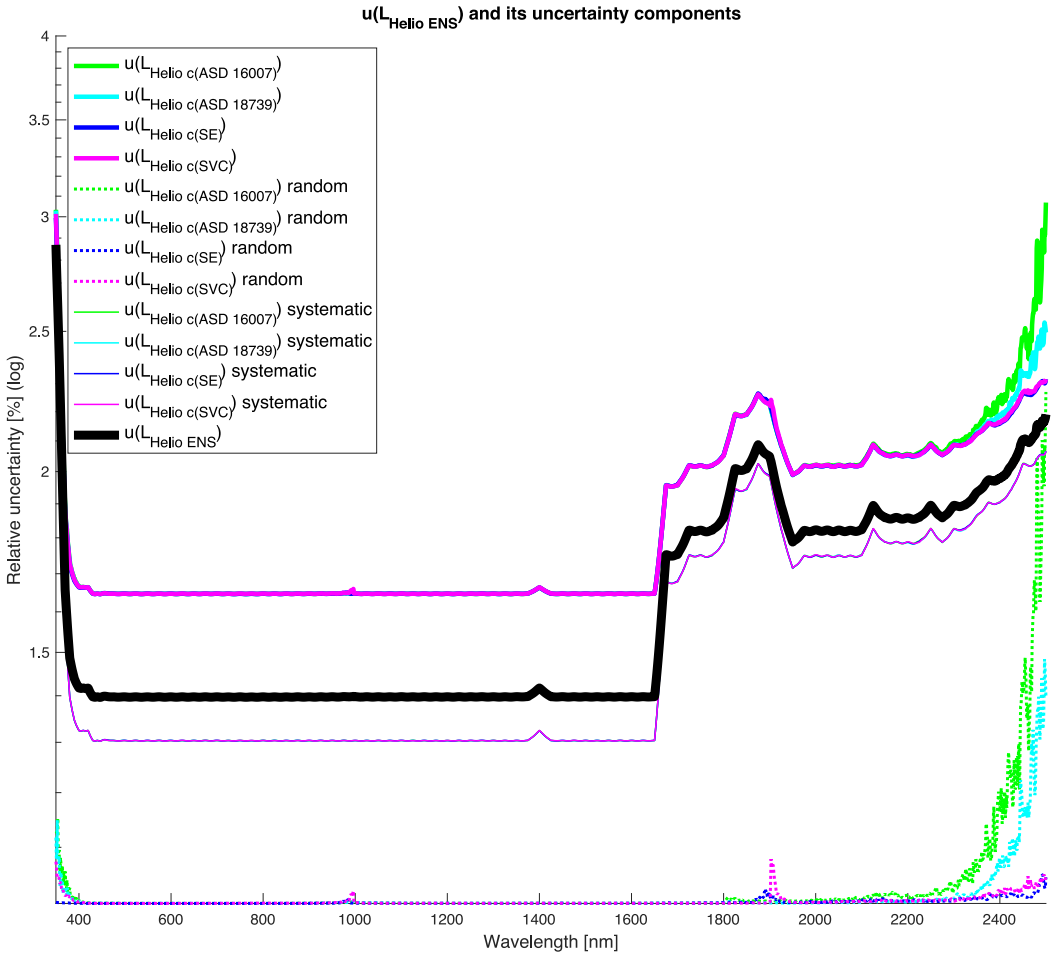

Fig. 11. Final propagated heliosphere radiance uncertainty (black line) and its components (different line styles) of each sensor applied as XFR (color coded). This plot shows how the single uncertainties of each XFR cumulate as they are propagated, whereas the combination of the four XRFs reduces the overall uncertainty.

II (FS4), which is mostly governed by systematic components with random noise influencing the uncertainties in the UV and at the end of SWIR II.

E. Heliosphere Ensemble Radiance

The uncertainties $u(L_{\mathrm{Helio}_{c(x)}})$ are estimated for every XFR, respectively (see Fig. 10). These uncertainties are propagated and combined to $u(L_{\mathrm{Helio}_{\mathrm{ENS}}})$. Fig. 11 shows the uncertainties contributing to the heliosphere ensemble radiance uncertainty (black line). Each XFR shown in Fig. 11 was given a color code, with ASD II FS4 discussed earlier being shown in cyan. Furthermore, the most central uncertainties, such as the uncertainties of the instrument-specific c_x components, are shown with different line styles but using the respective XFR color coding. An expected but nonetheless satisfying result is that the average uncertainty of the XFR ensemble ($u\ (L_{\mathrm{Helio}_{\mathrm{ENS}}}) = \ 1.6\%$) is 0.3% lower compared to the single propagated XFR uncertainties. This reduction in uncertainty is caused by the utilization of four different XFRs and demonstrates the reduction of the random uncertainties through averaging. Furthermore, the features found in the single XFR uncertainties are smoothed to some extent in $u(L_{\mathrm{Helio}_{\mathrm{ENS}}})$.

In the SWIR II region, the uncertainties increase drastically, especially for ASD I (FS3), which originates in the random component of the heliosphere measurement and is, thus, characterized by the XFR-specific noise.

IV. Discussion

Applying an XFR approach to calibrate the in-house heliosphere allows control over the process of calibration and helps establish both the traceability and the uncertainty budget. Utilizing four XFRs for heliosphere radiance calibration adds robustness to the approach and reduces the random noise component in the final ensemble radiance. There are, however, several shortcomings caused by different sources of error introducing uncertainties. One of these sources of uncertainty is the drift of the RASTA radiance source, which is monitored by filter radiometers but not corrected for as yet [9], [19]. A recalibration of RASTA at PTB or the development of a drift correction would decrease the uncertainty associated with the XFR calibration versus RASTA and, thus, improve the total propagated uncertainty. Similarly, the uncertainty introduced by the temporal stability of XFRs of 1% was only based on the best guess and is likely overestimated. Dedicated stability experiments could

help quantify and probably lower the influence of this source of uncertainty.

The random components of the RASTA measurement and the heliosphere measurement are related to the SNR of each XFR, which is defined by the optoelectronic characteristics of the instruments. Decreases in SNR in specific wavelength regions can be clearly linked to reduced quantum efficiencies of the detectors and to low signal intensities of the at-sensor radiance spectrum.

Establishing the repeatability of the integrating sphere attenuation revealed various sensor effects, showing that not all spectral bands are lending themselves to measuring the repeatability of the sphere. Instead, some wavelength regions showed clear impacts of signal drift due to the temperature-related efficiency of the read-out chip [20].

For calibration and uncertainty estimation, some assumptions had to be made, which had a significant influence on the total uncertainty budget. First, the environmental conditions were assumed to be stable throughout all measurements, adding only a little uncertainty to the total budget. Second, as pointed out above, the XFRs are thought to be temporally stable, assuming the internal optoelectronics have not experienced any drifts since the calibration factors are estimated at RASTA. As a conclusion, we find that the XFRs need to be characterized more thoroughly, including tests for various kinds of drifts (e.g., temporal, temperature, or pressure related), to make the associated uncertainties more precise and realistic.

Finally, the calibration discussed here was only performed at one attenuation level (76%), whereas the heliosphere has more attenuation settings and, consequently, more radiance intensities which need to be calibrated. In addition, the uncertainties associated with these different attenuation levels need to be established as well, requiring detailed characterizations of both heliosphere and XFR at the related radiance intensities [21].

V. Conclusion

With this work, we demonstrated how to perform an SI-traceable heliosphere calibration with a multiple XFR approach. Propagating the uncertainties clearly showed where the biggest sources of errors are and how to quantify and include them in the total uncertainty budget. Our well-documented approach should allow other users to take control of the traceability of their light sources, leading to a better understanding of the instrumentation at hand while providing the basis for the propagation of uncertainties to products derived from spectroradiometer measurements as well as future imaging spectroradiometer space missions. Time has, thus, come for the remote sensing community to embrace uncertainty approaches and also for field spectroscopy applications.

For a full heliosphere calibration and characterization, further experiments are needed to establish the radiances and their uncertainties related to attenuation settings other than the one chosen for the XFR calibration approach. The associated procedures and findings will be described in a future publication. With that groundwork completed, as laid out in parts I and II of this article series, we will be in the position of actually propagating the uncertainties to FSs calibrated against the now traceable heliosphere radiance, and from there onwards to biophysical products derived from spectroradiometer measurements.

Acknowledgment

The authors would like to thank A. Baumgartner at DLR for his great support and valuable input during field spectrometer calibration at DLR's RASTA, as well as T. Speckert, C. Meiller, and H. Kühnle for fruitful discussions and valuable input in the article.

References

[1] R. O. Green et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, pp. 227–248, 1998, doi: 10.1016/S0034-4257(98)00064-9.

[2] M. E. Schaepman, S. L. Ustin, A. J. Plaza, T. H. Painter, J. Verrelst, and S. Liang, "Earth system science related imaging spectroscopy—An assessment," *Remote Sens. Environ.*, vol. 113, pp. 123–137, 2009, doi: 10.1016/j.rse.2009.03.001.

[3] J. W. Chapman et al., "Spectral and radiometric calibration of the next generation airborne visible infrared spectrometer (AVIRIS-NG)," *Remote Sens. (Basel)*, vol. 11, no. 2129, pp. 1–18, 2019, doi: 10.3390/rs11182129.

[4] C. Kuenzer, A. Bluemel, S. Gebhardt, T. V. Quoc, and S. Dech, "Remote sensing of mangrove ecosystems: A review," *Remote Sens. (Basel)*, vol. 3, pp. 878–928, 2011, doi: 10.3390/rs3050878.

[5] K. Anderson and E. J. Milton, "On the temporal stability of ground calibration targets: Implications for the reproducibility of remote sensing methodologies," *Int. J. Remote Sens.*, vol. 27, no. 16, pp. 3365–3374, 2006.

[6] K. Staenz, A. Mueller, and U. Heiden, "Overview of terrestrial imaging spectroscopy missions," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2013, pp. 3502–3505.

[7] S. L. Ustin and E. M. Middleton, "Current and near-term advances in Earth observation for ecological applications," *Ecol. Processes*, vol. 10, no. 1, pp. 1–57, 2021, doi: 10.1186/s13717-020-00255-4.

[8] S. Cogliati et al., "Continuous and long-term measurements of reflectance and sun-induced chlorophyll fluorescence by using novel automated field spectroscopy systems," *Remote Sens. Environ.*, vol. 164, pp. 270–281, 2015, doi: 10.1016/j.rse.2015.03.027.

[9] JGCM, *Guide to the Expression of Uncertainty in Measurement*. Paris, France: BIPM, 2008.

[10] S. Sterckx et al., "Towards a European Cal/Val service for earth observation," *Int. J. Remote Sens.*, vol. 41, no. 12, pp. 4496–4511, 2020, doi: 10.1080/01431161.2020.1718240.

[11] S. A. Trim, K. Mason, and A. Hueni, "Spectroradiometer spectral calibration, ISRF shapes, and related uncertainties," *Appl. Opt.*, vol. 60, no. 18, pp. 5405–5417, 2021, doi: 10.1364/AO.425676.

[12] D. Schläpfer, A. Boerner, S. Bojinski, and M. Schaepman, "Calibration concept for the airborne PRISM experiment (APEX)," *Can. J. Remote Sens.*, vol. 26, no. 5, pp. 455–465, 2000.

[13] L. Ma et al., "Uncertainty analysis for RadCalNet instrumented test sites using the Baotou sites BTCN and BSCN as examples," *Remote Sens. (Basel)*, vol. 12, no. 11, 2020, Art. no. 1696, doi: 10.3390/rs12111696.

[14] L. Zhang, C. Huang, T. Wu, F. Zhang, and Q. Tong, "Laboratory calibration of a field imaging spectrometer system," *Sensors*, vol. 11, no. 3, pp. 2408–2425, 2011, doi: 10.3390/s110302408.

[15] N. Anderson, K. Thome, S. Biggar, and J. Czapla-Myers, "Design and validation of a transfer radiometer," in *Proc. Earth Observ. Syst. 13th*, 2008, vol. 7081, pp. 30–37, doi: 10.1117/12.795478.

[16] A. Baumgartner, P. Gege, C. Köhler, K. Lenhard, and T. Schwarzmaier, "Characterisation methods for the hyperspectral sensor HySpex at DLR's calibration home base," *Sensors, Syst., Next-Gener. Satellites 16th*, vol. 8533, pp. 371–378, 2012.

[17] C. Rammeloo and A. Baumgartner, "Spectroradiometer calibration for radiance transfer measurements," *Sensors*, vol. 23, pp. 1–13, 2023.

[18] E. Woolliams, A. Hueni, and J. Gorrono, "Intermediate uncertainty analysis for Earth observation—instrument calibration module—training course textbook," vol. 2014, pp. 1–121, 2015, doi: 10.5167/uzh-98386.

[19] E. Woolliams, J. Mittaz, C. J. Merchant, and P. Harris, "D2. 2a: Principles behind the FCDR effects tables," Tech. Rep., 2017.
[20] C. J. Merchant et al., "Uncertainty information in climate data records from Earth observation," *Earth Syst. Sci. Data*, vol. 9, no. 2, pp. 511–527, 2017, doi: 10.5194/essd-9-511-2017.
[21] J. Mittaz, C. J. Merchant, and E. R. Woolliams, "Applying principles of metrology to historical Earth observations from satellites," *Metrologia*, vol. 56, 2019, Art. no. 032002.
[22] D. R. Taubert et al., "Providing radiometric traceability for the calibration home base of DLR by PTB," in *Proc. Amer. Inst. Phys. Conf.*, 2013, vol. 1531, pp. 376–379, doi: 10.1063/1.4804785.
[23] W. Feller, "The fundamental limit theorems in probability," *Bull. Amer. Math. Soc.*, vol. 51, no. 11, pp. 800–832, 1945, doi: 10.1090/S0002-9904-1945-08448-1.
[24] A. Hueni and A. Bialek, "Cause, effect and correction of field spectroradiometer inter-channel radiometric steps," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 4, pp. 1542–1551, Apr. 2017, doi: 10.1109/JSTARS.2016.2625043.
[25] A. Hueni, K. Mason, and S. Trim, "Uncertainty support in the spectral information system SPECCHIO," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 2668–2680, 2023.

Mike Werfeli received the B.Sc. degree in geography in 2018 and the M.Sc. degree in geography with a focus on GIS in 2021 from the University of Zurich (UZH), Zurich, Switzerland.

Since 2018, he has been with the Remote Sensing Laboratory, University of Zurich, as a Scientific Assistant and Technician for APEX, AVIRIS–NG, CHIME, and TRISHNA projects. His research interests include the calibration of spectrometers, uncertainty analysis, and the design of experiments.

Andreas Hueni (Member, IEEE) received the B.Sc. degree in computer science from the University of Applied Science Brugg-Windisch, Windisch, Switzerland, in 1997, the P.G.Dipl. in geographic information systems and the M.Phil.(Sc.) degree in earth science from Massey University, Palmerston North, New Zealand, in 2005 and 2006, respectively, and the Ph.D. degree in geography from the University of Zurich (UZH), Zurich, Switzerland, in 2011.

Since 2007, he has been working on the APEX project and is currently an Academic Associate with the Remote Sensing Laboratories, UZH, responsible for APEX sensor and data calibration, and is the Project Manager with the Airborne Research Facility for the Earth System. His research interests include the calibration of spectrometers, uncertainty analysis, and the design of combined database and software systems, such as the spectral database SPECCHIO.

Dessislava Ganeva received the B.Sc. degree in computer science from the Université "Pierre et Marie Curie"- Paris 6, Paris, France, in 1992, the M.Sc. degree in GIS from Southampton University, Southampton, U.K., in 2015, and the Ph.D. degree in remote sensing from the Space Research and Technology Institute – Bulgarian Academy of Sciences (SRTI-BAS), Sofia, Bulgaria, in 2021.

Since 2021, she has been working with SRTI-BAS as an Assistant Professor. Her research interests include the retrieval of vegetation properties from satellites, drones, and in situ sensors for monitoring applications in agriculture and uncertainty analysis of the retrieved parameters.

Giulia Ghielmetti received the B.Sc. degree in micro- and nanosciences from the University of Neuchâtel, Neuchâtel, Switzerland, in 2007, the M.Sc. degree in physics with a focus on lasers physics from the University of Bern, Bern, Switzerland, in 2009, and the Ph.D. degree in physics from the University of Zurich, Zurich, Switzerland, in 2014.

Since 2014, she has been working with the Remote Sensing Laboratory, University of Zurich, as a Lab Engineer, where she helps in the design and execution of experiments related to the ongoing research. Her research interest focuses on the calibration of spectrometers as well as their applications in the field of remote sensing.

Laura Mihai (Member, IEEE) received the B.Sc. degree in physics (computational physics) in 2004, the M.Sc. degree in environmental physics and meteorology in 2009, and the Ph.D. degree in environmental physics and meteorology in 2012, all from the Faculty of Physics, University of Bucharest, Bucharest, Romania.

Since 2008, she has been working as a Researcher with the National Institute of Laser, Plasma, and Radiation Physics, Laser Metrology and Standardization Laboratory, which became in 2014 the Photonic Investigations Lab within the Center of Advanced Laser Technologies. Since then, she has been responsible for the radiometry and terahertz spectroscopy laboratory infrastructure. Over the last two years, as a project responsible, she developed a tool (VALGEOS) capable of validating the calibration of high spectral resolution optical sensors in their operation media. Her research interest includes the development of systems and protocols tailored for the comprehensive characterization, calibration, and validation of field systems employed in the domain of Earth Science.