

Influence of spectral sampling and wavelength calibration when comparing radiometry of two spectrometers for earth observation

Sanket B. Shah^{*a}, Gerard C.J. Otter^{*a}, Bob Kruizinga^a, James P.R. Day^a, Alison L. Gibbings-Bräuer^a

^a TNO, Unit HTI, Stieltjesweg 1, Delft 2628CK, NL

ABSTRACT |

Absolute radiometric calibration is of essence for many earth observation missions. If this needs to be done with a high level of accuracy, on-board calibration is needed to monitor the changes of the instrument over time. This necessitates a calibration unit that establishes a radiometric reference (sun diffuser), ideally in front of the instrument optics. In case of a large pupil, these calibration units can become a significant part of the instrument. We are developing a dedicated small instrument for absolute radiometric calibration called the Absolute Radiometric Reference Instrument (ARRI). This concept aims to provide reference spectra to other Earth observation instruments, thereby eliminating the need for their own individual radiometric calibration units. By doing so, the reference instrument may be significantly smaller than both the primary instrument and its typical calibration unit. Furthermore, a single reference instrument can deliver the reference data for multiple instruments on the same platform. In our earlier papers, it was already proven that a difference in the spatial sampling between two instruments would lead to very small errors, provided that the differences are not too extreme. However, until now it has not been investigated thoroughly how the differences in the spectral resolution and the spectral sampling influence the radiometric error, also considering the limitations of the wavelength calibration knowledge. In the case of different spectral sampling between the main instrument and ARRI, a resampling could be used to compare both the spectra. However, this can lead to errors even if the data and the wavelength calibration knowledge is perfect. We have simulated several cases of such spectral sampling differences between the two to find the limits of the accuracy that can be achieved for inter-instrument comparison. With the simulated results, we conclude on some potential improvements for the on-board metrology used for ARRI wavelength calibration. This compliments the already existing radiometric metrology on the instrument. Together with the other expected error contributors, we update the radiometric error budget for the reference radiance delivered by ARRI and present the inter model comparison.

Keywords: wavelength calibration, radiometry, reference instrument, in-orbit calibration

1. INTRODUCTION

For institutional Earth Observation instruments, the absolute radiometric calibration is mandatory for the quality of science products [1][2]. This radiometric response of an instrument is usually realized by the pre-launch characterisation and calibration of the instrument. However, it has to be monitored throughout the mission lifetime, as optical, mechanical and detector components degrade in orbit, leading to a change in the radiometric response. The on-orbit radiometric response monitoring relies on observations of stable reference sources or reference scenes. This can be a known area on ground or a special target on board of the instrument that is illuminated by a well know source i.e. the sun. That target itself is also subject to degradation so the target also requires monitoring [3]. A common practice is to use a secondary diffuser that has been substantially less exposed to solar radiation [4][5]. This approach assumes that accumulated exposure is the primary and dominant driver of degradation [6].

Another way of performing a radiometric calibration is using known reference scenes on the ground, which is referred to as vicarious calibration. The calibration knowledge is then dependent of the knowledge of the radiance of the ground target. This radiance will predictably vary due to the solar zenith angle but also changes difficult to predict such as the conditions of the target itself and the variation of the atmospheric column in between [7][8].

*sanket.shah@tno.nl; phone +31 888 665 912; tno.nl | gsserard.otter@tno.nl; phone +31 888 666 547; tno.nl

The ARRI concept combines the best of both the worlds. It synchronously observes the same Earth scene as the main instrument providing a calibrated reference radiance to the main instrument. ARRI provides the absolute reference by being equipped with its own calibration device using the sun as a reference. ARRI purely focuses on delivering calibration references and does therefore not require to have the same spatial or spectral resolution as the main instrument, making the ARRI potentially much smaller than the main instrument(s) [9].

In previous papers it was already proven that the inter instrument calibration will introduce small errors due to the different spatial resolution used [9]. But the impact of difference in the spectral calibration has not been addressed. In this paper, simulations have been performed to investigate the effect of difference in spectral resolution and sampling on the radiometric error for inter-instrument comparison. High spectral resolution radiance reference spectra are used in the simulations to achieve a realistic representation of the spectral variability and fine spectral structure expected in operational observations [10].

2. INTER INSTRUMENT CALIBRATION ACCURACY

In an ideal case, when one instrument is used as a reference for another instrument, there should be residual no radiometric error. In practice, this is unrealistic even if the design of the two instruments is exactly the same. There will always be differences due to tolerances and their state of degradation. These differences are minor relative to what is proposed with ARRI. It is intrinsic in the ARRI concept that the spatial resolution will be poorer than the main instrument, perhaps more than an order of magnitude. It has been proven that this will still enable a good radiometric comparison [9] provided the right analysis is performed.

For the spectral sampling, the difference may not be so large. The benefit of using ARRI is to remove the need for a calibration unit, which is most relevant for larger pupil instruments like CHIME [9]. The advantage of ARRI diminishes for very high spectral resolution instruments such as in Sentinel 5P [14] that have a smaller pupil and therefore a moderately sized calibration unit. In the current ARRI concept trade-offs, the design choice was made to maintain a spectral resolution comparable to instruments such as CHIME. However, ARRI would still remain useful for radiometric calibration of high spectral resolution instruments. It just means that the instrument does not provide the spectral details for the science. For radiometric comparison however the aim is to determine changes that are most likely caused by degradation, which are likely less divergent for very narrow spectral lines.

The ARRI can measure the earth reflectance purely based on the measurements of the different metrology lines on board as described in [9]. The ARRI will have its own characterization and calibration, following an approach comparable to that used for other instruments. Consequently, ARRI is expected to achieve a similar level of measurement accuracy. On top of that the inter instrument comparison will have an additional contribution to further improve the accuracy.

The accuracy of the inter-instrument comparison depends on the magnitude of the differences between the instrument being evaluated and ARRI. It was already explored in [9] the impact of the difference in spatial sampling on the error, but not that of spectral calibration. There are two aspects to this; the difference in wavelength calibration and the knowledge of that difference.

A difference in spectral sampling (and spectral resolution) will change the details that are visible in the spectrum. The earth and sun spectra have a lot of spectrally narrow absorption lines in their spectra (Fraunhofer lines and atmospheric absorption lines). These lines are only visible when measuring with a very high resolution spectrometer. Most current generation atmospheric instruments do resolve all the lines relying on the structures created in the lower resolution spectra they measure. Nevertheless, these spectral features become less pronounced as the spectrometer's resolution decreases. This is demonstrated in Figure 1 where the radiometric response of an earth reference spectrum (Tropical Dark) is shown for different spectral resolutions. The raw spectrum was sampled at 0.04 nm (light grey).

In addition, an erroneous shift in the wavelength calibration will lead to an inaccurate assignment of the radiometric output as a function of the wavelength. In case of a spectrally flat radiometric scene, this will not lead to any errors, but in reality such scenes do not exist. The more spectrally dynamic the scene is, the larger the potential errors become.

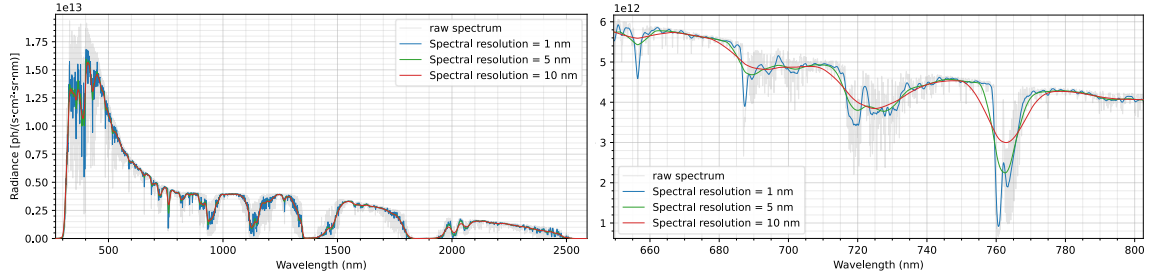


Figure 1. Full raw spectrum after convolution with a spectral resolution simulated with a triangular ISRF of varying FWHM of 1, 5 and 10 nm (left), looking near the O2-A band (right).

3. WAVELENGTH CALIBRATION OF ARRI

To quantify how spectral distortions degrade the accuracy of the measured spectrum, we developed a simulation that compares the output of a distorted instrument against an accurate, high-resolution reference spectrum. The simulation takes as its starting point Sentinel-5 reference spectra sampled at 0.04nm resolution [10] and treated as the ground truth. An instrument is then modelled as an observer, so any discrepancy between the instrument's output and the truth can be attributed unambiguously to the instrument's own spectral response and wavelength calibration, rather than to the differences in the scene content.

Two instrument configurations were considered: a reference instrument, based on CHIME's characteristics, representing a well-characterised instrument and ARRI, representing the radiometric refence for CHIME with a coarser spatial resolution but a slightly higher spectral resolving power. Both instruments observe the same underlying scene; only their spectral response functions, detector sampling, and injected wavelength distortions differ.

The instrument spectral response function (ISRF) is the response of an individual spectral channel (pixel) as function of wavelength normalised to the overall energy. The full width half maximum of this function is considered the spectral resolution. The instrument spectral response function (ISRF) is build up by convolution of individual contributions of the point spread function (PSF), the pixel response and the slit function. The ISRF function is by definition energy normalized so convolving a spectrum with this function will conserve energy.

$$\text{ISRF}(\lambda) = K_{\text{psf}}(\lambda) \otimes K_{\text{pixel}}(\lambda) \otimes K_{\text{slit}}(\lambda) \quad (1)$$

The convolution of the $\text{ISRF}(\lambda)$ with the high resolution spectrum $S(\lambda)$ is then used to create the spectrum as observed by the instrument, which is then sampled in discrete intervals of sampling intervals Δ_i .

$$S_{\text{conv}}(\lambda) = S(\lambda) \otimes \text{ISRF}(\lambda) \quad (2)$$

$$\bar{S}_i = \frac{1}{\Delta_i} \int_{\lambda_i}^{\lambda_{i+1}} S_{\text{conv}}(\lambda) d\lambda \quad (3)$$

The distortion model used is similar to prior work done for wavelength calibration of GOME instrument using spectral shift and squeeze [11]. Considering only one spatial pixel, we inject 3 variables: spectral shift (α), spectral squeeze/stretch factor (β) about a central wavelength (γ_{nm}). For a wavelength axis λ with span $\Delta\lambda_{\text{in}} = \lambda_{\text{max}} - \lambda_{\text{min}}$:

- Target span: $\Delta\lambda_{\text{target}} = \Delta\lambda_{\text{in}} + \alpha$
- Squeeze factor: $\beta = \Delta\lambda_{\text{target}} / \Delta\lambda_{\text{in}}$
- Warped wavelength axis: $\lambda' = \alpha + \gamma_{\text{nm}} + (\lambda - \gamma_{\text{nm}}) \cdot \beta$

Repeating the convolution and sampling with this warped wavelength axis gives you the output expected without wavelength calibration. For example, Figure 2 (top-left) is an example output without any warping (blue plot) and with (red) warping introduced with -1.5nm shift (α), 2.5nm of squeeze ($\beta = 1.001031$) about 1400nm (γ_{nm}). This is with an ISRF modelled as a triangular function (oversampling of 2) with a FWHM of 5nm and energy maintained to unity (top right), convolved with the raw spectrum of Tropical Dark from 300 to 2500nm. The radiometric error is shown (bottom

left) with the blue or red indicating which value is higher (undistorted or distorted, respectively). The region with Fraunhofer lines such as Oxygen A, B, Hydrogen α show the impact of spectral dynamics with poor wavelength calibration (bottom right). State-of-the-art EO spectrometers with 5-10 nm spectral resolution, may experience wavelength distortions in the order of 1-5 nm, which are corrected by wavelength calibration and they report a post-calibration wavelength knowledge better than 0.1-0.5 nm [12][13]. This magnitude of calibration aligns well with the magnitude of distortions introduced in this work.

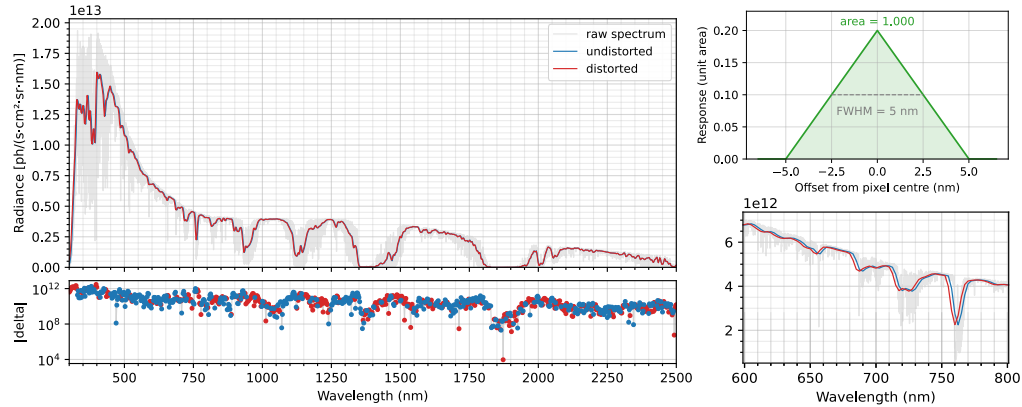


Figure 2. Introduction of wavelength errors into the high resolution raw radiance values (300 to 2500nm) and convolution with a triangular kernel. Top left: raw radiance spectrum (faint grey), convolution with the ISRF kernel sampled at 2.5nm of raw spectrum (blue), of distorted raw spectrum (red). Top right: ISRF used with a FWHM of 5nm representing the instrument normalizing the total energy of 1 pixel. Bottom left: the absolute radiance error (undistorted – distorted) below in log scale. Bottom right: zoomed in region near the Oxygen-A band showing the major influence of shift and squeeze.

The proposed ARRI's architecture allows simultaneously looking at the diffused sun spectrum which is typically the reference for radiometric calibration. We also tested adding equivalent spectral distortion in the solar spectrum and tested the wavelength calibration with commonly used approach using the Fraunhofer absorption lines in the solar spectrum Figure 3 [14][15]. We were able to achieve better than 0.01nm fit purely looking at the solar spectrum using known approaches with a radiometric error of less than 0.12% RMS. Similarly, atmospheric band-weighted wavelength correction have been performed on several instruments achieving sub-nm accuracies [16][17].

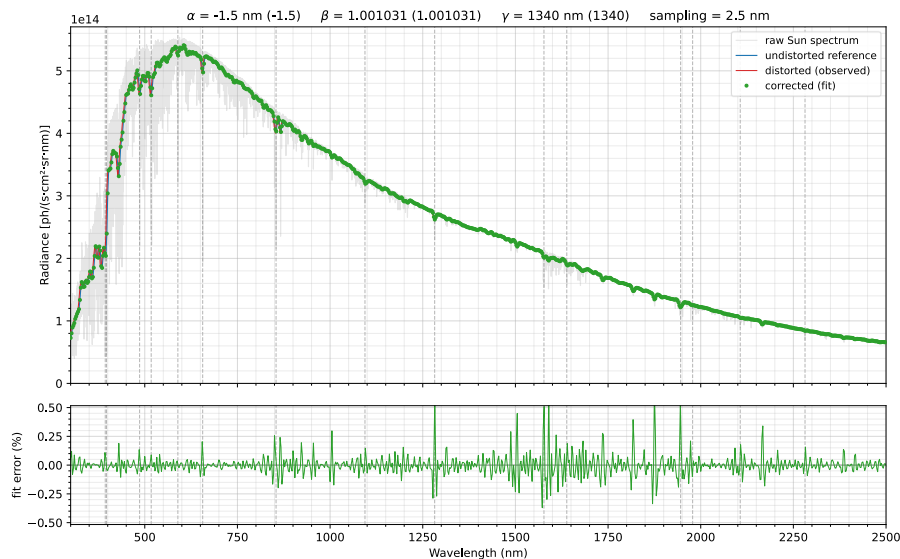


Figure 3. Raw Sun spectrum when convolved with a triangular ISRF with a FWHM of 5nm and sampled at 2.5nm with injected distortions and fitting based on 15 Fraunhofer lines (vertical dashed lines). Total radiometric error after wavelength calibration fit was 0.12% RMS ($=2.09 \times 10^{11}$ ph/s/cm²/sr/nm).

We applied a spectral shift of this magnitude on Earth spectrum (Tropical Dark) to see its impact on the radiometry. The results in Figure 4 show three cases of spectral shift ranging from 0.1 to 0.3 nm. The observed error magnitude scales with the spectral dynamic (high $dL/d\lambda$), as expected. In most cases however the error is below 0.1% as can be seen in the distribution plots at the right side of the figure (0.3nm extreme case, $1\sigma = 0.11\%$).

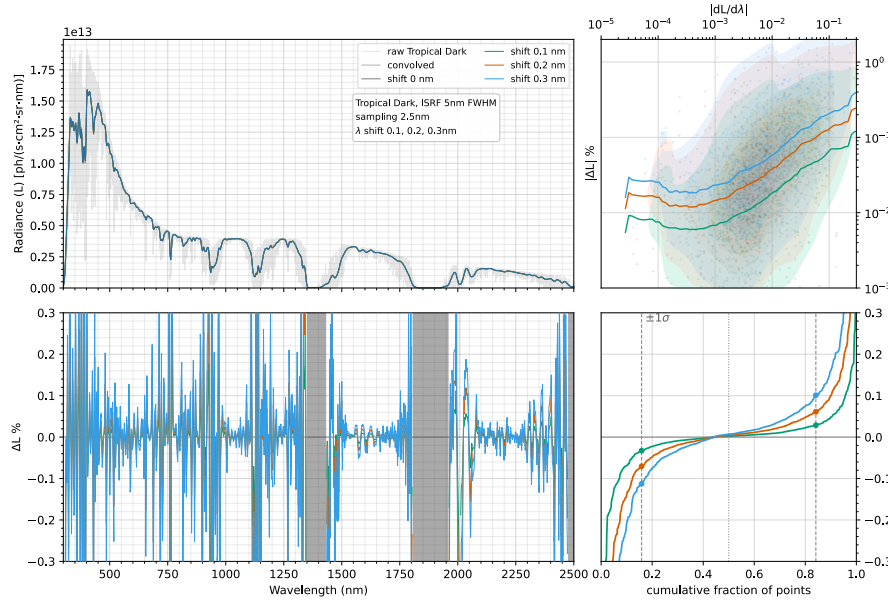


Figure 4. Raw spectrum Tropical Dark when convolved with a 5nm FWHM triangular ISRF, oversampled at 2.5nm, and with spectral shifts of 0.1nm introduced (top right). The relative radiometric error for each new spectra obtained with unused atmospheric absorption bands greyed (bottom left). The cumulative fraction vs error (bottom right, shared y-axis with error plot, ΔL %). The absolute the absolute spectral derivative (slope) vs error of the reference spectrum (top left).

4. MATCHING THE SPECTRA

Previously in Figure 1 we already saw that the resolution of the spectrometer strongly influences the shape of the spectrum measured. Therefore when comparing different instruments we need to somehow match the spectral resolution as well as the sampling. This data handling to compare can itself introduce radiometric discrepancies. To match the spectral resolution we will convolve the both spectra with an artificial ISRF two or more times the resolution of the courser spectrometer.

We simulated the two possibilities when there is a mismatch: a case where ARRI is used to match with a CHIME like spectrometer, where the CHIME like spectrometer has the courser resolution (8.5nm vs 5nm, Figure 5), and an inverse case where ARRI is used for a Sentinel 5 like spectrometer for the full range, where ARRI has the courser resolution (0.195 nm vs 5nm, Figure 6). In both cases we simulated different ISRF FWHM values to match the spectra (where the FWHM is at least twice or larger than the lower resolution instrument) and resampled the convolved data to a common spectral oversampling which is twice the smallest common ISRF FWHM value.

In these plots, the relative error is calculated as in Equation 4 after the convolution with new ISRF & resampling for each new sample. It is the radiometric error between the two samples divided by the mean value of the two instrument samples.

$$\Delta L \% = 100 \times \frac{(L_{\text{CHIME, isrf2}} - L_{\text{ARRI, isrf2}})}{((L_{\text{CHIME, isrf2}} + L_{\text{ARRI, isrf2}})/2)} \quad (4)$$

We can clearly observe that the radiometric error between the two instruments reduces with a larger FWHM of the common ISRF at the cost of spectral details. Also, the errors increase significantly where the spectral derivative (slope) is higher, i.e. where larger spectral dynamics exist. The trend is less pronounced in the second case where the instrument Sentinel-5 has much higher native resolution compared to ARRI. This can be explained by the larger differences in spectral resolution and sampling because the spectrum of Sentinel-5 is very close to the highly resolved raw reference spectrum.

When the sampling is comparable, the more resolved spectrum does not provide sufficient detail resulting in larger errors. This can be mitigated by going for much lower resolution for both of the spectra, which we also have done as can be seen in the same figures. As can be seen in the CHIME case, the error drops significantly when using lower resolution common ISRF. This lower ISRF can be justified if we are looking at the effects with little spectral dynamics, which is the case for degradation related errors.

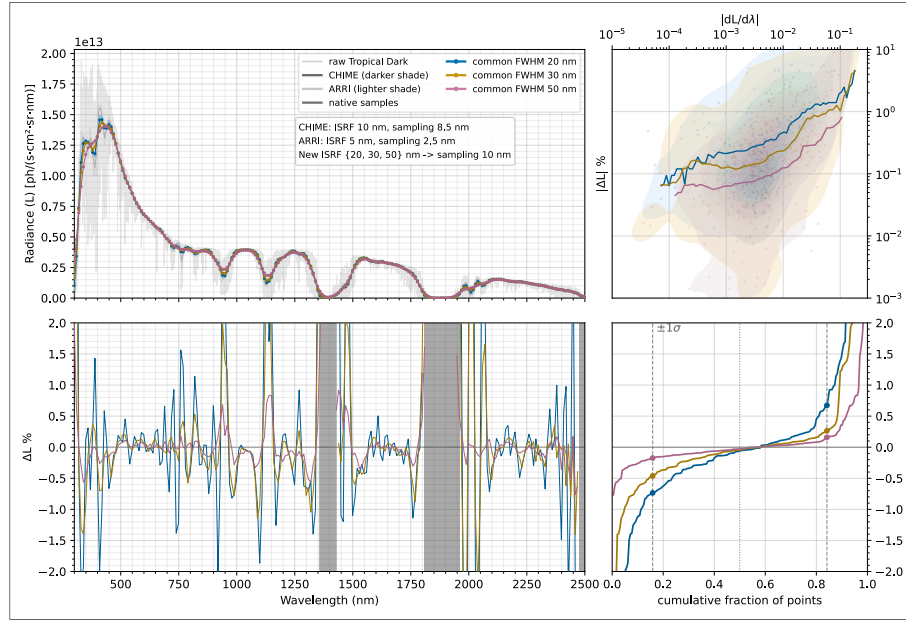


Figure 5. Matching spectra between simulated CHIME and ARRI. Top right plot shows the radiometric spectra obtained in the different cases of the common ISRF (see legend). The relative mean radiometric error for each new spectra obtained with unused atmospheric absorption bands greyed (bottom left). The cumulative fraction vs error (bottom right, shared y-axis with error plot, ΔL %). The absolute the absolute spectral derivative (slope) vs error of the reference spectrum (top left).

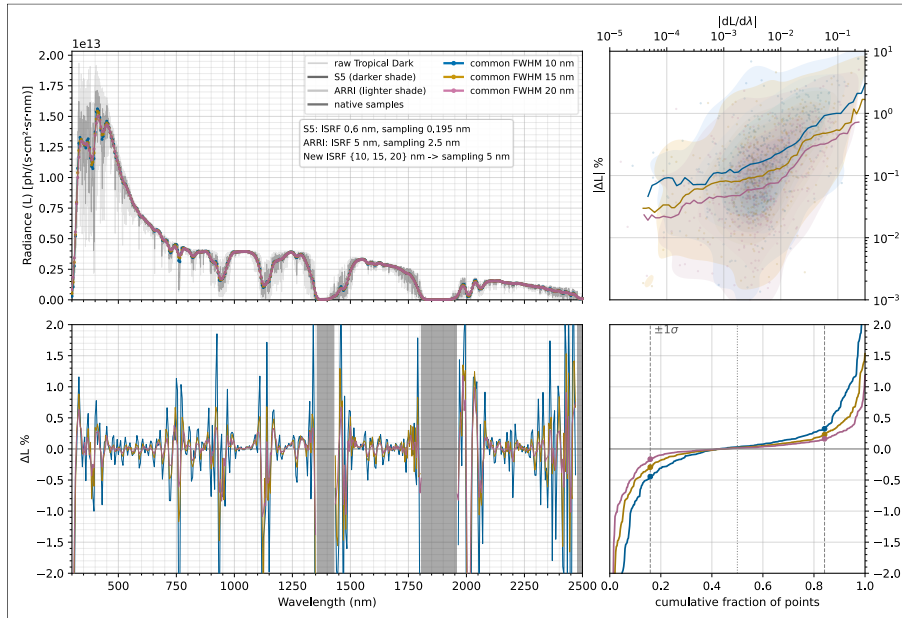


Figure 6. Matching spectra between simulated Sentinel-5 and ARRI. The plots arranged same as in Figure 5. The mean errors are slightly lower in magnitude than the simulated CHIME case.

5. ONBOARD METROLOGY FOR MONITORING WAVELENGTH CALIBRATION

As has been proven both by our own analysis and from literature, very accurate wavelength calibration is possible using the features from the measured spectra. The on-ground characterization should provide a sufficiently accurate starting point for the wavelength calibration in orbit. The on-ground must be achieved with the same accuracy as required for in orbit wavelength calibration which is easily achievable with the current techniques. Nevertheless, we summarize the state-of-the-art in-orbit monitoring of the wavelength calibration in orbit using on board technologies.

For in-orbit wavelength calibration, off-the-shelf laser diodes exist that can provide a very accurate wavelength knowledge and also scan over a small range enabling to determine the ISRF. This method is used in the Sentinel-5P mission for the shortwave infrared bands [14]. This method can be easily adapted and implemented in the existing calibration unit of the ARRI. However, the total spectral coverage needs to be verified and it remains to be evaluated if range of the laser diode can realistically accurately calibrate the entire spectrum of the instrument.

A solution used in MERIS is a doped diffuser (pink diffuser). This is a diffuser that has Erbium doping that creates a distinct spectral lines that can be used as spectral reference [18]. This is similar to using a hollow cathode lamp used on GOME-2 where atomic emission lines are used [11]. In both the cases it is not free to choose the position of the spectral lines or features available since they are limited by the physical characteristics of the hardware used. It also requires some additional voluminous hardware straining the trade-offs further. These methods are considered less suitable for a small instrument like ARRI.

6. CONCLUSION

The ARRI instrument is designed to function as a radiometric reference for any other instrument measuring in the same wavelength range. This other instrument will not have the same spectral resolution and spectral sampling as ARRI, which can lead to errors. Additionally, there will be some error introduced due to the inaccuracies in the wavelength calibration of ARRI relative to the main instrument. The nature of these errors have been explored and their expected magnitude quantified.

For the magnitude of the wavelength calibration error, we looked at available literature and we built our own model of the measured spectra to see how accurate the wavelength calibration can be. Based on our own model we achieved a spectral calibration accuracy between 0.1 and 0.2 nm for the ARRI resolution of 5 nm with 2.5 nm sampling. This value is in line with what can be found in literature. This small wavelength calibration error can lead to a moderate radiometric error of 0.11% 1-sigma in a typical Earth scene.

The difference in the spectral resolution and sampling requires a good resampling strategy to accurately compare the inter-instrument spectra. Even with minimal resampling, the error can be up to 0.8% 1-sigma. When a more rigorous resampling, this error can be reduced to 0.2% 1-sigma at the cost of less detailed spectral information. The large errors have a strong correlation with the dynamics in the spectrum and total error can be mitigated since these parts of the spectrum can be easily be identified. In that case, a stronger resampling might not be required, maintaining detailed spectral information.

The overall error budget for ARRI as a reference is estimated using the results found here and with earlier results. For the reflectance calibration of ARRI itself, the error budget stands at 0.86% [9]. With the inter model comparison we have a spatial re-sampling error of 0.20% [9], a spectral shift error of 0.11% and a spectral re-sampling error of 0.20%. This gives an overall error 0.91% 1-sigma for the ARRI instrument. The reflectance calibration error of the ARRI instrument remains as the dominant contributor.

REFERENCES

- [1] Fox, N., Theocharous, M., Harris, P., et al., "A traceable radiometry architecture to underpin the development of climate data records from Earth observations," *Metrologia* 48(2), S155-S163 (2011).
- [2] Merchant, C. J., Embury, O., Bulgin, J. R., et al., "Uncertainty information in climate data records from Earth observation," *Int. J. Remote Sens.* 30(2), 325-341 (2009).

- [3] Datla, R. U., Rice, J. P., Lykke, K. R., et al., “Best practice guidelines for pre-launch characterization and calibration of instruments for passive optical remote sensing,” *J. Res. Natl. Inst. Stand. Technol.* 116, 621-646 (2011).
- [4] Xiong, X., Sun, J., Barnes, W., Salomonson, V., Esposito, J., Erives, H., and Guenther, B., “MODIS solar diffuser stability monitor sun view modeling,” *IEEE Trans. Geosci. Remote Sens.* 43(2), 355-365 (2005).
- [5] Twedt, K. A., Angal, A., Chen, H., and Xiong, X., “Comparison of MODIS solar diffuser stability monitor calibration results for different operational configurations,” *Proc. SPIE 11127, Earth Observing Systems XXIV*, 1112728 (2019).
- [6] Xiong, X., Xie, X., Angal, A., et al., “Characterization of MODIS solar diffuser on-orbit degradation,” *Proc. SPIE 6677, Earth Observing Systems XII*, 66770O (2007).
- [7] Thome, K. J., Whittington, E. E., Smith, N., Nandy, P., and Zalewski, E. F., “Ground-reference techniques for the absolute radiometric calibration of MODIS,” *Proc. SPIE 4135, Earth Observing Systems V*, 624-638 (2000).
- [8] Helder, D. L., Thome, K. J., Mishra, N., et al., “Absolute radiometric calibration of Landsat using a pseudo-invariant calibration site,” *IEEE Trans. Geosci. Remote Sens.* 51(3), 1360-1369 (2013).
- [9] Otter, G., Day, J., Di Iorio, E., et al., “Absolute radiometric reference instrument (ARRI),” *Proc. SPIE 11858, Sensors, Systems, and Next-Generation Satellites XXV*, 118580D (2021).
- [10] Sentinel-5 Team, “Sentinel-5 Phase B2, C/D and Support to Phase E1: Reference Spectra Specifications,” S5-RS-ESA-SY-0003, Issue 2, Rev. 0, ESA Technical Report (2013).
- [11] Munro, R., Lang, R., Klaes, D., Poli, G., et al., “The GOME-2 instrument on the MetOp series of satellites: instrument design, calibration, and level 1 data processing – an overview,” *Atmos. Meas. Tech.* 9(3), 1279-1301 (2016).
- [12] Carmona, E., Bachmann, M., et al., “The DESIS spaceborne hyperspectral instrument calibration,” CEOS WGCW IVOS-35 Workshop, Oberpfaffenhofen, Germany, 27 Sept. 2023.
- [13] Baur, S., Mücke, M., Sang, B., Wachter, R., et al., “Pre-flight calibration and characterization of the EnMAP sensor,” *Proc. SPIE 12777, International Conference on Space Optics (ICSO 2022)*, 127774V (2023).
- [14] Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., et al., “In-flight calibration results of the TROPOMI payload on board the Sentinel-5 Precursor satellite,” *Atmos. Meas. Tech.* 13(7), 3561-3580 (2020).
- [15] Hilbig, T., Weber, M., Bramstedt, K., Noël, S., Burrows, J. P., et al., “The new SCIAMACHY reference solar spectral irradiance and its validation,” *Sol. Phys.* 293, 125 (2018). doi:10.1007/s11207-018-1339-9.
- [16] Gao, B.-C., Montes, M. J., and Davis, C. O., “Refinement of wavelength calibrations of hyperspectral imaging data using a spectrum-matching technique,” *Remote Sens. Environ.* 90(4), 424-433 (2004).
- [17] Kuhlmann, G., Hueni A., Damm A., and Brunner D., “An algorithm for in-flight spectral calibration of imaging spectrometers,” *Remote Sens.* 8(12), 1017 (2016).
- [18] Delwart, S., Preusker, R., Bourg, L., et al., “MERIS in-flight spectral calibration,” *International Journal of Remote Sensing* 28(3-4), 479-496 (2007).