# Vicarious calibration as a method for inflight monitoring of atmospheric composition small-sat instrumentation

Kempen, T.A. van; Hoogeveen, R.; Hees, R.M. van; Kempers, N.; Veldhuis, C.E.; Rotmans, T.J.; ... ; Petrozzi-Ilstad, M.

# VICARIOUS CALIBRATION AS A METHOD FOR IN-FLIGHT MONITORING OF ATMOSPHERIC COMPOSITION SMALL-SAT INSTRUMENTATION

**Tim A. van Kempen(1), R. Hoogeveen(1), R. M. van Hees(1), N. Kempers(2,1), C. Veldhuis(2,1), T. J. Rotmans(3,1), F. Oggionni(3,1), J. Landgraf(1), I. Aben(1,4)**

*(1)SRON Netherlands Institute for Space Research, Niels Bohrweg 4, 2333 CA, Leiden, the Netherlands, t.a.van.kempen@sron.nl*
*(2)Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA, Leiden, the Netherlands*
*(3)Technical University Delft, Aerospace Engineering Faculty, Kluyverweg 1, 2629 HS, Delft, the Netherlands*
*(4)Vrije Universiteit, De Boelelaan 1081, 1081 HV, Amsterdam, the Netherlands*

## PAPER

Due to their dimensions and weight constraints, atmospheric composition small-sat missions typically lack onboard calibration capabilities, forcing them to resort to vicarious calibration methods.
The sole instrument on the Sentinel-5 Precursor mission, TROPOMI, has been performing excellently since its launch in 2017. The SWIR module has degraded less than 0.2%.
TROPOMI provides an excellent test-case for the applicability of vicarious calibration for atmospheric composition small-sat application such as the upcoming TANGO mission.
Here we present the ongoing efforts to validate both the relative radiometric calibration (i.e. instrument stability) using desert PICS, as well as the absolute radiometric calibration using ground-measurement of the Railroad Valley playa. Such independent validation constrains TROPOMI to 0.5% relative and 6% absolute. Several significant limitations have been identified; First, results of large pixel sizes deviate from the assumption of surface homogeneity, in particular smaller sites such as RRV. Modelling of the BRDF (Bidirectional Reflectance Distribution Function) needs to be accurate to account for non-Lambertian effects. For small-sat applications with much higher spatial resolution, homogeneity is much simpler to guarantee. We also recommend soundings be done as close to nadir-view as possible to minimize other uncertainties due to the BRDF.

## 1 INTRODUCTION

In addition to the instrument calibration, the verification of instrument stability in time is an essential component for trustworthy datasets probing the Earth's atmospheric chemical composition. Due to their dimensions and accompanying weight constraints, missions in the so-called 'small-sat' or 'cube-sat' categories typically lack on-board facilities necessary for such activities.
Vicarious calibration [1] may be a viable alternative for such instruments. These methods rely on repeated measurements of the (continuum) radiance over pseudo-invariant calibration sites (PICS) that determine (limits on) instrument degradation, perform cross- calibration, or, if ground-based measurements are available, can be used to measure the absolute radiometric calibration.
PICS are sites that are a) as cloud free as possible, b) dry, and c) vegetation-free and d) devoid of human activity. With these condition, the bi-directional reflectance distribution function (BRDF) is

Small Satellites Systems and Services Symposium (4S 2024), edited by Max Petrozzi-Ilstad,
Proc. of SPIE Vol. 13546, 135460N  · 0277-786X · doi: 10.1117/12.3061019

assumed to be invariant over time. In addition, spectral variation at the SWIR wavelength is smooth and relatively low compared to the spectral resolution of the spectrometers to be tested. Well-known examples of PICS are RailRoad Valley (also known as RRV, located in central Nevada, USA) or remote desert regions in the Sahara or on the Arabian peninsula [1][2][3].

Vicarious calibration at these wavelengths has been done for atmospheric composition missions such as OCO-2 and GOSAT with moderate success [2][4].

In recent years, the SWIR 2.3 micron band has been extensively used by TROPOMI to probe Earth's atmospheric composition. TROPOMI (The Tropospheric Measurement Instrument) is the sole instrument onboard the Sentinel-5 Precursor mission [5][6][7]. It consist of two modules: the UVN module, covering the UV, Visible, and near-IR wavelengths, and the SWIR module, covering the 2.3 micron wavelength band in the short wavelength infrared (SWIR).
Due to the excellent instrument stability during nominal operations [6][8], the results of the TROPOMI-SWIR are well-suited to probe the (limitations of) the vicarious calibration method as applied on SWIR spectrometer soundings, in particular for small-sat applications.
In this contribution, we will present results from the vicarious calibration efforts done at SRON Netherlands Institute for Space Research using the TROPOMI-SWIR radiance results and consider the results in possible applications for small-sat instrumentation in the SWIR wavelength region, such as the recently accepted TANGO [9] constellation.
In section 2, details are provided on the TROPOMI instrument and the observed instrument stability of the SWIR module. Section 3 presents the methodology behind vicarious calibration as used on the TROPOMI data. Accompanying results are given in sections 4 (relative radiometric calibration) and 5 (absolute radiometric calibration). Since the goals of relative and absolute radiometric calibration are different, it is worth discussing the results separately. Section 6 considers both sets of results for small-sat application. Conclusions are given in section 7.

## 2 TROPOMI SWIR

TROPOMI is the only payload on the Sentinel-5 Precursor (S5p) mission [5]. Launched in October 2017 and operational since April 2018, it has produced a wealth of products that provide daily global coverage on several greenhouse gases in Earth's atmosphere. The SWIR module is focussed on the CO and $CH_4$ column densities [15][16]. By covering the 2.3-micron window with a 80 nm bandpass at a 0.25 nm spectral resolution, it uses the 130 degree wide swath to carry out its measurements using an innovative immersed grating.

### 2.1 TROPOM-SWIR Stability

The instrument stability of TROPOMI-SWIR is verified using on-board lights (i.e., a white light source, several LEDs and five laser diodes) and daily measurements of the solar irradiance [6]. It has shown remarkable stability over the six years in operations [6][8], with a limit on the transmission degradation of 0.3%, a pixel loss of less than 0.5% and a noise increase of less than 0.5%, perhaps even no increase. All other parameters (e.g., dark current, WLS stability) relevant to instrument stability have not shown any changes beyond the measurement uncertainties.

An example of this high level of stability over nearly six years of inflight operations can be seen in Figure 1, where the time-series of the solar irradiance over the main diffuser (QVD1) is shown.
This stability is a key initial condition for vicarious calibration with TROPOMI. Instead of not knowing the origin of any uncertainty (i.e., instrumental or due to the methodology), we can probe the effectiveness of the vicarious calibration method at SWIR wavelengths.

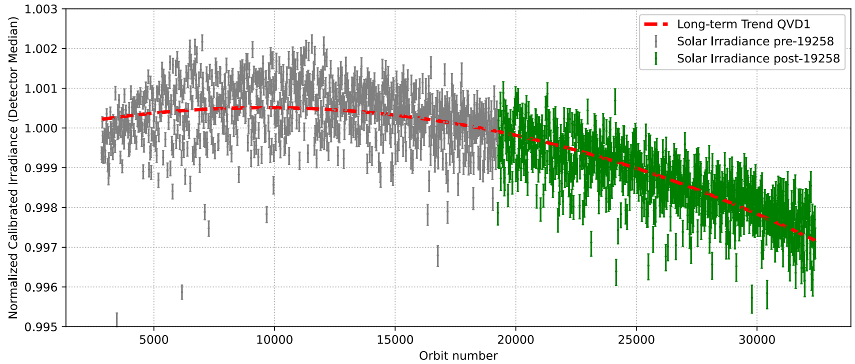


Fig. 1: Daily solar irradiance measurements, normalized to the start of nominal operations. These show the remarkable stability of the TROPOMI-SWIR module (<0.3% degradation). The observed curve is not caused by a degradation in the solar diffuser (QVD1, see [8]). During orbit 19258, the calibration key data for the solar irradiance were revised, indicated with a color change from grey to green.

## 3 METHODOLOGY

In its essence, Vicarious calibration is a relatively simple method. For absolute calibration, a Top-of-Atmosphere measurement (ToA, e.g., from TROPOMI-SWIR) is compared to an independently obtained reference spectrum over a PICS [3]. The reference spectrum can be calculated from the reflectance of the ground and propagated using radiative transfer to derive an equivalent ToA atmosphere. It is assumed that the measured reflectance is representative of the observable surface throughout the pixel. In other words, the surface within a pixel is homogeneous and isotropic. Alternatively, the ToA spectra of the instrument over a PICS itself can be compared to themselves over (long) periods of time. Although this eliminates the need for a reference, only the relative calibration can be performed. However, relative calibration is particularly useful to monitor instrument degradation [4,12].

### 3.1 Method

Vicarious calibration relies on two elements. First is the change (or lack thereof) in time. Secondly is the comparison to a reference (either in time or with a different instrument). Both elements have been corrected with the necessary corrections.
For the time series, we correct the measured radiance for

- The position of the sun
- The viewing angles of the instrument
- Additional seasonal variations

If a surface is assumed to be non-Lambertian, it is also corrected for

- BRDF effects

BRDF effects are corrected using a term *nBRDF* which is defined as BRDF (θ,ϕ) /BRDF (0,0) for a given position of the Sun. Viewing angles of the instrument are θ (zenith) and ϕ (azimuth).
For absolute calibration, the following corrections are done on top of these.

- Temporal variation correction between sounding and reference measurement time

In addition, to obtain a ToA (Top of Atmosphere) spectrum from a measured reflectance, a simulated 2.3-micron spectrum is produced using the radiative transfer code RemoteC [14].
The full method is described in detail in [1], [8] and [12].

### 3.2 PICS

A Pseudo-Invariant Calibration Site, or PICS, is a site that is well-suited for vicarious calibration. These sites should be invariant in time (i.e. unaffected by human activity as well as devoid of vegetation) and preferably lack altitude differences on the scales of the instrument spatial pixel extend. An inventory was made by [3] on suitable sites at scales of 20 and 100 km for various instrument with the aim to characterize the (lack of) spatial variability. See [3] for the exact limits on spatial inhomogeneity allowed. These sites are all located in the Saharan, Namib and Arabian deserts, as shown in Fig. 2 (taken from [12]). None of the sites shown is easily accessible. All are far from human habitation, lacking even basic infrastructure. Placement of ground equipment is thus not a realistic option. These sites are thus well-suited for relative calibration (as well as potential cross-calibration with other space-based instruments, if only the differential between two instruments is required.)

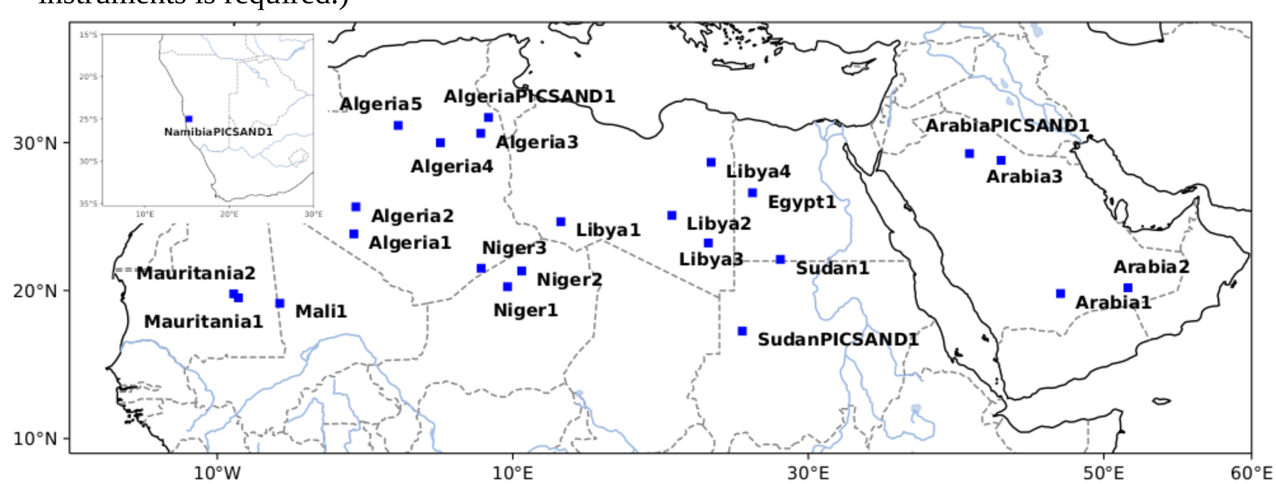


Fig. 2 (taken from [12]): Distribution of PICS in the Saharan, Namib and Arabian deserts that are suitable at 20 and 100 km spatial scales.

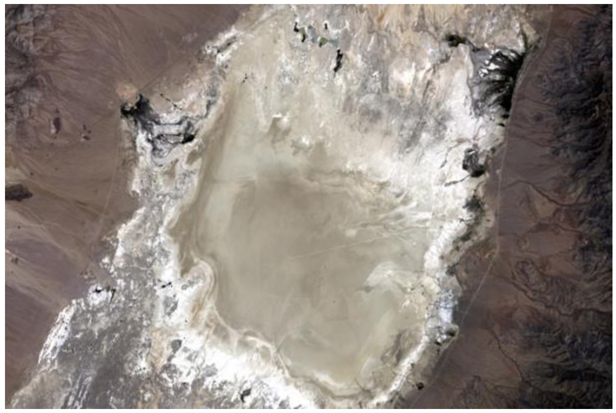

Fig. 3: Satellite view of RRV in the visible wavelength range.

### 3.3 An Accessible PICS: the Railroad Valley playa

PICS that are more easily reachable do exist all over the world. However, most lack the required spatial homogeneity across the spatial extend of pixels of SWIR wavelength spectrometers. This can be due to multiple reasons; Either due to their own inhomogeneity (e.g., mountainous structure), temporal variations (e.g. vegetation in the Australian outback) or very high latitudes (i.e. Arctic and Antarctic sites qualify, but lack yearly coverage.). Most are simply too small for typical atmospheric sounders.

One of the more suitable sites is the Railroad Valley playa in the centre of Nevada, USA [2], often referred to as RRV. RRV is a dried-up salt basin extending to about 10 km in diameter. Although it is not completely circular, it is large enough that it can fill a significant portion of a TROPOMI pixel projected on the Earth. It also is relatively close to the American highway infrastructure. In April 2023, it was designated for NASA satellite calibration by the Bureau of Land Management[1]. At the centre of the playa, the RADCALNET initiative placed instrumentation to obtain reference data continuously [13]. It is the only site of this network to do so at IR wavelengths. The OCO and GOSAT instrument teams conduct yearly campaigns since the 2000s to obtain reference measurements of the reflectance usable for vicarious calibration, which has been freely distributed [2][4][8].

### 3.4 Relative versus Absolute radiometry

As seen above, one of the large differentiations between available sites in the world is their capacity for human access. Most suitable sites are very remote, typically located far from any human activity. Natural conditions can be extremely harsh for ground measurements that can act as a reference, or natural and/or political barriers exist for an effective deployment of ground equipment. For vicarious calibration, we must make the distinction between sites aimed at relative (where no ground equipment can be deployed) and absolute (where ground equipment can be deployed) radiometric calibration.

For the relative case, satellite measurements are normalized and compared within a time series. All required corrections are associated with the passage of time. As sites should be inherently

[1] (see https://www.nasa.gov/science-research/earth-science/nearly-23000-acres-designated-for-nasa-satellite-calibration/).

invariant, the corrections are reduced to the viewing parameters of the instrument and the position of the sun. These are four parameters in total: the zenith angle of the sun and the instrument and the azimuth angle of the sun and the instrument. It is also assumed that the incident beam (i.e., solar variation) is invariant in time at the observed wavelengths. The only exception is the yearly variation due to Earth's Keplerian orbit. Given known orbit parameters, all viewing parameters are easily determined. The largest uncertainty remains from the uncertainty in the BRDF representation of the site. These cannot be directly measured.

For sites to be used for absolute calibration, this independent measurement can be obtained. Typically, the measurement is the nadir reflectance of the incoming beam. At times, a full characterization on-ground of the BRDF can be obtained, but this is a very time intensive measurement. In addition, it is not strictly necessary within the assumptions in place. Satellite measurements such as from MODIS and VIIRS can complement the angle dependencies of the ground BRDF [1,2].

Although great care is taken to ensure the time of the ground measurements corresponds to the time of the overpass sounding, this has proven to be logistically difficult. Due to the Earth's rotation, a ten-minute difference already corresponds to a 2.5 degree change in the position of the Sun. As such, the introduction of a reference measurement does effectively introduce time dependencies in the viewing angles discussed above [8].

Absolute calibration does allow the calibration and/or stability to be related to other instruments.

### 3.5 BRDF

The largest non-instrumental influence on variations of measured radiance[2] is caused by the properties of the reflective surface. Expressed as the Bi-directional Reflectance Distribution Function, or BRDF, this determines the scattering properties of a parallel beam of incident sunlight into all directions of the hemisphere as seen from the ground. If the BRDF is devoid of dependencies (i.e. light is scattered equally to all directions) the BRDF is reduced to a dependency to amount of light absorbed, and the surface is considered to be Lambertian.

However, surfaces of invariant sites are not perfectly Lambertian, as shown in Fig. 4 (reproduced from [1]). Measurements show a clear dependency on the zenith and azimuth angles, resulting in preferential off-nadir directions in which light is reflected.

BRDF is also a property of the surface composition, and thus location dependent.

[2] Note that although spectrometers are used throughout this work, we assume the radiance to be obtained from a line-free part of the spectrum. For the SWIR wavelength range as observed by TROPOMI this corresponds to a region between 2312.7 and 2313 nm.

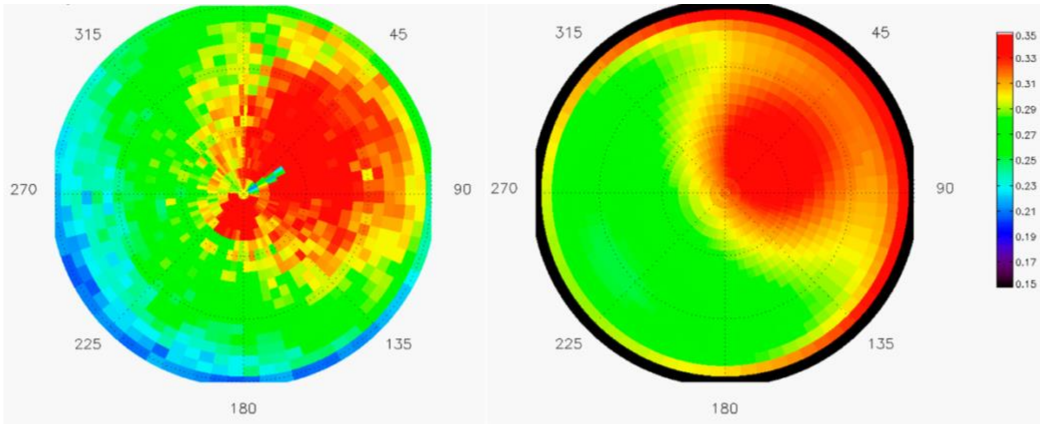


Fig. 4 (taken from [1]): PARABOLA raw data as acquired on 28 June 2018, 581 nm band at the Railroad Valley Playa. The Sun is in the southwest, and the hotspot is in the northeast in the back-scatter direction; (Right) data smoothed following a modified Rahman–Pinty–Verstraete (mRPV fit) [11].

## 4 RESULTS ON RELATIVE RADIOMETRY

Using the Saharan, Namib and Arabian desert sites, a first study was done in [12] to attempt to verify the instrument stability seen in Fig. 1. Using the assumption that the desert surface was Lambertian, time series were made over the first two years of the nominal operations of TROPOMI SWIR. Figure 5 shows the results over four sites: Egypt1, Algeria3, Libya3 and Namibia_PICS_AND1.

Two conclusions are clear from this data. First, the instrument stability could be independently verified to a value less than 0.8%, about a factor 10 worse than the observed instrument degradation over the probed period (see Fig. 6). Second, the spread in measurements was significant (4%) revealing the assumption of a Lambertian surface to be relatively poor and impacting the method. Recently, this analysis over this sample of desert PICS was being repeated with six years' worth of data instead of the initial two years.

Figure 7 shows this update over the time series over Egypt 1. A fitted slope of 0.06% per year is found, much flatter with the longer period, again verifying the lack of instrument degradation. Figure 8 shows the distribution of slopes and spreads (assuming a Lambertian surface) over the lifetime of TROPOMI. Here we derive limits of 0.08% per year and a median standard deviation of 4.5%. These results validate both the strong constraints on instrument degradation obtained from TROPOMI itself, as well as the initial results from [12]. They also indicate the spread remains significant and does not reduce in time. Thus this originates in the assumption of a Lambertian surface.

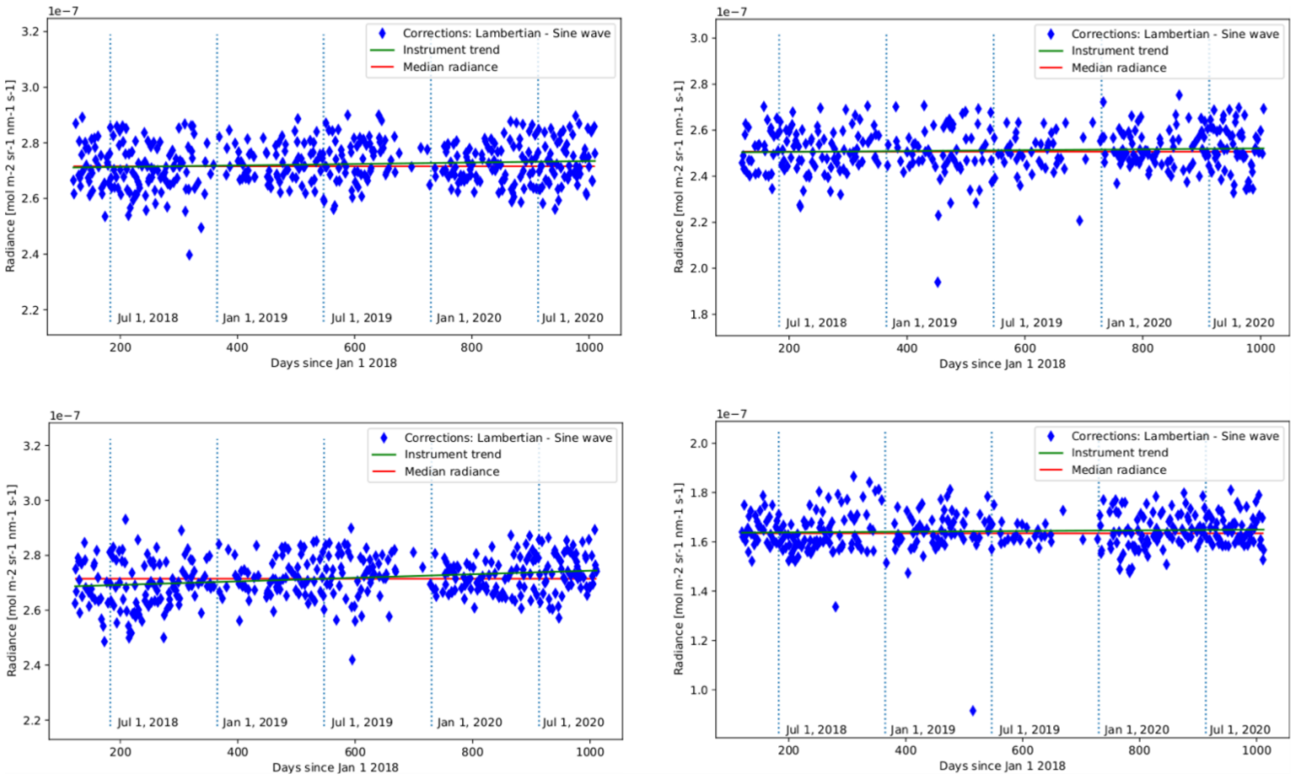


Fig. 5 (taken from [12]): Time series of the radiance of four PICS (Egypt1 in the top left, Algeria3 in the top right, Libya3 in the bottom left, Nambian_PICS_AND1 in the bottom right) as measured by TROPOMI-SWIR.

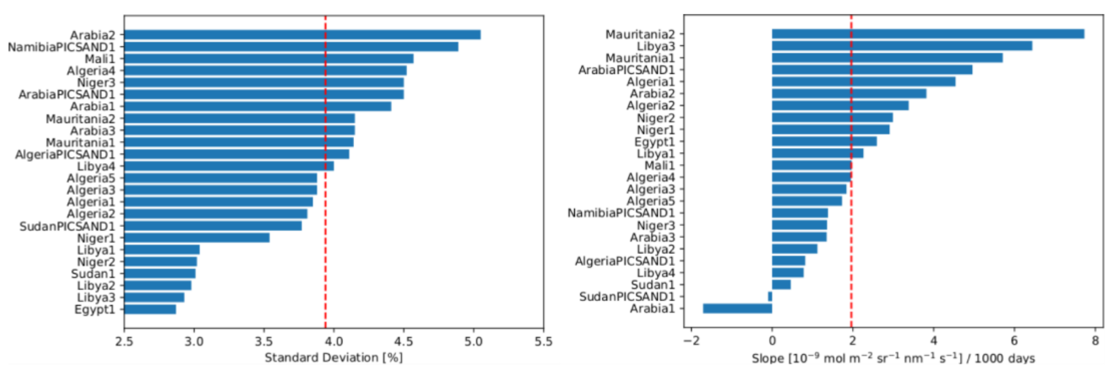


Fig. 6 (taken from [12]): The standard deviation (left) and fitted slope of the time series over a number of PICS. The slope provides an upper limit to an independently verified instrument degradation (0.8%) and the standard deviation indicates that the assumption of a Lambertian surface may limit interpretation.

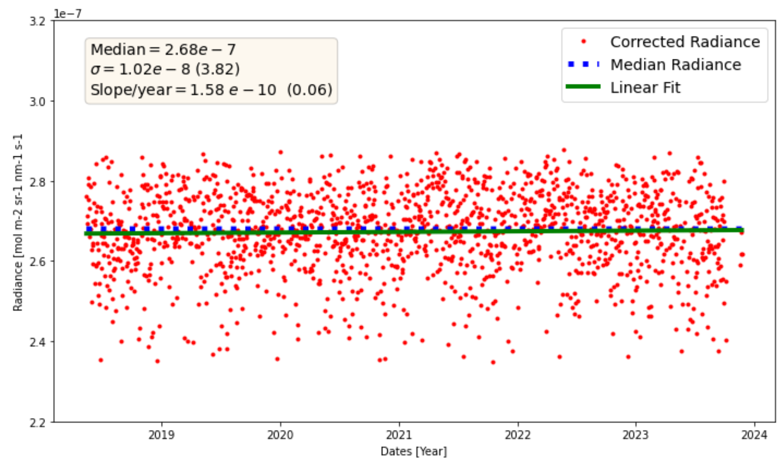


Fig. 7: Time series of the corrected radiance of TROPOMI-SWIR measured over Egypt (filtered for clouds) over the full mission lifetime. A linear fit is given in green, showing an increase with a slope of 0.06% /year.

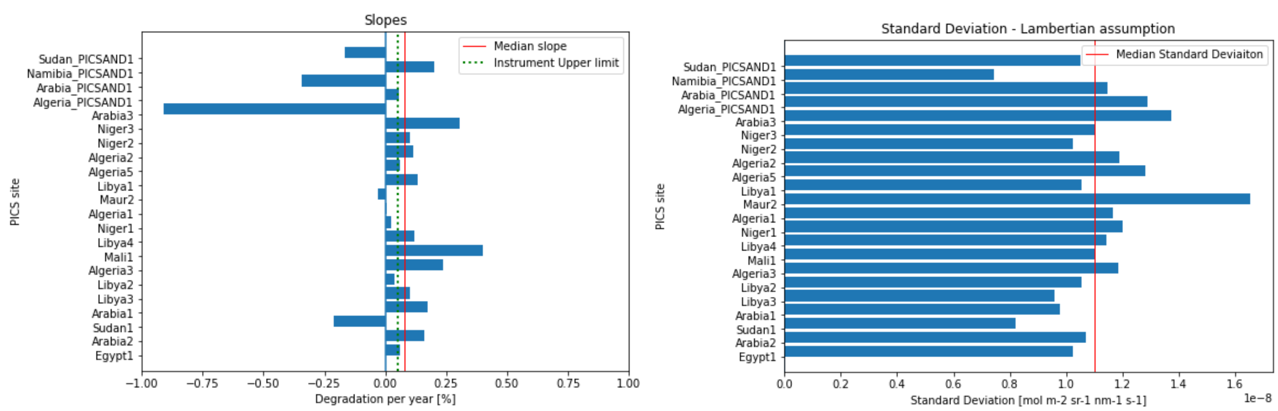


Fig. 8: Overview of the derived slopes and standard deviations of the PICS sites over the full mission lifetime.

## 5 ABSOLUTE RADIOMETRY

Absolute calibration is not possible at any of the Saharan sites due to accessibility constraints, either natural or political. Campaigns with ground equipment are regularly carried out at RRV by the GOSAT and OCO instrument teams [1, 2, 4, 8].
These data contain the reflectance at nadir, which can be used to simulate a 2.313 micron continuum radiance. At this wavelength, little to no (<1%) molecular absorption exists from the various species.

The BRDF as a function of solar and viewing angles was obtained from three sources.

1) A theoretical model using the mRPV model [11]. This follows the equation:

$$BRF\ (\theta_{sol}, \theta_i, \varphi) = r_0 (\cos\theta_i \cos\theta_{sol}\ (\cos\theta_i + \cos\theta_{sol}))^{k-1}\ h\ e^{(-b\cos\xi)}$$

where

$$h = 1 + \frac{1 - r_0}{1 + G}$$

$$G = \sqrt{\tan^2\theta_i + \tan^2\theta_{sol} - 2\tan\theta_i \tan\theta_{sol}\cos\varphi}$$

$$\cos\xi = \cos\theta_i \cos\theta_{sol} + \sin\theta_i \sin\theta_{sol}\cos\varphi$$

$$\varphi = \varphi_{sol} - \varphi_i$$

In these equations, $\theta$ is the zenith angle and $\varphi$ the azimuth angle. *Sol* represents the sun, *i* the instrument or viewing direction. $r_0$, *k* and *b* are free parameters.
These free parameters originate in the surface composition. Although theoretical assumptions can be made, it was preferred to obtain them from satellite measurements of the MISR instrument at 886 nm [1]. It is assumed the free parameters are invariant with respect to wavelength [1].

2) The BRDF product of MODIS at 2.1 micron (Product MCD43A1).
3) The BRDF product of VIIRS at 2.1 micron (Product VNP143IA1)

For both 2) and 3), one must either correct for the wavelength difference to 2.3-micron or adopt the same assumption as for 1). More importantly, one must as well consider how to average or co-add the representative area of RRV relative to the TROPOMI-SWIR pixel. After deriving a simulated spectrum using the measured reflectances with corrections for BRDF effects and differences in measurement times, the ratio of the simulated to measured radiance can used. Figure 9 shows an example of the two spectra.

Regardless of origin (mRPV, MODIS or VIIRS), the correction factor is obtained by taking ratio of the BRDF at nadir over the BRDF at the instrument viewing angle. The solar illumination angles are the same for both. Figure 10 shows the ratio of the corrected simulated radiance over the observed radiance as a function of either time (top row) or instrument viewing angle (bottom row). This is done for all three BRDF origins in the three columns; mRPV at the left, MODIS in the middle and VIIRS to the right.

The results show that the mRPV model is clearly superior over the satellite measurements of MODIS and VIIRS. Ignoring clear outliers (i.e. a ratio larger than two, likely attributed to saturation in MODIS and/or VIIRS measurements), the median absolute difference is 6.1% for the mRPV, 11.9% for VIIRS and 15.5% for MODIS .

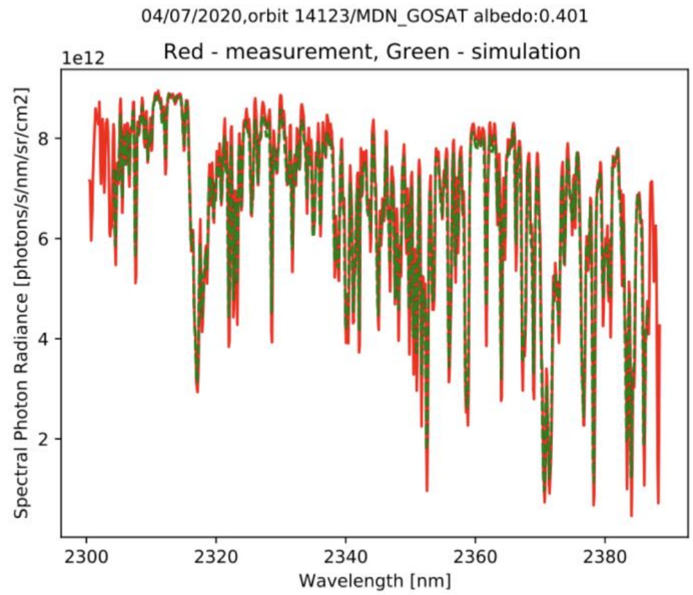


Fig. 9: Comparison of the simulated (green) and measured (red) 2.3 micron spectrum over RRV on July 4th 2020. The measurement originates from TROPOMI-SWIR, the simulated spectrum uses the ground measurement of the OCO-GOSAT team and various correction factors for the BRDF that uses MISR measurements.

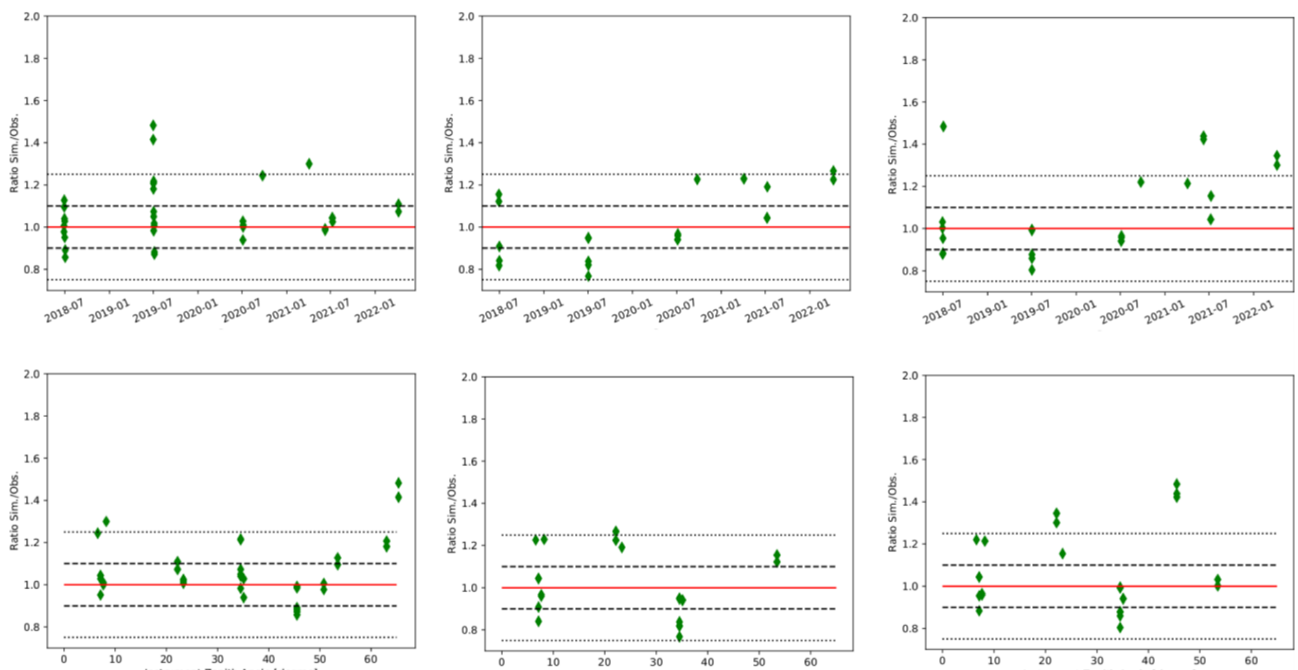


Fig. 10: Ratio of simulated over measured 2.313 micron radiance. The top row shows the variation in time, and the bottom row with viewing zenith angle. The BRDF effects are corrected with a mRPV (left column), MODIS (middle column) and VIIRS (right column).

It is also clearly visible that the large viewing angles possible due to the large TROPOMI swath are not representative of RRV. Significant deviations are seen for the larger angles. This has two origins:

1) The BRDF is not well represented at large angle by either the VIIRS and/or MODIS measurements. Even the mRPV model may not correctly describe the behaviour at large (60 degrees and higher) angle.
2) Due to the effective dependency of the size of the TROPOMI pixel on viewing angle, the observed area for pixels at large angles include a significant region outside of the RRV playa and include the mountains surrounding the playa. Both the reflectance and overall BRDF are not correct w.r.t. the measured playa reflectance.

The resulting 6% for the mRPV model (which can be improved to 4.5% if one considers and removes outliers for large angles or other origins) is in line with the accuracy seen for OCO-2 and GOSAT [2,4].

# 6 CONSIDERATIONS FOR SMALL-SAT APPLICATIONS

## 6.1 Limitations of TROPOMI

Despite its success, the TROPOMI-SWIR instrument has been pushed to its limits regarding vicarious calibration in the results presented above. These limitations are caused almost exclusively by a single instrumental property: TROPOMI-SWIR pixels are large, 5.5x7 $km^2$ at nadir and significantly larger at the edge of the swath. This limitation has two possible effects relating to the assumption of homogeneity and isotropy that were adopted at the start.

1) A reflectance and/or BRDF measured from the ground is not always representative for the average value across the TROPOMI pixel. E.g., the extend of RRV barely fills one pixel. Different terrain from the surroundings dilutes the measured values.
2) Even if the site is relatively homogeneous (i.e. the same surface material), height differences (e.g. sand dunes) impact the optical path as the surface observed likely is not flat at these scales. Local inclination may complicate the reflective properties.
   Although the TROPOMI overpass times are near midday, inclined slopes may already cause shadowing. See e.g. the satellite image of NAMIBIA_PICSAND1 in Fig. 6 of [3].
   Shadowing effects or inclinations are not included in the representation by the BRDF products of MODIS or VIIRS. Neither are they treated in the mRPV equations.
   Effects are clearly dependent on the solar zenith angle, which itself is dependent on the overpass time of a satellite as well as the latitude of the PICS site.

Soundings with <u>a small solar zenith angle</u> (i.e. near the equator and/or near a early afternoon crossing time) are thus recommended.

## 6.2 Viewing Angles

Related to limitation on large TROPOMI pixels are the instrument viewing angles of TROPOMI. These also become large when the target location is at the edge of the swath. For small-sat applications, it is recommended that calibration soundings are done <u>with a viewing zenith angle as small as possible (i.e. near nadir)</u>. This has two benefits:

1) It minimizes the uncertainties in (various) BRDF models, which dominate correction factors. BRDF models are more reliable near nadir. In addition, the correction factor is much smaller

near nadir measurements (note that the correction factor is one by definition for nadir measurements).

2) It facilitates the repeatability of a calibration measurement, effectively reducing the dependency of the reflectance to two instead of four parameters (i.e. only the solar angles and not the viewing angles). Most of the observed uncertainty (i.e. standard deviation of a time series in Figure 5, or the systematic differences in Fig. 9) originate from the difference in viewing angle between TROPOMI and the nadir reflectance. Respective soundings taken at significantly different viewing angles are more difficult to compare.

### 6.3 Aerosol Optical Depth

One parameter that has not yet been discussed in-depth is the aerosol optical depth. For the 2.3-micron window, the effective aerosol optical depth (AOD) over various sites can be safely ignored due to its very low value. Although included in the radiative transfer calculations to simulate a ToA spectrum, its effects are negligible at longer wavelengths.

However, this may not be the case for missions at shorter wavelengths. Instruments observing in the visible (e.g. $NO_2$), near-IR ($O_2$-A band) or possibly even in the SWIR-1 band should consider the effects of aerosols and include the AOD in the reproduction of the ToA spectra.

Of note is that it is especially important the day-to-day *variations* of the AOD need to be included correctly when considering (long) time series. Variations of AOD linearly increase or decrease the spread seen in time series.

### 6.4 Spatial resolution and Homogeneity

Besides the surface BRDF (which can be improved by adopting the mRPV model), one of the main contributors to the uncertainty in vicarious calibration is the assumed homogeneity of a PICS surface[3]. Sites such as RRV are clearly not. Much higher spatial resolution may thus improve results focussing on a truly homogeneous site. In [3] site homogeneity at scales of 20 and 100 km was quantified. Spatial homogeneity (*SHOM*) at scales of *X* kilometres was derived using the MODIS MCD43A3 product (white-sky albedo) following their equation (1):

$$SHOM_\lambda\ (X\ km) = 100\ \frac{\sigma(\rho_\lambda - \overline{\rho_{\lambda,X\ km}})}{\overline{\rho_{\lambda,X\ km}}}$$

In this equation, $\rho_\lambda$ is the surface reflectance and $\sigma$ the standard deviation. $\overline{\rho_{\lambda,X\ km}}$ is the mean reflectance in spectral band λ over its surrounding neighborhood within *X* kilometers distance. Note that the original resolution of the surface reflectance in the MODIS product is 500 meters.

The final number *SHOM* was a percentage, with 0 % corresponding to a completely homogeneous surface. For the adaptation done in [12] it was assumed that although the surface reflectance $\rho_\lambda$ is dependent on wavelength, the measured inhomogeneity is not. I.e., Even with changes in the surface reflectance, similar changes occur over $\overline{\rho_{\lambda,X\ km}}$ .

In [3] homogeneity improved at smaller scales; I.e., $SHOM_{100} > SHOM_{20}$. To estimate the homogeneity at even smaller spatial scales (e.g., 300 meter scale), we consider a limit in which the

[3] In truth, this should apply to the full extent of the point spread function.

homogeneity improves linearly with the same improvement seen from 100 km to 20 km applied to 300 meters, until either an upper limit of *SHOM* = 0 is reached or $SHOM_{0.3}$ converges.

One can estimate the level of improvement that is possible for small-sat applications using an estimation of $SHOM_{0.3}$. This would linearly scale the values found for TROPOMI with the improvement of $SHOM_{0.3}$ w.r.t. $SHOM_{20}$. An example is shown in Fig. 11 for Egypt1. Here the linear interpolation results in a $SHOM_{0.3}$ of 0.40, 25% better than $SHOM_{20}$.

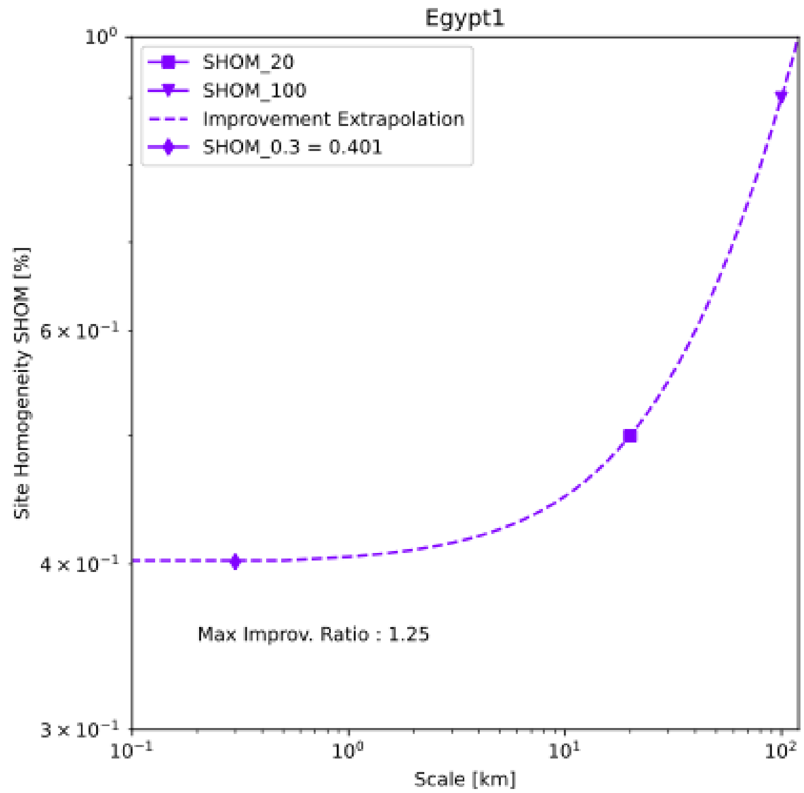


Fig. 11: Extrapolation of the spatial homogeneity from 20 and 100 km scales to 300 meter scales (i.e. the TANGO spatial scale).

Mathematically this affects the vicarious calibration values such that:

$$Ratio_{0.3} = Ratio_{TROPOMI} / IF\_SHOM_{0.3}$$

$$STDEV_{0.3} = STDEV_{TROPOMI} / IF\_SHOM_{0.3}$$

With

$$IF\_SHOM_{0.3} = SHOM_{20} / SHOM_{0.3}$$

And

$SHOM_{0.3}$ linearly extrapolated from $SHOM_{20}$ and $SHOM_{100}$.

$Ratio_{0.3}$ and $STDEV_{0.3}$ are estimations for small-sat instruments.

In the results above, *$Ratio_{Tropomi}$* and *$STDEV_{TROPOMI}$* are 6.1% and 5.5% respectively for RRV, an *$IF\_SHOM_{0.3}$* of 1.5 (i.e. an improvement of 50%) would reduce *$Ratio_{0.3}$* and *$STDEV_{0.3}$* to 3% and 2.8%.

In Table 1 and Figure 12, we give the possible improvement factors for the sample of PICS from [12] with the observed spread and a possible spread. One should take note that a bias is introduced. Libya3, Arabia3, Niger1, Arabia_PICSAND1 or Namibia_PICSAND1 appear to show dramatic improvements. These are not physical, but dominated by a lack of homogeneity at large scales.

| SITE | SHOM 0.3 | RATIO | SPREAD [1E$^{-8}$ MOL M$^{-2}$ SR$^{-1}$ NM$^{-1}$ S$^{-1}$] | POT. SPREAD [1E$^{-8}$ MOL M$^{-2}$ SR$^{-1}$ NM$^{-1}$ S$^{-1}$] |
|---|---|---|---|---|
| Arabia2 | 0.15 | 1.33 | 1.07 | 0.80 |
| Sudan1 | 0.33 | 1.53 | 0.82 | 0.54 |
| Arabia1 | 0.15 | 2.60 | 0.98 | 0.38 |
| Egypt1 | 0.40 | 1.25 | 1.02 | 0.82 |
| Libya3 | 0* | Inf | 0.96 | n.a. |
| Libya2 | 0.23 | 1.76 | 1.05 | 0.60 |
| Algeria3 | 0.53 | 1.33 | 1.18 | 0.89 |
| Maur1 | 0.53 | 1.14 | n.a. | n.a. |
| Mali1 | 0.27 | 1.33 | 1.10 | 0.83 |
| Libya4 | 0.50 | 1.20 | 1.14 | 0.95 |
| Niger1 | 0.08* | 3.82 | 1.20 | n.a. |
| Algeria1 | 0.30 | 1.65 | 1.17 | 0.71 |
| Maur2 | 0.45 | 1.33 | 1.65 | 1.24 |
| Algeria4 | 0.16* | 3.22 | n.a. | n.a. |
| Libya1 | 0.43 | 1.40 | 1.05 | 0.75 |
| Algeria5 | 0.58 | 1.38 | 1.28 | 0.93 |
| Algeria2 | 0.45 | 1.54 | 1.19 | 0.77 |
| Niger2 | 0.28 | 2.14 | 1.02 | 0.48 |
| Niger3 | 0.71 | 1.70 | 1.10 | 0.65 |
| Arabia3 | 0* | Inf | 1.37 | n.a. |
| Algeria_PICSAND1 | 0.48 | 1.25 | 1.29 | 1.03 |
| Arabia_PICSAND1 | 0.08* | 6.14 | 1.15 | n.a. |
| Namibia_PICSAND1 | 0* | Inf | 0.74 | n.a. |
| Sudan_PICSAND1 | 0.45 | 1.54 | 1.05 | 0.68 |

*Table 1:* Homogeneity at 300 meters, the improvement ratio, and measured and potential spread. *Site dominated by large systematic differences between $SHOM_{20}$ and $SHOM_{100}$. No improvement to be derived.

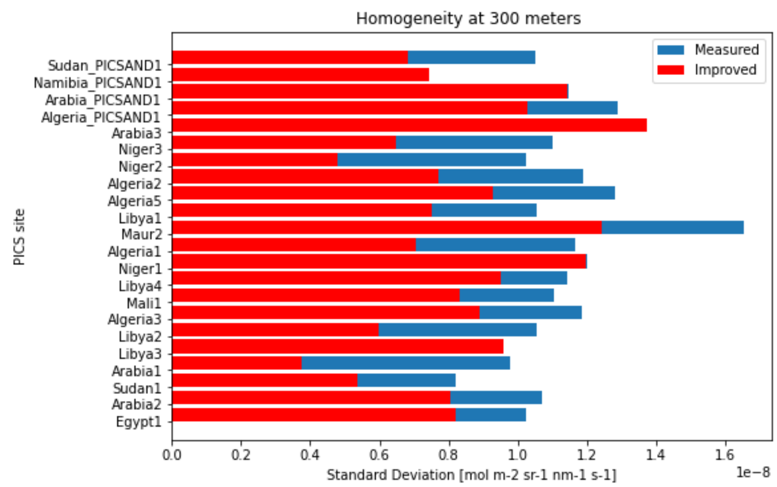


Fig. 12: Standard deviation as measured by TROPOMI-SWIR and the potential improvement due to increased homogeneity.

## 7 CONCLUSIONS

In this contribution, the stability of the TROPOMI-SWIR module was used to explore the applicability of vicarious calibration to small-sat instruments such as the planned TANGO mission. The main conclusions are:

- Many of the limitations and uncertainties for TROPOMI-SWIR vicarious calibration derive from its large pixel size.
- Uncertainties appear to be dominated by the non-Lambertian effects described in the BRDF of desert surfaces. Theoretical approximations such as the mRPV model appear to best describe the reflectances seen over deserts.

For small-sat instruments three recommendations can be made:

- Soundings taken at or very close to, nadir are strongly preferred. Nadir measurements eliminate many BRDF effects with respect to ground measurements. Accurate treatment of the BRDF effects also minimizes many of the necessary correction factors. The mRPV model is preferred.
- Homogeneity of a PICS at the intended spatial resolution should be known before implementation. Choice of purely homogeneous sites should improve the reliability of monitoring instrument stability.
- Improvements for small-sat vicarious calibration applications w.r.t. TROPOMI should be made by implementing both (i.e., homogeneity and BRDF)

## 8 REFERENCES

[1] Bruegge, C.J., Coburn, Elmes, et al., *Bi-Directional Re- flectance Factor Determination of the Railroad Valley Playa*, Remote Sensing, 2019, 11, 2601

[2] Bruegge C.J., Chrien N.L., Ando R.R., et al., *Vicarious Cal- ibration of Orbiting Carbon Observatory-2,* IEEE Transactions on Geoscience and Remote Sensing, 2019, 57, 5135

[3] Bacour C., Briottet X., Bréon F.-M., et al., *Revisiting Pseudo Invariant Calibration Sites (PICS) Over Sand Deserts for Vicarious Calibration of Optical Imagers at 20 km and 100 km Scales***,** Remote Sensing, 2019, 11, 1166

[4] Kuze A., Taylor T.E., Kataoka, F., et al., *Long-Term Vicarious Calibration of GOSAT 45 Short-Wave Sensors: Techniques for Error Reduction and New Estimates of Radiometric Degradation Factors*, IEEE Transactions on Geoscience and Remote Sensing, 2014, 52, 3991

[5] Veefkind J.P., Aben I., McMullan K., et al., *TROPOMI on the ESA Sentinel-5 Precur- sor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications*, RSE, 2012, 120, 70

[6] van Kempen T.A., van Hees R.M., Tol. P.J.J., et al., *In-flight calibration and monitor- ing of the Tropospheric Monitoring Instrument (TROPOMI) short-wave infrared (SWIR) module*, Atmospheric Measurement Techniques, 2019, 12, 6827

[7] Ludewig A., Kleipool Q., Bartstra R., et al., *In-flight calibration results of the TROPOMI payload on board the Sentinel-5 Precursor satellite*, Atmospheric Measurement Techniques, 2020, 13, 3561

[8] van Kempen T.A., Rotmans T.J., van Hees R.M, et al. *Vicarious Calibration of the TROPOMI-SWIR module over the Railroad Valley playa*, Atmospheric Measurement Techniques, 2023

[9] Landgraf J., Rusli S., Cooney R., et al., *The TANGO mission: A satellite tandem to measure major sources of anthropogenic greenhouse gas emissions*, EGU General Assembly, 2020, EGU2020-19463

[11] Rahman H., Verstraete M.M. and Pinty B., *Coupled surface-atmosphere reflectance (CSAR) model: 1. Model description and inversion on synthetic data*, Journal of Geophysical Research, 1993, 98 D11, 20779

[12] van Kempen T.A., Oggionni F. and van Hees R.M, *Monitoring the Tropospheric Monitoring Instrument (TROPOMI) short-wave infrared (SWIR) module instrument stability using desert sites*, AMT, 2021, 14, 6711

[13] Bouvet, Thome, Berthelot, et al., *A Radiometric Calibration Network for Earth Observing Imagers Operating in the Visible to Shortwave Infrared Spectral Range*, Remote sensing, 2019, 11, 2401

[14] Hu H., Hasekamp O., Butz A., et al., *The operational methane retrieval algorithm for TROPOMI,* Atmospheric Measurement Techniques, 2016, 9, 5423

[15] Borsdorff T, aan de Brugh J., Hu H., et al., *Mapping carbon monoxide pollution from space down to city scales with daily global coverage*, AMT, 2018, 11, 5507

[16] Hu H., Landgraf J., Detmers. R., et al., *Toward Global Mapping of Methane With TROPOMI: First Results and Intersatellite Comparison to GOSAT*, Geophysical Research Letters, 2018, 45, 3682

[17] van Amerongen A.H., Visser H., Vink R.J.P. et al., *Development of immersed diffraction grating for the TROPOMI-SWIR spectrometer*, Proceedings of SPIE, 2010, 7826, 1