| Customer: | ESA | Document Ref.: | OMPC.ACR.APR.002 |
| Contract No.: | 4000136252/21/I-BG | Date: | 17/03/2023 |
| Issue: | 1.1 |

| Project: | COPERNICUS SPACE COMPONENT SENTINEL OPTICAL IMAGING MISSION PERFORMANCE CLUSTER SERVICE |
| Title: | Annual Performance Report |
| Author(s): | S3 OLCI, SLSTR and SYN ESL team |

| Approved by: | L. Bourg, OPT-MPC OLCI-SYN ESL Coordinator |
| | D. Smith, OPT-MPC SLSTR ESL Coordinator |
| | S. Clerc, OPT-MPC Optical ESL Coordinator |
| Authorized by | C. Henocq, OPT-MPC S3 Technical Manager |
| | J. Bruniquel, OPT-MPC Service Manager |

| Distribution: | ESA, published in Sentinel Online |
| Accepted by ESA | S. Dransfeld, ESA TO |

| Filename | OMPC.ACR.APR.002 - i1r1 - S3 Optical Annual Performance Report 2022.docx |

Disclaimer

The views expressed herein can in no way be taken to reflect the official opinion of the European Space Agency or the European Union.
## Changes Log

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>28/02/2023</td>
<td>First version</td>
</tr>
<tr>
<td>1.1</td>
<td>17/03/2023</td>
<td>Version updated to take account of ESA comments</td>
</tr>
</tbody>
</table>

## List of Changes

<table>
<thead>
<tr>
<th>Version</th>
<th>Section</th>
<th>Answers to RID</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>6.3.2.2</td>
<td></td>
<td>Clarification of the caption of Figure 150</td>
</tr>
</tbody>
</table>
|         | 7.2.3.1 |                | Clarification related to Figure 180 and Figure 181  
|         |         |                | Correction of caption of Figure 185          |
# Table of content

1 INTRODUCTION .................................................................................................................. 1
   1.1 SCOPE OF THE DOCUMENT .................................................................................. 1
   1.2 APPLICABLE AND REFERENCE DOCUMENTS .................................................. 1
   1.3 ACRONYMS AND ABBREVIATIONS ..................................................................... 1

2 EXECUTIVE SUMMARY .................................................................................................... 2
   2.1 OLCI ......................................................................................................................... 2
   2.2 SLSTR ....................................................................................................................... 4
      2.2.1 SLSTR-A .......................................................................................................... 4
      2.2.2 SLSTR-B .......................................................................................................... 5
   2.3 SYN .......................................................................................................................... 7

3 PROCESSING BASELINE DESCRIPTION ....................................................................... 8
   3.1 OLCI ......................................................................................................................... 8
   3.2 SLSTR ....................................................................................................................... 9
   3.3 SYN .......................................................................................................................... 10

4 CALIBRATION AND CHARACTERISATION CHANGES ................................................ 12
   4.1 OLCI ......................................................................................................................... 12
      4.1.1 Instrument settings .......................................................................................... 12
      4.1.2 Evolutions in Radiometric Calibration of EO data ....................................... 12
   4.2 SLSTR ....................................................................................................................... 12
      4.2.1 Instrument settings .......................................................................................... 12
      4.2.2 Evolutions in Radiometric Calibration of EO data ....................................... 12

5 SUMMARY OF PERFORMANCES – OLCI ................................................................. 13
   5.1 INSTRUMENT PERFORMANCES ........................................................................ 13
      5.1.1 Temperature stability ..................................................................................... 13
      5.1.2 Signal to noise ratio ...................................................................................... 15
      5.1.3 Spectral Calibration ....................................................................................... 21
      5.1.4 Radiometric stability ..................................................................................... 26
      5.1.5 Ageing of radiometric diffuser ..................................................................... 30
   5.2 L1 PRODUCTS PERFORMANCES ...................................................................... 35
      5.2.1 Geometric Performance ............................................................................... 35
      5.2.2 Radiometric Gain Model Performance ....................................................... 43
5.2.3 Radiometric Validation

5.3 L2 PRODUCT PERFORMANCES

5.3.1 Level 2 cloud screening

5.3.2 Integrated Water Vapour (IWV)

5.3.3 OLCI Global Vegetation Index (OGVI), a.k.a. GIFAPAR, and OLCI Terrestrial Chlorophyll Index (OTCI).

6 SUMMARY OF PERFORMANCES — SLSTR

6.1 INSTRUMENT PERFORMANCES

6.1.1 Instrument temperatures

6.1.2 Detector Temperatures

6.1.3 Scanner performance

6.1.4 Black- Bodies

6.1.5 IR Channels

6.1.6 VIS/SWIR Channels

6.2 L1 PRODUCTS PERFORMANCES

6.2.1 TIR Channel Calibration

6.2.2 VIS/SWIR Channel Calibration

6.2.3 Geometric Calibration

6.2.4 Cloud Screening

6.2.5 References

6.3 L2 PRODUCT PERFORMANCES

6.3.1 Land Surface Temperature (LST)

6.3.2 Fire Radiative Power (FRP)

7 SUMMARY OF PERFORMANCES — SYN

7.1 L1 PRODUCTS PERFORMANCES

7.2 L2 PRODUCT PERFORMANCES

7.2.1 SYN L2 SDR quality assessment: intercomparisons with other missions

7.2.2 SYN VGT: consistency evaluation with SPOT/Vegetation and PROBA-V

7.2.3 SYN L2 global Aerosol performance

8 PROBLEMS ENCOUNTERED IN THE REPORTING PERIOD

8.1 PRODUCT NOTICES REPORTS

8.1.1 OLCI

8.1.2 SLSTR
8.1.3 SYN ......................................................................................................................... 239
8.2 INSTRUMENT ANOMALIES ......................................................................................... 239
  8.2.1 OLCI .......................................................................................................................... 239
  8.2.2 SLSTR ....................................................................................................................... 240
  8.2.3 SYN .......................................................................................................................... 241
List of Figures

Figure 1: Daytime combined SLSTR-A and SLSTR-B Level-1 image for SWIR channels on 15th November 2022.--------------------------------- 6

Figure 2: long term monitoring of CCD temperatures using minimum value (top), time averaged values (middle), and maximum value (bottom) provided in the annotations of the Radiometric Calibration Level 1 products, for the Shutter frames, all radiometric calibrations so far.--------------------------------- 13

Figure 3: long term monitoring of OLCI-B CCD temperatures using minimum value (top), time averaged values (middle), and maximum value (bottom) provided in the annotations of the Radiometric Calibration Level 1 products, for the Shutter frames, all radiometric calibrations so far except the first one (absolute orbit 167) for which the instrument was not yet thermally stable.--------------------------------- 14

Figure 4: OLCI-A Signal to Noise ratio as a function of the spectral band for the 5 cameras. These results have been computed from radiometric calibration data. All calibrations except first one (orbit 183) are present with the colours corresponding to the orbit number (see legend). The SNR is very stable with time: the curves for all orbits are almost superimposed. The dashed curve is the ESA requirement. ------------ 16

Figure 5: OLCI-A long-term stability of the SNR estimates from Calibration data, example of channel Oa01. -------------------- 17

Figure 6: OLCI-B Signal to Noise ratio as a function of the spectral band for the 5 cameras. These results have been computed from radiometric calibration data. All calibrations except first one (orbit 167) are present with the colours corresponding to the orbit number (see legend). The SNR is very stable with time: the curves for all orbits are almost superimposed. The dashed curve is the ESA requirement. ---- 19

Figure 7: OLCI-A across track spectral calibration from all S02/S03 sequences since the beginning of the mission. Upper top plot is spectral line 1; middle plot is spectral line 2 and bottom plot spectral line 3. On-ground spectral characterisation is in red and nominal spectral assignment is in red dotted line. -------- 22

Figure 8: OLCI-A camera averaged spectral calibration evolution as a function of absolute orbit number (all spectral S02/S03 calibrations since the beginning of the mission are included except the very first one (1rst March 2016, orbit 195)). The data are normalized with the first Spectral Calibration of the plot, which is from 18 April 2016 (orbit 881). The last spectral Calibration is from 07 January 2023 (orbit 35886/35887). ------------------------------------------ 23

Figure 9: OLCI-A line-averaged spectral calibration relative to the one acquired on 4th May 2016 (orbit 1107), as a function of time derived from all S09 sequences. The last calibration is from 07 January 2023 (orbit 35885). For each camera, the spectral evolution derived from spectral lines at 485 nm, 656 nm, 770 nm and 854 nm have been averaged. ------------------------------------------------ 23

Figure 10: OLCI-B across track spectral calibration from all S02/S03 sequences since the beginning of the mission. Upper top plot is spectral line 1; Middle plot is spectral line 2 and bottom plot spectral line 3. On-ground spectral characterisation is in red and nominal spectral assignment is in red dotted line. ------- 24

Figure 11: OLCI-B camera averaged spectral calibration evolution as a function of absolute orbit number (all spectral S02/S03 calibrations since the beginning of the mission are included). The data are normalized
with the first Spectral Calibration. The first (reference) calibration is from 8 May 2018 (orbit 182), the last 17 Jan. 2023 (orbit 24634). For each camera, the spectral evolution derived from spectral lines at 485 nm, 656 nm, 770 nm and 854 nm have been averaged. The data are normalized with the first Spectral Calibration. For better comparability, the orbit range (abscissa) is equal to that of OLCI-A (Figure 8).

Figure 12: OLCI-B camera averaged spectral calibration evolution as a function of absolute orbit number from S09 calibrations since the beginning of the mission. The first (reference) calibration is from 8 May 2018, the last is from 17 Jan. 2023 (orbit 24634). For each camera, the spectral evolution derived from spectral lines at 485 nm, 656 nm, 770 nm and 854 nm have been averaged. The data are normalized with the first Spectral Calibration. For better comparability, the orbit range (abscissa) is equal to that of OLCI-A (Figure 8).

Figure 13: OLCI-A camera-averaged instrument evolution since channel programming change (25/04/2016) and up to most recent calibration (26/01/2023) versus wavelength.

Figure 14: OLCI-A camera averaged gain relative evolution with respect to calibration of 25/04/2016 (change of OLCI channel settings), as a function of elapsed time since launch; one curve for each band (see colour code on plots), one plot for each module. The diffuser ageing has been taken into account. Early mission data (16 Feb. to 25 April) is not available due to missing information required for accurate gain computation.

Figure 15: OLCI-A gain relative evolution with respect to calibration of 25/04/2016 (change of OLCI channel settings), as a function of elapsed time since launch (x axis) and spatial pixel (y axis) for Channel Oa2 (412.5 nm), Camera 1.

Figure 16: OLCI-B camera-averaged instrument evolution since channel programming change (18/06/2018) and up to most recent calibration (30/01/2023) versus wavelength.

Figure 17: OLCI-B camera averaged gain relative evolution with respect to first calibration after channel programming change (18/06/2018), as a function of elapsed time since the beginning of the mission; one curve for each band (see colour code on plots), one plot for each module. The diffuser ageing has been taken into account.

Figure 18: OLCI-A Diffuser 1 ageing as a function of wavelength (or spectral band). Ageing is clearly visible in spectral band #1 to #6.

Figure 19: Slope of ageing fit (% of loss per exposure) vs wavelengths, using all the available ageing sequence at the time of the most recent monthly Data Quality Report with an ageing measurement (January 2023 = red curve), and at the time of the 23 previous cycles with an ageing sequence (see legend above the curves).

Figure 20: OLCI-B Diffuser 1 ageing as a function of wavelength (spectral bands).

Figure 21: OLCI-B: Slope of ageing fit (% of loss per exposure) vs wavelengths, using all the available ageing sequence at the time of the current cycle (red curve) and at the time of previous cycle for which an ageing sequence was measured (see legend within the figure).

Figure 22: Ageing rates (in % per exposure) as derived using nominal methods (at cycles 28, 46 and 62), and successive upgrades (Yaw Manoeuvres and Equal SZA at 3 different dates).

Figure 23: overall OLCI-A georeferencing RMS performance time series over the whole monitoring period (left) and number of validated control points corresponding to the performance time series (right).
Figure 24: across-track (left) and along-track (right) georeferencing biases time series for Camera 1 (starting 01/03/2018). 36
Figure 25: same as Figure 24 for Camera 2. 37
Figure 26: same as Figure 24 for Camera 3. 37
Figure 27: same as Figure 24 for Camera 4. 37
Figure 28: same as Figure 24 for Camera 5. 38
Figure 29: OLCI-A spatial across-track misregistration at each camera transition (left) and maximum amplitude of the across-track error within each camera (left). 38
Figure 30: OLCI-A spatial along-track misregistration at each camera transition (left) and maximum amplitude of the along-track error within each camera (left). 39
Figure 31: overall OLCI-B georeferencing RMS performance time series (left) and number of validated control points corresponding to the performance time series (right) over the whole monitoring period. 40
Figure 32: across-track (left) and along-track (right) georeferencing biases time series for Camera 1. 40
Figure 33: same as Figure 32 for Camera 2. 40
Figure 34: same as Figure 32 for Camera 3. 41
Figure 35: same as Figure 32 for Camera 4. 41
Figure 36: same as Figure 32 for Camera 5. 41
Figure 37: OLCI-B spatial across-track misregistration at each camera transition (left) and maximum amplitude of the across-track error within each camera (left). 42
Figure 38: OLCI-B spatial along-track misregistration at each camera transition (left) and maximum amplitude of the along-track error within each camera (left). 42
Figure 39: RMS performance of the Gain Model of current Processing Baseline as a function of orbit. 44
Figure 40: For the 5 cameras: Evolution model performance, as camera-average and standard deviation of ratio of Model over Data vs. wavelength, for each orbit of the test dataset, including 17 calibrations in extrapolation, with a colour code for each calibration from blue (oldest) to red (most recent). 45
Figure 41: RMS performance of the OLCI-B Gain Model of the current processing baseline as a function of orbit. 46
Figure 42: For the 5 cameras: Evolution model performance, as camera-average and standard deviation of ratio of Model over Data vs. wavelength, for each orbit of the test dataset, including 10 calibrations in extrapolation, with a colour code for each calibration from blue (oldest) to red (most recent). 47
Figure 43: comparison of OSCAR and DIMITRI results for the various methods. 49
Figure 44: Time-series of the elementary ratios (observed/simulated) signal from S3A/OLCI for (top to bottom) Band Oa03 and band Oa18 respectively, over Six PICS Cal/Val sites. Dashed-green and orange
lines indicate the 2% and 5% biases respectively. Error bars indicate the desert methodology uncertainty.

Figure 45: Time-series of the elementary ratios (observed/simulated) signal from S3B/OLCI for (top to bottom) Band Oa03 and band Oa18 respectively over Six PICS Cal/Val sites. Dashed-green and orange lines indicate the 2% and 5% biases respectively. Error bars indicate the desert methodology uncertainty.

Figure 46: The estimated gain values for (top) S3A/OLCI and (bottom) S3B/OLCI from Glint, Rayleigh and PICS methods as a function of wavelength. We use the gain value of Oa8 from PICS method as reference gain for Sunglint method. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the methods uncertainties.

Figure 47: Ratio of observed TOA reflectance to simulated one for (black) MERIS, (pale-green) S2A/MSI, (white) S2B/MSI, (blue) S3A/OLCI, (green) S3B/OLCI, (red) S3A/SLSTR-NADIR, and (cyan) S3B/SLSTR-NADIR averaged over the six PICS test sites as a function of wavelength.

Figure 48. New CHL climatology for the Rayleigh calibration sites based CMEMS GlobColour products.

Figure 49: OSCAR Rayleigh S3A and S3B Calibration results for 2022 as a function of wavelength.

Figure 50: OSCAR Rayleigh OLCI-A Calibration results 2021 and 2022 as a function of wavelength.

Figure 51: OSCAR Rayleigh OLCI-B Calibration results for 2021 and 2022 as a function of wavelength.

Figure 52: OSCAR Glitter S3A and S3B Calibration results (adapted to Rayleigh result at 665 nm) for 2022 as a function of wavelength.

Figure 53: Sky camera setup.

Figure 54: Validation results of RF classifier for SC1.

Figure 55: Validation results of RF classifier for SC2.

Figure 56: 2021 test validation results for OLCI FR using reference data from SC1.

Figure 57: Comparison between PixBox and Sky Camera validation.

Figure 58: Confusion matrix showing validation results for OLCI L2 cloud screening including margin against SC1 automated classification for the year 2022.

Figure 59: Confusion matrix showing validation results for OLCI L2 cloud screening excluding margin against SC1 automated classification for the year 2022.

Figure 60: Validation results for OLCI A+B separated between harmonized and non-harmonized data, as well as all surfaces, as well as land and water surfaces.

Figure 61: Validation results for OLCI A separated between harmonized and non-harmonized data, as well as all surfaces, as well as land and water surfaces.

Figure 62: Validation results for OLCI B separated between harmonized and non-harmonized data, as well as all surfaces, as well as land and water surfaces.

Figure 63: Upper left: Scatter plot of the IWV products, derived from OLCI A above land and from SUOMI NET GNSS measurements. Upper right: Histogram of the difference between OLCI and GNSS (blue: original
OLCI, orange: bias corrected OLCI). Lower left: Temporal evolution of different quality measures (from top to bottom: systematic deviation factor, bias, root mean squared difference (with and without bias correction), explained variance (number in boxes are the numbers of matchups)). Lower right: Positions of the GNSS stations (grey: no valid matchup) 73

Figure 64: Upper left: Scatter plot of the IWV products, derived from OLCI A above land and from AMR MWR. Upper right: Histogram of the difference between OLCI and ARM (blue: original OLCI, orange: bias corrected OLCI). Lower left: Temporal evolution of different quality measures (from top to bottom: systematic deviation factor, bias, root mean squared difference (with and without bias correction), explained variance (number in boxes are the numbers of matchups)). Lower right: Position of ARM SGP. 74

Figure 65: Upper left: Scatter plot of the IWV products, derived from OLCI A above land and from AERONET. Upper right: Histogram of the difference between OLCI and AERONET (blue: original OLCI, orange: bias corrected OLCI). Lower: Positions of the used AERONET stations (grey: no valid match). 75

Figure 66: Scatter plot of the IWV products, derived from OLCI B above land and from SUOMI NET GNSS measurements (upper left), from ARM MWR (upper right) and AERONET (lower left). Lower right: Temporal evolution of different quality measures from the SUOMI-NET comparison (from top to bottom: systematic deviation factor, bias, root mean squared difference (with and without bias correction), explained variance (number in boxes are the numbers of matchups)). 77

Figure 67: Upper: Scatter plot of the IWV products, derived from OLCI (A left, B right) above ocean and from SUOMI NET GNSS measurements. Lower: Positions of the GNSS (A: left, B: right). 78

Figure 68: Upper: Scatter plot of the IWV products, derived from OLCI (A left, B right) above ocean and from Aeronet-OC v.3 L1.5 measurements. Lower: Positions of the used Aeronet match ups (A: left, B: right). 79

Figure 69: Time-series OGVI and OTCI and corresponding scatterplot of monthly mean for site BE-Braschaat, Belgium, land cover Needle-leaved, evergreen. A and C represent S3A; B and D represent S3B. 84

Figure 70: Time-series OGVI and OTCI and corresponding scatterplot of monthly mean for site DE-Haininich, Deutschland, land cover Broadleaved, deciduous, closed. A and C represent S3A; B and D represent S3B. 85

Figure 71: Time-series OGVI and OTCI and corresponding scatterplot of monthly mean for site FR-EstreesMons, France, land cover Cultivated and managed areas. A and C represent S3A; B and D represent S3B. 86

Figure 72: Comparison of OTCI-MTCI (a) and OGVI-MGVI (b). Points in the scatterplot represent the monthly mean of all available S3A and MERIS archive over 42 validation sites. Red and grey lines represent the modelled and 1:1 line respectively. The scatterplots are updated to include data up to May 2022. 87

Figure 73: Comparison of GBOV vs OLCI-A (left) and GBOV vs OLCI-B (right). Points in the scatterplot represent the monthly mean of all available S3A and GBOV data (3×3) over 13 validation sites. Red and black lines represent the modelled and 1:1 line, respectively. 89
Figure 74: Time series of GBOV FIPAR LP4 (red/green) and OLCI-A GIFAPAR (grey) for sites Steigerwaldt (left) and Guanica (right) ......................................................... 89
Figure 75: same as Figure 74 for sites Talladega (Left) and Tumbarumba (right) ................................................................. 90
Figure 76: same as Figure 74 for sites Bartlett (Left) and Oak (right) ................................................................. 90
Figure 77: same as Figure 74 for sites Hainich (Left) and Harvard (right) ................................................................. 91
Figure 78: same as Figure 74 for site Smithsonian ................................................................. 91
Figure 79: same as Figure 74 for sites Moab (Left) and CPER (right) ................................................................. 92
Figure 80: same as Figure 74 for sites Jornada (Left) and Litchfield (right) ................................................................. 92
Figure 81: Time series of GBOV FIPAR LP4 (red/green) and OLCI-A GIFAPAR (grey) for sites Steigerwaldt (left) and Guanica (right) ................................................................. 93
Figure 82: same as Figure 81 for sites Talladega (Left) and Tumbarumba (right) ................................................................. 94
Figure 83: same as Figure 81 for sites Bartlett (Left) and Oak (right) ................................................................. 94
Figure 84: same as Figure 81 for sites Hainich (Left) and Harvard (right) ................................................................. 94
Figure 85: same as Figure 81 for site Smithsonian ................................................................. 95
Figure 86: same as Figure 85 for sites Moab (Left) and CPER (right) ................................................................. 95
Figure 87: same as Figure 81 for sites Jornada (Left) and Litchfield (right) ................................................................. 95
Figure 88: Scatter plots between OLCI-A and OLCI-B of equal dates for GBOV sites (left) and Land Cover types (right) ................................................................. 96
Figure 89: Time series of monthly mean GIFAPAR and OTCI for S3A and S3B for site US-Bartlett. Top-left and right represent S3A; Bottom-left and right represent S3B ................................................................. 97
Figure 90: Time series of monthly mean GIFAPAR and OTCI for S3A and S3B for site CA-Mer-Blue. Top-left and right represent S3A; Bottom-left and right represent S3B ................................................................. 98
Figure 91: Time series of monthly mean OTCI and GIFAPAR for S3A and S3B for site AU-Rushworth. Top-left and right represent S3A; Bottom-left and right represent S3B ................................................................. 99
Figure 92: Time series of monthly mean OTCI and GIFAPAR for S3A and S3B for site AU-Warra Tall. Top-left and right represent S3A; Bottom-left and right represent S3B ................................................................. 100
Figure 93: Baffle temperature trends for SLSTR-A (left) and SLSTR-B (right) from 1st Jan 2022 to end of Jan 2023. The vertical dashed lines indicate the start and end of each cycle ................................................................. 101
Figure 94: OME temperature trends for SLSTR-A (left) and SLSTR-B (right) from Jan 2022 to end of Jan 2023, showing the paraboloid stops and flip baffle (top two plots) and optical bench and scanner and flip assembly (lower two plots). The vertical dashed lines indicate the start and end of each cycle ................................................................. 102
Figure 95: SLSTR-A detector temperatures for each channel from 1st Jan 2022 to end of Jan 2023. The discontinuity for the infrared channels occurs where the FPA was heated for decontamination in July. The vertical dashed lines indicate the start and end of each cycle ................................................................. 104
Figure 96: SLSTR-B detector temperatures for each channel from Jan 2022 to end of Jan 2023. The discontinuity in the IR channels in June 2022 is a result of the anomaly and subsequent decontamination. The vertical dashed lines indicate the start and end of each cycle.

Figure 97: SLSTR-A histogram of standard deviation with respect to the linear control law of the scanners and flip mirror in each orbit for the year 2022 for nadir view (left) and oblique view (right).

Figure 98: SLSTR-A scanner and flip jitter from January 2022 to January 2023, showing mean (red) and stddev (blue) compared to the expected one for the nadir view (left) and oblique view (right). The vertical dashed lines indicate the start and end of each cycle.

Figure 99: SLSTR-B histogram of standard deviation with respect to the linear control law of the scanners and flip mirror for the year 2022.

Figure 100: SLSTR-B scanner and flip jitter, showing mean (red) and stddev (blue) compared to the expected one for the oblique view for March 2022 to February 2023. The vertical dashed lines indicate the start and end of each cycle.

Figure 101: SLSTR-A blackbody temperature and baseplate gradient trends for Jan 2022 to Jan 2023. The vertical dashed lines indicate the start and end of each cycle. Discontinuities are caused by the decontamination in August 2022, and the black-body crossover test at the end of September.

Figure 102: SLSTR-B blackbody temperature and baseplate gradient trends for Jan 2022 to Jan 2023. The vertical dashed lines indicate the start and end of each cycle. Discontinuities are caused by the decontamination of June, and the black-body crossover test of September.

Figure 103: BT differences vs time for all of the blackbody cross-over tests performed to date (including pre-launch measurements) for SLSTR-A (left) and SLSTR-B (right). The part 1 crossover is shown in the top plots, and part 2 in the lower plots. Different symbols indicate different channels (S7: square, S8: triangle, S9: diamond) and different colours indicate nadir (red) and oblique (blue) views. Error bars are derived from the blackbody temperature gradients and standard deviations of the BB signals during the cross-over.

Figure 104: SLSTR-A gain (left) and offset (right) trends for the TIR channels in nadir view. The different colour symbols show the response for each of the detector elements and integrators in the channels. The discontinuity is due to the decontamination (July 2022) and the blackbody crossover test (end Sept 2022). The offset plot for S8 and F2 also shows a step at the end of January 2021 when the commanded detector offset was updated for one detector.

Figure 105: SLSTR-B gain (left) and offset (right) trends for the TIR channels in nadir view. The different colour symbols show the response for each of the detector elements and integrators in the channels. The discontinuity in June is due to the decontamination while the spike in Sept 2022 is due to the blackbody crossover test.

Figure 106: Ratio between nadir and oblique view gains for TIR channels for SLSTR-A (left) and SLSTR-B (right). The different colour symbols show each of the detector elements and integrators in the channels.

Figure 107: Dynamic range lower limit for SLSTR-A (left) and SLSTR-B (right) for channels S8, S9 and F2. The different colour symbols show each of the detector elements and integrators in the channels.
Figure 108: NEDT trend for the thermal channels for SLSTR-A (left) and SLSTR-B (right). Blue points were calculated from the cold blackbody signal and red points from the hot blackbody.

Figure 109: Gain trend for VIS channels (nadir view) for SLSTR-A. The data have been adjusted to allow for the variation of the solar intensity. The oscillations in the signal are due to the build-up of a thin condensation layer causing a thin film interference effect. The different colour symbols show the response for each of the 4 detector elements in the VIS channels.

Figure 110: Gain trend for SWIR channels (nadir view) for SLSTR-A. Outliers in the plots are due to gaps in L0 data or decontamination cycles. The different colour symbols show the response for each of the 8 detector elements of the A and B stripes of the SWIR channels.

Figure 111: Gain trend for VIS channels (nadir view) for SLSTR-B. The data have been adjusted to allow for the variation of the solar intensity. The oscillations in the signal are due to the build-up of a thin condensation layer causing a thin film interference effect. The different colour symbols show the response for each of the 4 detector elements in the VIS channels.

Figure 112: Gain trend for SWIR channels (nadir view) for SLSTR-B. Outliers in the plots are due to gaps in L0 data or decontamination cycles. The different colour symbols show the response for each of the 8 detector elements of the A and B stripes of the SWIR channels.

Figure 113: Dark signal trend for VIS channels (nadir view) for SLSTR-A. The different colour symbols show the signal for each of the 4 detector elements in the VIS channels.

Figure 114: Dark signal trend for SWIR channels (nadir view) for SLSTR-A. The different colour symbols show the signal for each of the 8 detector elements of the A and B stripes of the SWIR channels. The gap/jump in April is due to the decontamination.

Figure 115: Dark signal trend for VIS channels (nadir view) for SLSTR-B. The different colour symbols show the signal for each of the 4 detector elements in the VIS channels.

Figure 116: Dark signal trend for SWIR channels (nadir view) for SLSTR-B. The different colour symbols show the signal for each of the 8 detector elements of the A and B stripes of the SWIR channels.

Figure 117: SLSTR-A VIS and SWIR channel signal-to-noise. Different colours indicate different detectors.

Figure 118: SLSTR-B VIS and SWIR channel signal-to-noise. Different colours indicate different detectors.

Figure 119: Comparisons of SLSTR-A (red) and SLSTR-B (blue) S1-S3 and S5a and S5b channels vs. the corresponding channels for AATSR over desert sites.

Figure 120: Inter-comparisons between SLSTR-A and SLSTR-B S1-S3 and S5, oblique view and AATSR.

Figure 121: Time-series of the elementary ratios (observed/simulated) signal from SLSTR-A for (top to bottom) bands S01, S02 and S03 respectively over January - December 2022 from the six PICS Cal/Val sites. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the desert methodology uncertainty.

Figure 122: Time-series of the elementary ratios (observed/simulated) signal from SLSTR-B for (top to bottom) bands S01, S02 and S03 respectively over January - December 2022 from the six PICS Cal/Val sites.
sites. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the desert methodology uncertainty.--------------------------------------------------136

Figure 123: SLSTR-A measured TOA reflectance with respect to the simulated TOA reflectance over sunglints using DIMITRI for the Nadir view.-----------------------------------------------138

Figure 124: SLSTR-B measured TOA reflectance with respect to the simulated TOA reflectance over sunglints using DIMITRI for the Nadir view.-----------------------------------------------139

Figure 125: The estimated gain values for (top) S3A/SLSTR and (bottom) S3B/SLSTR from Glint, Rayleigh and PICS methods as a function of wavelength. We use the gain value of S2 from PICS method as reference gain for Sunglint method. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the methods uncertainties.-----------------------------------------------140

Figure 126: Summary of comparisons of SLSTR VIS/SWIR channel reflectances vs. Reference methods used to provide vicarious correction factors.------------------------------------------------------143

Figure 127: Daily offset results from the GEOCAL Tool analysis for Nadir view along and across track (top two plots) and Oblique view along and across track (bottom two plots) for SLSTR-A.------------------146

Figure 128: Daily offset results from the GEOCAL Tool analysis for Nadir view along and across track (top two plots) and Oblique view along and across track (bottom two plots) for SLSTR-B.------------------147

Figure 129: In situ validation of S3A SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022. 1st row: Bondville (left), Desert Rock (right); 2nd row: Fort Peck (left), Goodwin Creek (right); 3rd row: Penn State (left), Sioux Falls (right); 4th row: Table Mountain (left); Southern Great Plains (right); 5th row: North Slopes Alaska (left), Kanpur (right).-----------------------------------------------152

Figure 130: In situ validation of S3B SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022. 1st row: Bondville (left), Desert Rock (right); 2nd row: Fort Peck (left), Goodwin Creek (right); 3rd row: Penn State (left), Sioux Falls (right); 4th row: Table Mountain (left); Southern Great Plains (right); 5th row: North Slopes Alaska (left), Kanpur (right).-----------------------------------------------153

Figure 131: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Svartberget between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively.-----------------------------------------------157

Figure 132: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Svartberget between October 2021 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B.-----------------------------------------------157

Figure 133: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Svartberget between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup pair added together in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot. --158
Figure 134: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Hyytiälä between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively.--------------------------------- 159

Figure 135: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Hyytiälä between October 2021 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B. --------------------------------- 159

Figure 136: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Hyytiälä between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot. --------------------- 160

Figure 137: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from the KIT forest site between August 2020 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. --------------------- 161

Figure 138: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from the KIT forest site between August 2020 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B. --------------------------------- 161

Figure 139: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from the KIT forest site as in Figure 138, but using a 3 x 3 grid to calculate $\sigma_{\text{space}}$ in Equation 1. --------------------------------- 162

Figure 140: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over the KIT forest site between August 2020 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot. --------------------- 162

Figure 141: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Robson Creek between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. --------------------------------------------------------------- 164

Figure 142: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Robson Creek between October 2021 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B. --------------------------------------------------------------- 164
Figure 143: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Robson Creek between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot.  

Figure 144: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Puéchabon between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively.  

Figure 145: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Puéchabon between October 2021 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B.  

Figure 146: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Puéchabon between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot.  

Figure 147: Areas of high fire activity selected for the inter-comparison. The basemap shows night-time fires detected by SLSTR (top) and MODIS (bottom) for the months of respectively (1) April, May – (2) July, August, September and (3) October, November and December 2022.  

Figure 148: The basemap shows night-time fires detected by Sentinel 3A (top) and Sentinel 3B (bottom) for the months of October and December 2022.  

Figure 149: Cluster formation for a specific area.  

Figure 150: Scatterplot between the number of fires detected by both sensors (night-time). SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. (1) April-May, (2) July, August and September, (3) October, November and December. Colour of circles represents the SLSTR fire radiative power while diameter represents the number of occurrences.  

Figure 151: Distribution of the FRP for fire clusters detected by both sensors. The lines represent kernel density estimates (KDE) computed over the barplot.  

Figure 152: Scatterplot between the FRP of fire clusters detected by both sensors. SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. The regression line (with a shaded 0.05 confidence interval) is obtained with a robust linear method (RLM), which is less affected by outliers. The scatter points are color-coded according to their cluster number, so that they can be traced back to the original datasets. (1) April, May R=0. 785; (2) July, August and September R=0.835 and (3) October, November and December R=0.925.  

Figure 153: Areas of high fire activity selected for the inter-comparison. The basemap shows day-time fires detected by SLSTR (top) and MODIS (bottom) for the months of respectively (1) April, May – (2) July, August, September and (3) October, November and December 2022.
Figure 154: Scatterplot between the number of fires clusters detected by both sensors (day-time). SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. (1) April-May, (2) July, August and September, (3) October, November and December.

Figure 155: Scatterplot between the FRP of fire clusters detected by both sensors (with an FRP < 500MW). SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. (1) April, May R=0.915; (2) July, August and September R=0.925 and (3) October, November and December R = 0.775.

Figure 156: SLSTR vs MODIS clusters (FRP < 200) comparisons with regression line – night-time (left) and day-time (right). (1) April-May, (2) July, August and September, (3) October, November and December.

Figure 157: Selected regions across the globe for inter-comparison of SYN SDR products.

Figure 158: density scatter plots of inter-comparisons of SYN SDR, Oa17 with that of observed from MODIS, b02, for the regions (a) Congo, (b) Europe, (c) Africa, (d) Australia, (e) Central USA and (f) region of France and for randomly selected days in the year 2022.

Figure 159: density scatter plots of inter-comparisons between SYN SDR (OLCI), bands Oa08, Oa17, Oa04, Oa06 and MODIS, b01, b02, b03, b04 for the region of Australia and for a selected day during the year 2022.

Figure 160: density scatter plots of inter-comparisons of SYN SDR (OLCI), Oa17 with that of observed from VIIRS, bands M7, for the regions (b) Africa, (d) Australia. Shown in the same figure, the counter parts from MODIS, b02 for the same regions, (a) Africa, (c) Australia and for randomly selected days during the year 2022.

Figure 161: (top 4 plots) time series of surface reflectance (SR) of MODIS b01, b02, b03 and b04 and their Sentinel3-A SYN SDR (OLCI) counterpart Oa08, Oa17, Oa04 and Oa06; and (bottom 4 plots) time series of the comparison statistics in-terms of A,P,U for the selected bands, for the station Talladega, United States (needle-leaved evergreen).

Figure 162: same as Figure 161 but for the station Robson-Creek, Australia (Broadleaved Evergreen).

Figure 163: same as Figure 161 but for the station Estrees-Mons, France (Cultivated).

Figure 164: same as Figure 161 but for the station Lison, Italy (Cropland).

Figure 165: same as Figure 161 but for the station MOAB, United States (Shrub closed open deciduous).

Figure 166: same as Figure 161 but for the station Dahra, Senegal (Cultivated).

Figure 167: same as Figure 161 but for the station Mata-Seca, Brazil (Herbaceous).

Figure 168: Scatter plot of intercomparisons between Sentinel3-A SYN SDR (OLCI) and MODIS for selected band pairs: b01 vs Oa08 (a), b02 vs Oa17 (b), b03 vs Oa04 (c) and b04 vs Oa06 (d).

Figure 169: Scatter plot of intercomparisons between Sentinel3-B SYN SDR (OLCI) and MODIS for selected band pairs: b01 vs Oa08 (a), b02 vs Oa17 (b), b03 vs Oa04 (c) and b04 vs Oa06 (d).
Figure 170: Scatter plot of intercomparisons between Sentinel3-B SYN SDR (OLCI) and MODIS for the selected band pair b01 vs Oa08 and for selected biomes: (a) Broad evergreen (b), Needle evergreen (c) Shrub and (d) Cultivated.

Figure 171: Scatter plot of intercomparisons between Sentinel3-A SYN SDR (OLCI) and MODIS for the selected band pair b01 vs Oa08 and for selected biomes: (a) Broad evergreen (b), Needle evergreen (c) Shrub and (d) Cultivated.

Figure 172: Number of pairwise TOA observations used in the intercomparison between SY_VGP and PROBA-V L2A from 01/07/2020 – 30/09/2020.

Figure 173: Distribution of Accuracy (A), Precision (P) and Uncertainty (U) between SY_VGP and PROBA-V L2A match-ups (July-September 2020) for S3A (blue) and S3B (orange).

Figure 174: Spatial artefacts related to projection to the fixed grid and large ground sampling distances at the edge of the swath. S3A_SY_VG1 product of 20220726, focus on an area around 58°N 45°E.

Figure 175: Blue TOC (left) and AOT (right, missing pixels in dark grey) for S3A_SY_VG1 product of 20220405. Focus on an area around 24°N 2°E. High AOT values (right) introduce artefacts in the Blue TOC (left) over a relatively homogeneous area.

Figure 176: GMR scatter density plots, frequency histograms and bias frequency plots between S3A (top) and S3B (bottom) SY_V10 (2022) and SPOT-VGT S10-TOC LTS (2009-2013), for Blue, Red, NIR and SWIR bands.

Figure 177: GMR scatter density plots, frequency histograms and bias frequency plots between S3A (top) and S3B (bottom) SY_V10 (2022) and PROBA-V S10-TOC LTS (2014-2018), for Blue, Red, NIR and SWIR bands.

Figure 178: GMR scatter density plot (left), frequency histogram (middle) and bias frequency plot (right) between S3A resp. S3B SY_V10 (2022) and SPOT/VGT S10 TOC NDVI LTS (2009-2013) (top) and PROBA-V S10-TOC NDVI LTS (2014-2018) (bottom).

Figure 179: Land and ocean regions defined for this study (as in Sogacheva et al., 2020): Europe (Eur), Boreal (Bor), northern Asia (AsN), eastern Asia (AsE), western Asia (AsW), Australia (Aus), northern Africa (AFN), southern Africa (AFS), South America (SA), eastern North America (NAE), western North America (NAW), Indonesia (Ind), Atlantic Ocean dust outbreak (AOD), Atlantic Ocean biomass burning outbreak (AOB). In addition, Southeast China (ChinaSE), which is part of the AsE region, marked with a blue frame, is considered separately. Land, ocean and global AOD were also considered.

Figure 180: Globally, for S3A (left) and S3B (right), scatter density plots for AOD$_{550}$ matchups between Sy$_2$ and AERONET for different groups of products: all matchups, pixels retrieved with dual, single applied to nadir (singleN) and single applied to oblique (singleO) approaches (top down).

Figure 181: Left: For S3A, binned (based on AERONET AOD) Sy$_2$ AOD offsets (magenta dots in Figure 180) for different groups of matchups: all, dual, singleN and singleO. Right: For S3A and S3B, for different groups of matchups, binned offset to AERONET (colours for S3A are as on the left panel; S3B is shown in grey).

Figure 182: Binned syAOD offset for selected regions (see legend) for dual (upper panel) and singleN (lower panel) groups of matchups.
Figure 183: Global difference (dAOD$_{550}$) between syAOD and aAOD for aAOD binned in 0.1 intervals: median bias (circles) and bias standard deviation (error bars) for all and background (aAOD ≤ 0.2) AOD types (purple), aerosol fine-dominated AOD (blue) and coarse-dominated AOD (green). The fraction (F) of points in each bin from the total number of matchups is represented with orange bars. The fraction of fine-dominated matchups in each bin is shown as the blue dashed-line.---------------------------------226

Figure 184: For S3A, seasonal scatter density plots with corresponding validation statistic.-----------------------------227

Figure 185: For S3A (left panel) and S3B (right panel), annual scatter density plots with corresponding validation statistic for three consequent years, 2020, 2021 and 2022.-----------------------------------228

Figure 186: For S3A (left) and S3B (right), annual binned syAOD offsets for years 2020, 2021 and 2022 -------------------------------------------------------------229

Figure 187: For S3A, scatter density plot for syAOD$_{550}$ “extreme” (>1.0AOD) outliers July, years 2020, 2021 and 2022 (upper panel) and for AERONET stations fraction of outliers for corresponding periods (lower panel)--------------------------------------------------229

Figure 188: Examples of syAOD spatial distribution for the cases when matchup with AERONET station (pixels within the magenta circles) is classified as an “extreme” syAOD outlier.---------------------------------------------230

Figure 189: $\chi^2$ for binned aAOD for all available matchups (magenta line) and after the outliers of the individual weighted deviations ($d_i$>10) are removed (red line). Density scatter plot for PU and syAOD.---------------------------------------------------------------------------------------------------------------231

Figure 190: Comparison between prognostic uncertainties, AOD error distribution and theoretical Gaussian distribution for the whole product (left panel), dual- (middle panel) and singleN (right panel) groups of matchups.-----------------------------------------232

Figure 191: Comparison between normalized error and theoretical Gaussian distribution for the whole product (left panel), dual- (middle panel) and singleN (right panel) groups of matchups.-----------------------------------------232

Figure 192: Percentile plots of absAODerrors at 38% (black), 68% (red) and 95% (blue) as function of binned expected discrepancy.-------------------------------------------------------------------------------------------------------------------------------------233

Figure 193: For S3A (left panel) and S3B (right panel), global scatter plots between syAE$_{550-865}$ and aAE$_{500-870}$ for different groups of products (top-down: all, dual, singleN and singleO).---------------------------------------------234

Figure 194: For S3A, regional scatter plots between syAE$_{550-865}$ and aAE$_{500-870}$-------------------------------------------------------------235

Figure 195: Scatter density plots for S3A syFMAOD and corresponding aFMAOD for matchups available globally.-------------------------------------------------------------------------------------------------------------------------------------236

Figure 196: OLCI anomalies/events in 2021----------------------------------------------------------------------------------240

Figure 197: SLSTR anomalies/events in 2021----------------------------------------------------------------------------------241

Figure 198: SYN anomalies/events in 2021-----------------------------------------------------------------------------------242
List of Tables

Table 1: OLCI Processing Baseline  ......................................................................................................................................................... 8
Table 2: SLSTR Processing Baseline .................................................................................................................................................... 9
Table 3: SYN Processing Baseline ....................................................................................................................................................... 10

Table 4: OLCI-A SNR figures as derived from Radiometric Calibration data. Figures are given for each camera (time average and standard deviation), and for the whole instrument. The requirement and its reference radiance level are recalled (in mW.sr⁻¹.m⁻².nm⁻¹). ........................................................................................................ 18

Table 5: OLCI-B SNR figures as derived from Radiometric Calibration data. Figures are given for each camera (time average and standard deviation), and for the whole instrument. The requirement and its reference radiance level are recalled (in mW.sr⁻¹.m⁻².nm⁻¹). ........................................................................................................ 20

Table 6: Synthesis of the DIMITRI results: estimated gain values for S3A/OLCI from Glint, Rayleigh and PICS over the period January 2022- December 2022. .................................................................................................................. 52

Table 7: Synthesis of the DIMITRI results: estimated gain values for S3B/OLCI from Glint, Rayleigh and PICS over the period January 2022- December 2022. .................................................................................................................. 53

Table 8: S3ETRAC Rayleigh Calibration sites ....................................................................................................................................... 55

Table 9. OSCAR Rayleigh and Glitter calibration results for S3A and S3B for 2022 ............................................................................. 60

Table 10: Validation sites analysed in report S3A 78/S3B 59. Land cover data from GLC2000 grouped according to the International Geosphere-Biosphere Programme (IGBP) designations. .................................................................................. 82

Table 11: Comparison statistics between monthly S3A/B OLCI land products and MERIS archive data. - 83

Table 12: GBOV validation sites analysed. ................................................................................................................................................. 88

Table 13: NEDT for SLSTR-A in cycles 067-078 averaged over all detectors for both Earth views towards the hot +YBB (top) and the cold -YBB (bottom). .................................................................................................................. 114

Table 14: NEDT for SLSTR-B in cycles 048-059 averaged over all detectors for both Earth views towards the hot +YBB (top) and the cold -YBB (bottom) ........................................................................................................ 115

Table 15: The list of these sites and their geographical limits ........................................................................................................... 130

Table 16: Summary of Vicarious Radiometric Validation Results performed by ARGANS using DIMITRI. Results presented here are the ratios R_{meas}/R_{ref} averaged over the period January-December 2022.---- 141

Table 17: Summary of Vicarious Radiometric Calibration Results performed by all groups. Comparisons are performed by comparing the measured reflectance vs. reference reflectance. Results presented here are the ratios R_{meas}/R_{ref}. .................................................................................................................. 142

Table 18: Proposed VIS-SWIR Calibration Adjustments Based on Vicarious Calibration analysis. Note S4 is not included because the vicarious calibration techniques do not extend to this band. ................................. 144

Table 19: SLSTR-A temporal drift rates for each year since the start of the mission (in %) ............................ 144
Table 20: SLSTR-B temporal drift rates for each year since the start of the mission (in %) 145

Table 21: Statistics of In situ validation for S3A SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022 154

Table 22: Statistics of In situ validation for SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022 154

Table 23: The results of the validation of Sentinel-3A and 3B LST using in-situ measurements. 156

Table 24: Summary of the inter-comparison between night-time SLSTR FRP and MODIS FRP – October, November and December 2022. Results from previous comparison obtained between July and August 2022, April and May 2022 also obtained with FRP have been added respectively in the second and third column for information and comparison. 170

Table 25: Summary of the inter-comparison between SLSTR FRP and MODIS FRP – From October to December 2022 data. 182

Table 26: List of CEOS LPVs used for intercomparison of SYN SDR products. 198

Table 27 APU statistics for S3A and S3B SY_V10 products vs. SPOT/VGT LTS and PROBA-V LTS 218

Table 28: List of OLCI Product Notices issued in 2022 238

Table 29: List of SLSTR Product Notices issued in 2022 238

Table 30: List of SYN Product Notices issued in 2022 239
1 Introduction

This document is the 2022 Annual Performance Report of the Sentinel-3 optical mission prepared by the OPT-MPC consortium led by ACRI-ST in the frame of the “COPERNICUS SPACE COMPONENT SENTINEL OPTICAL IMAGING MISSION PERFORMANCE CLUSTER SERVICE”, ESA contract 4000136252/21/I-BG.

1.1 Scope of the document

This document provides a summary of the end-to-end mission performance of Sentinel-3 from the 1st of January 2022 until the 31st of December 2022 carried out by the Optical Mission Performance Cluster during its first year of the routine operations phase.

Please note that a report addresses the Sentinel-2 mission (ref. OMPC.CS.APR.001).

1.2 Applicable and Reference documents

The full Applicable Documents (AD) and Reference Documents (RD) ID correspondence is provided in the Configuration Item Data List (OMPC.ACR.LST.001).

1.3 Acronyms and abbreviations

The definition of the acronyms and abbreviations used in this document is provided in the List of Acronyms and Definitions (OMPC.ACR.LST.002).
2 Executive Summary

2.1 OLCI

Instrument performance

The OLCI-A and OLCI-B instrument health is excellent. The sensors temperatures are perfectly well controlled. The nominal radiometric diffusers ageing shows the expected magnitude and spectral behaviours: around 0.5% after 5 years for OLCI-A at 400 nm (Oa01), down to 0.1% at 560 nm (Oa06) and undetectable above; below 0.35% for OLCI-B. The instrument sensitivity evolution so far is limited to about 3% (OLCI-A) and no evidence of severe degradation can be demonstrated: the variation of the instrument sensitivity seems more correlated with a potential spectral evolution of the correcting filters – inside the spectrometers – than to darkening of the optics or loss of sensitivity of the CCD sensors. Sensitivity evolution of OLCI-B is similar to that of OLCI-A, and maybe with a slightly higher magnitude for the 400 nm channel (up to 4.5%). The regularly monitored instrument SNR performance is well within requirement and shows a very good stability in time.

Spectral Calibration is monitored thanks to dedicated acquisition campaigns. The in-flight spectral campaigns reveal a high agreement of the in-flight characterisation with the pre-flight spectral calibration for both A and B sensors, with differences of the OLCI channels centre smaller than 0.1 nm, except for channels Oa01 (400 nm) and Oa21 (1020 nm), with up to 0.2 nm. A small temporal evolution is observed, different for each camera but approximately identical at all wavelengths; the observed changes for OLCI-A after 4 years are smaller than 0.15 nm (except camera 5 at 0.23 nm); observed changes for OLCI-B are within 0.25 nm for all cameras.

Level 1 products performance

The geometric performance is monitored using the ESA GeoCal tool CFI. It is currently fully compliant for OLCI-A and OLCI-B to the 0.5 pixel RMS requirement. The significant along-track drift of OLCI-B cameras assessed over the first 12 to 16 months of mission, that required frequent geometric re-calibration, seems now stabilized.

The OLCI-A and OLCI-B Radiometric Gain Models (gain at reference date + time drift) are used to calibrate Earth Observation data at any date. Their current performance is better than 0.1% RMS (0.12 for OLCI-B channel Oa01).

Absolute and inter-band calibration performance is monitored by indirect methods over natural targets. Three methods are used within S3-MPC: the “Rayleigh” method (molecular atmospheric backscattering over clear sky off-glint open ocean) provides absolute calibration in the blue-to-red spectral domain; the “Glint” method (spectral dependency of the Sun specular reflection over ocean) provides inter-band calibration; and the PICS method (Pseudo-Invariant Calibration Sites, temporarily stable desert areas) provides absolute calibration over the whole spectral domain as well as cross-mission comparisons for sensors with comparable channels. Two of these methods, Rayleigh and Glint, are undertaken by two different implementations providing very consistent results.

All methods point out an excess of brightness for OLCI-A radiances. Results are in pretty close agreement around 2-3% between 560 and 900 nm (Oa06 to Oa19). Rayleigh gives higher biases in the blue-green (about 6 % while PICS remains around 2%) but this method is suspected to overestimate the simulated
signal at those wavelengths so PICS are considered more reliable. Channel Oa21 (1020 nm) is only addressed by the Glint interband method, and the results are much worse: 4 to 8%, depending on the reference band. **Radiometric validation for OLCI-B indicates performance within the 2% requirement** for all bands from 560 nm (Oa05) to 940 nm (Oa20). As for OLCI-A, the PICS method shows compliance also in the blue region (Oa1 to Oa4, 400 to 510 nm) while the Rayleigh method shows biases of about 3 to 5%, depending on implementation. The OLCI-B 1020 nm (Oa21) has a similar performance than its OLCI-A counterpart.

**Level 2 products performance**

**Integrated Water Vapour**

Integrated Water Vapour has been validated against available in-situ data, according to the surface type: GNSS and AERONET networks over Land, AERONET (coastal stations), AERONET-OC and AERONET Maritime networks over water.

Validation over Land demonstrates that the product is of high quality (bias corrected RMS difference of ~ 0.8 to 1.5 kg/m²) for retrievals above land surfaces, but there is a systematic overestimation of 9% to 13%. Validation for OLCI-B gives similar results.

The comparison with GNSS stations close to water shows a larger wet bias for the ocean retrievals (up to 25%), and in particular in transition zones between glint and off glint.

**Land Products**

**Green Instantaneous Fraction of Absorbed Photosynthetically Available Radiation (GIFAPAR, formerly referred to as the OLCI Global Vegetation Index, O-GVI)**

Validation against in-situ data is now possible thanks to the Ground-Based Observations for Validation (GOBV) of Copernicus Global Land Products project. It provides in particular regular FIPAR measurements at a number at stations (see GBOV-ATBD-RM4-RM6-RM7_v2.1-Vegetation for the definition of the FIPAR, Fraction of Intercepted PAR and some considerations on differences with GIFAPAR, in particular the contribution of non-photosynthetic vegetation (NPV) elements, such as stems and branches, not accounted for by GIFAPAR).

OLCI reproduces the temporal variations of GBOV values satisfactorily for almost all sites. However, GBOV provide systematically higher values than satellite products in forest classes. On the contrary, for shrubland and grasslands, satellite products tend to present higher values than GBOV. OLCI-A and OLCI-B provide very similar comparison patterns. Overall, OLCI GIFAPAR compares with high correlation, > 0.85 (when sufficient dynamics are present) and reasonably good RMSD (~0.2) considering the inherent differences between FIPAR and GIFAPAR.

In addition, OLCI GIFAPAR is regularly compared to MERIS 10-years climatology. There is a fairly good agreement, accounting for the methodology limitations, with high correlation, > 0.9 (when sufficient dynamics are present) and good RMSD (<0.1). Consistency between OLCI-A and OLCI-B is very good, with high correlation (0.9), low bias (0.01) and dispersion (NRMSE=0.18).
OLCI Terrestrial Chlorophyll Index (O-TCI)

Quantitative validation against in-situ data is not possible so far, as no in-situ station provides directly comparable products. **Comparison with MERIS TCI (M-TCI) climatology** is regularly done over a number of sites, showing high correlation, > 0.9 (when sufficient dynamics are present) and good RMDS (<0.1). **Consistency between OLCI-A and OLCI-B is very good**, with high correlation (0.95), low bias (0.03) and dispersion (NRMSE=0.07).

### 2.2 SLSTR

#### 2.2.1 SLSTR-A

**Instrument performance**

The SLSTR-A instrument has performed exceptionally well for another year, with all parameters within safe limits. There have been no major anomalies, and only short gaps in data coverage due to ground station issues, manoeuvres or calibration observations.

The cooler has been performing well, with the IR detectors maintained at a stable temperature.

Radiometric noise levels for the TIR and VIS/SWIR channels have remained stable throughout at pre-launch values. NEDT for the S8 and S9 channels are below 20 mK with no indication of degradation.

Blackbody temperatures have shown a seasonal cycle on top of the daily/orbital temperature cycles, with the highest temperatures reached during December. The commanded heater power to the blackbody was reduced by one bit at the end of September 2021 in order to reduce the maximum temperature.

The VISCAL system is illuminated by the Sun once per orbit and Vicarious calibration results suggest that the system is not degrading significantly over time. The stability is much better than that observed for AATSR on ENVISAT.

The scanners continue to perform well, with orbital mean deviation from the expected position for both nadir and oblique scanners less than 1.5”, and a standard deviation less than 4”. The flip mirror orbital mean deviation is less than 1” with a standard deviation <6” in the nadir position and <15” in oblique position. The worst instantaneous jitter encountered is as good, or better, than previous years.

**Level 1 products performance**

Validation of the absolute radiometric calibration of the IR channels has been carried out at EUMETSAT using comparisons against IASI-A and B in 2018. The stability of the flight gains, radiometric noise and instrument temperatures suggest that the calibration has not drifted significantly since then.

The VIS and SWIR channels are calibrated via an on-board Solar diffuser-based calibration system. Evaluation of the radiometric calibration has used the techniques developed for AATSR and MERIS and show that the calibration system is stable. Assessment of the VIS channels S1-S3 show good agreement with OLCI and AATSR. At the SWIR wavelengths, there is a significant discrepancy between SLSTR and AATSR and MODIS that must be taken into account in any L2 processing. An adjustment to the L1 processing to correct the main calibration difference is foreseen. The root cause of the anomaly has not been found and is still under investigation.
Geometric calibration is monitored using the GEOCAL tool. Average absolute geometric offsets <0.1 km are achieved for the nadir view and oblique view across-track and <0.2 km for the oblique view along-track.

**Level 2 products performance**

**Land Products**

The SLSTR-A SL_2_LST product from SLSTR went operational in the Sentinel 3 PDGS on 5th July 2017 with PB 2.16. No additional updates to the retrieval algorithm have been implemented in the IPF since. However, Processing Baseline 2.29 released on 4th April 2018 included the new Probabilistic Cloud Mask implemented in the IPF at Level-1 and carried through to Level-2. Furthermore, from 26th February 2019 an updated ADF of retrieval coefficients has been implemented in PB 2.47, IPF 06.14. An updated probabilistic cloud coefficients ADF was applied on 23rd October 2020 in PB 2.73. Matchups against twelve "Gold Standard" in situ stations show that the overall absolute daytime accuracy is 0.86 K and the absolute night-time accuracy is 0.66 K, both of which are within the mission requirements for LST. Comparisons with respect to the operational LSA SAF LST product are within the uncertainty range when considering the uncertainties from the reference products, and thus the products can be interpreted as consistent with each other. Overall, the SL_2_LST product is performing in line with the 1 K mission requirement for LST.

A major upgrade of the FRP algorithm was deployed in March 2022. This updated algorithm is based on the previous one, with similar night-time algorithm, but includes improved thermal fire detection over daytime products. In addition, fire detection is now also performed using SWIR channels.

The FRP product quality is assessed via inter-comparisons with MODIS products. The inter-comparison highlights a generally good consistency. MODIS detects slightly fewer fires in the Northern hemisphere at night-time, whereas results are closer in the Southern hemisphere. For day-time fires, MODIS detects clearly more fires in the Northern hemisphere. More investigations are needed to better assess those differences.

**2.2.2 SLSTR-B**

**Instrument performance**

Instrument and blackbody temperatures for SLSTR-B have been stable on top of the daily/orbital and seasonal trends, and consistent with those for SLSTR-A. The cooler has been performing well, with the IR detectors maintained at a stable temperature.

The visible channel radiometric gain shows a variation from orbit to orbit especially in channels S1 and S2. The reason for this behaviour is thought to be due to partial motional chopping of the VIS detectors by an internal aperture in the VIS FPA. If this is correct, the effect will be present on the earth scene data for S1 and S2.

The NEDT levels are roughly consistent between SLSTR-A and SLSTR-B, except for F1, which shows more orbit-to-orbit variation and higher noise values. This may be caused by motional chopping of the SLSTR-B F1 detectors, which are known to be close to edge of the aperture for SLSTR-B.

The SLSTR-B scanner and flip mean and standard deviations from their expected positions are broadly consistent with SLSTR-A, although the oblique scanner has a slightly larger mean deviation of <3".
**Level 1 products performance**

Initial validation of the absolute radiometric calibration of the IR channels has been carried out by EUMETSAT using comparisons against IASI-A and B. Analysis from the tandem phase comparisons show that the in-flight calibration of SLSTR-B is consistent with that of SLSTR-A.

The S3A and S3B satellites are configured to be 140 degrees out of phase in order to observe complimentary portions of the earth. Figure 1 shows an example combined Level-1 SLSTR-A/SLSTR-B image (daytime only) to show the combined daily SLSTR coverage during one day.

![Daytime combined SLSTR-A and SLSTR-B Level-1 image](image)

*Figure 1: Daytime combined SLSTR-A and SLSTR-B Level-1 image for SWIR channels on 15th November 2022.*

**Level 2 products performance**

**Land Products**

The S3B SL_2_LST product from SLSTR went operational in the Sentinel 3 PDGS on 26th February 2019 with PB 1.19 IPF 06.14. An updated probabilistic cloud coefficients ADF was applied on 23rd October 2020 in PB 1.50. Matchups against ten "Gold Standard" in situ stations show that the overall absolute daytime accuracy is 0.81 K and the absolute night-time accuracy is 0.60 K, both of which are within the mission requirements for LST. As for SLSTR-A, comparisons with respect to the operational LSA SAF LST product are within the uncertainty range when considering the uncertainties from the reference products, and thus the products can be interpreted as consistent with each other. Overall, the SL_2_LST product is performing in line with the 1 K mission requirement for LST.

The FRP algorithm was upgraded in March 2022. The performance of FRP products for SLSTR-B, as assessed through intercomparisons with MODIS, is in line with that of SLSTR-A products.
2.3 SYN

Concerning SYN L2 SDR products, SYN L2 Surface directional reflectance and normalized SDR MODIS and VIIRS provides a global satisfying correlation between both datasets, with low dispersion except over desert.

Similarly, we observe a high correspondence between TOA reflectances provided by SYN VGT-like products and the ones provided by PROBA-V products. A regression slope close to 1 is observed on the BLUE, Red and NIR radiometric measurements. For SWIR measurements however, systematic large differences are observed and could be linked to the SLSTR calibration of SWIR channels.

A high consistency between SYN VGT-like and PROBA-V TOC reflectances can be observed with very small differences directly linked to calibration issues. Similarly, when considering the TOC NDVI dataset, both products show strong agreement ($R^2 = 0.975$).

Validation of the SYNERGY L2 global aerosol product against AERONET data shows a global performance of reasonable quality with an RMSE on AOD of 0.25 and a correlation coefficient $R$ of 0.6, with slightly better results for S3B. However, this hides variable performances performance depending on underlying surface – as expected – but also on the retrieval approach (single or dual view). If the correlation coefficient can reach 0.9 with RMSE of 0.1 for single view oblique retrieval, it degrades to 0.6 and 0.3 for other retrieval conditions. The conclusions of the detail analysis that AOD over land is below GCOS requirements and that Cloud Screening should be considerably revised.
3 Processing baseline description

This section lists all processing baselines that have been delivered between the 1st of January 2022 and the 31st of December 2022, corresponding to year 1 of the OPT-MPC service contract.

3.1 OLCI

All OLCI processing baselines are listed in Table 1.

<table>
<thead>
<tr>
<th>Processing Baseline</th>
<th>Delivered to PDGS</th>
<th>Deployed in Land PDGS</th>
<th>Changes</th>
</tr>
</thead>
</table>
| OL1 IPF v6.12       | 26/11/2021        | 20/01/2022            | ▪ Inclusion of updated processing baseline number in manifest and products  
▪ Update of OLCI-A and OLCI Radiometric Gain Models |
| OL2_LND IPF v6.16   | 13/12/2021        | 26/01/2022            | ▪ Inclusion of updated processing baseline number in manifest and products  
▪ Inclusion of L1 Degradation Flags in L2 associated products |
| OL1 IPF v6.13       | 03/06/2022        | 19/07/2022            | ▪ Inclusion of radiometric uncertainty computation inside OLCI L1 IPF (however, this computation is still disabled)  
▪ Update of OLCI-A and OLCI Radiometric Gain Models |
| OL1 IPF v6.13       | 07/07/2022        | 22/08/2022 (S3A) 30/08/2022 (S3B) | ▪ Activation of the Uncertainties Computation in OLCI L1 IPF |
| OL2_LND IPF v6.16   | 07/07/2022        | 22/08/2022 (S3A) 05/09/2022 (S3B) | ▪ Configuration of processing site in S3 L2 product manifest |
| OL1 IPF v6.14       | 14/11/2022        | Not yet deployed      | ▪ Add the support to S3C/D to L1 processing.  
▪ Detection/flagging of OLCI L1b Pixels affected by partial saturation of microbands associated to a nominal OLCI channel.  
▪ OLCI L1b Geometric Correction change and Frame Offset Orbit Evolution  
▪ Update of OLCI-A and OLCI Radiometric Gain Models  
▪ Correction of a Bug in SRS with usage of E0 not seasonally corrected |

Note that this version has been correcting by a patch delivered on 05/01/2023
3.2 SLSTR

All SLSTR processing baselines are listed in Table 2.

<table>
<thead>
<tr>
<th>Processing Baseline</th>
<th>Delivered to PDGS</th>
<th>Deployed in Land PDGS</th>
<th>Changes</th>
</tr>
</thead>
</table>
| SL1_IPF v6.19       | 15/10/2021        | 09/02/2022            | ▪ Inclusion of updated processing baseline number in manifest and products  
▪ Correction of the SLSTR Calibration Issue  
▪ Correction of several Met field issues  
▪ Correction of the False flagging S8 of Saturation in S8 (S3A only)  
▪ Inclusion of a switch to disable Bayesian and Probabilistic cloud module inside SLSTR L1 processing |
| SL2_LND IFP v6.20   | 25/01/2022        | 09/02/2022            | ▪ Inclusion of updated processing baseline number in manifest and products  
▪ Inclusion of the SLSTR Probabilistic Cloud module with RTTOV v12.3  
▪ Replacement of SLSTR L2 LST summary cloud by probabilistic cloud mask  
▪ Inclusion of L1 Degradation Flags in L2 associated products  
▪ Correction of L1 Degradation Flags in L2 associated products  
▪ Correction of block processing error occurring on SWIR detection |
| SL2_FRP_LND IPF v01.07 | 25/01/2022        | 28/02/2022            | ▪ Inclusion of updated processing baseline number in manifest and products  
▪ Inclusion of L1 Degradation Flags in L2 associated products  
▪ Inclusion of FRP V2 algorithm (with specific daytime processing over saturated scenes + updated SWIR detection)  
▪ Correction of Sunglint test wrt to S2 saturated pixels  
▪ Inclusion of a clustering approach between SWIR and TIR fires  
▪ Correction of block processing error occurring on SWIR detection |
### Processing Baseline

<table>
<thead>
<tr>
<th>Processing Baseline</th>
<th>Delivered to PDGS</th>
<th>Deployed in Land PDGS</th>
<th>Changes</th>
</tr>
</thead>
</table>
| SL2_LND IPF v6.21 SL_LST_004.007 Global PB_ID 3.11 | 07/07/2022 | 23/08/2022 (S3A) 05/09/2022 (S3B) | • Correction of Blocks processing issues impacting tie point interpolation and cloud module processing  
• Configuration of processing site in S3 L2 product manifest |
| SL2.FRPLND IPF v01.08 FRP_NTCL004.08.00 Global PB_ID 3.11 | 07/07/2022 | 23/08/2022 (S3A) 05/09/2022 (S3B) | • Correction of a (non-impacting) processing issue when using a deficient product to get first margin.  
• Configuration of processing site in S3 L2 product manifest |
| SL1 IPF v6.20 SL_L1__004.05.00 Global PB_ID 3.14 | 20/10/2022 | 12/12/2022 (S3A) 10/01/2023 (S3B) | • Support to S3C/D to L1 processing  
• Handling of Ungracious (seg. fault or hang-over) failures during Moon Calibration  
• Correction of minor error generating S3A and S3B SL_1_VSC_AX during soler eclipse  
• Correction of the wrong dimension for bayes_orphan dataset |
| SL2_LND IPF v06.21 SL_LST_004.007.01 Global PB_ID 3.17 | 09/12/2022 | Not yet Deployed | • Correction of the wrong dimension for bayes_orphan dataset |
| SL2.FRPLND IPF v01.08 FRP_NTCL004.08.01 Global PB_ID 3.17 | 09/12/2022 | Not deployed | • Correction of the wrong dimension for bayes_orphan dataset |

### 3.3 SYN

All SYN processing baselines are listed in Table 3.

**Table 3: SYN Processing Baseline**

<table>
<thead>
<tr>
<th>Processing Baseline</th>
<th>Delivered to PDGS</th>
<th>Deployed in Land PDGS</th>
<th>Changes</th>
</tr>
</thead>
</table>
| SY2 IPF v06.22 SYN_L2__.002.15.00 Global PB_ID 3.03 | 13/12/2021 | 27/01/2022 | • Inclusion of updated processing baseline number in manifest and products  
• Inclusion of L1 Degradation Flags in L2 associated products.  
• Update IPF facility file to include the new production services |
| SY2_VGS IPF v06.10 SYN_L2V.002.07.00 Global PB_ID 3.03 | 13/12/2021 | 27/01/2022 | • Inclusion of updated processing baseline number in manifest and products  
• Inclusion of L1 Degradation Flags in L2 associated products.  
• Update IPF facility file to include the new production services |
| SY2_AOD IPF v01.06 AOD_NTC.0002.06.00 Global PB_ID 3.03 | 13/12/2021 | 27/01/2022 | • Inclusion of updated processing baseline number in manifest and products  
• Inclusion of L1 Degradation Flags in L2 associated products.  
• Update IPF facility file to include the new production services |
### Processing Baseline

<table>
<thead>
<tr>
<th>Processing Baseline</th>
<th>Delivered to PDGS</th>
<th>Deployed in Land PDGS</th>
<th>Changes</th>
</tr>
</thead>
</table>
| SY2 IPF v06.23      | 14/07/2022        | 23/08/2022 (S3A)     | ▪ Correction of the Empty crs dataset in all VGT-like products  
▪ Correction of the Data gap observed on high latitude on the eastern part of OLCI orbits.  
▪ Correction of the VGT status map when VGT radiometry is set to NaN  
▪ Correction of attributes and manifest data associated with of VGT-P reflectances  
▪ Configuration of processing site in S3 L2 product manifest |
| SYN_L2V.002.16.00   |                   | 05/09/2022 (S3B)     |         |
| Global PB_ID 3.11   |                   |                       |         |
| SY2_VGS IPF v06.11  | 14/07/2022        | 23/08/2022 (S3A)     | ▪ Configuration of processing site in S3 L2 product manifest  
▪ Correction of the Empty crs dataset in all VGT-like products  
▪ Correction of the VGT status map when VGT radiometry is set to NaN  
▪ Correction of attributes and manifest data associated with of VGT-P reflectances |
| SYN_L2V.002.08.00   |                   | 05/09/2022 (S3B)     |         |
| Global PB_ID 3.11   |                   |                       |         |
| SY2_AOD IPF v01.06  | 14/07/2022        | 23/08/2022 (S3A)     | ▪ Configuration of processing site in S3 L2 product manifest |
| AOD_NTC.002.06.01   |                   | 05/09/2022 (S3B)     |         |
| Global PB_ID 3.11   |                   |                       |         |
| SY2 IPF v06.24      | 09/12/2022        | Not yet deployed      | ▪ Suppression of the SLSTR contribution into B2 and B3 composition  
▪ Update of the Spectral Response Function ADF  
▪ Reading of the OLCI L1 uncertainties inclusion and transfer into SYN L2 products when appropriate  
▪ Adaptation to the OLCI Geometric Correction Change and frame offset new format.  
▪ Adaptation to the OLCI partial saturation new flags  
▪ Inclusion of DOI in the manifest  
▪ Inclusion of SLSTR calibrations factors into SYN L2  
▪ Inconsistency between Time variables attributes and data type  
▪ Inclusion of dedicated saturation flags in VGT products |
| SYN_L2V.002.17.00   |                   |                       |         |
| Global PB_ID 3.17   |                   |                       |         |
| SY2_VGS IPF v06.12  | 09/12/2022        | Not yet deployed      | ▪ Inclusion of dedicated saturation flags in VGT products |
| SYN_L2V.002.09.00   |                   |                       |         |
| Global PB_ID 3.14   |                   |                       |         |
| SY2_AOD IPF v01.07  | 09/12/2022        | Not yet deployed      | ▪ Reading of the OLCI L1 uncertainties inclusion and transfer into SYN L2 products when appropriate  
▪ Adaptation to the OLCI Geometric Correction Change and frame offset new format.  
▪ Adaptation to the OLCI partial saturation new flags  
▪ Recommendations from LAW project: inclusion of SLSTR calibration factors and application of a filtering over pixel associated with high SZA |
| AOD_NTC.002.07.00   |                   |                       |         |
| Global PB_ID 3.17   |                   |                       |         |
4 Calibration and characterisation changes

4.1 OLCI

4.1.1 Instrument settings
There has been no change to the OLCI-A or OLCI-B instrument setting during the reporting period.

4.1.2 Evolutions in Radiometric Calibration of EO data

4.1.2.1 OLCI-A
The following evolutions of the EO radiometric calibration auxiliary data have been implemented since beginning of 2022:

❖ 20/01/2022: PB OL__L1_.002.22.00 updated the Radiometric Gain Model
❖ 19/07/2022: PB OL__L1_.002.23.00 updated the Radiometric Gain Model

4.1.2.2 OLCI-B
The following evolutions of the EO radiometric calibration auxiliary data have been implemented since beginning of 2022:

❖ 20/01/2022: PB OL__L1_.002.22.00 updated the Radiometric Gain Model
❖ 19/07/2022: PB OL__L1_.002.23.00 updated the Radiometric Gain Model

4.2 SLSTR

4.2.1 Instrument settings

4.2.1.1 SLSTR-A
The S8 detector offset voltage was adjusted on 26th January 2021 to correct the lower limit of the dynamic range following the cooler cold tip temperature increase in October 2020.

The commanded blackbody heater power was reduced by one bit on 28th September 2021. The effect of this commanded change was to reduce the hot blackbody temperature by ~0.3 K.

4.2.1.2 SLSTR-B
The commanded blackbody heater power was reduced by one bit on 1st October 2021. This change had very little effect on the blackbody temperature.

4.2.2 Evolutions in Radiometric Calibration of EO data
No updates to the SLSTR-A or SLSTR-B radiometric calibration parameters have been applied in the reporting period.
5 Summary of performances – OLCI

5.1 Instrument performances

5.1.1 Temperature stability

5.1.1.1 OLCI-A

CCD temperatures are monitored on the long-term using data from Radiometric Calibration acquisitions (see Figure 2). Variations are very small (0.09 °C peak-to-peak) and no trend can be identified.

Figure 2: Long term monitoring of CCD temperatures using minimum value (top), time averaged values (middle), and maximum value (bottom) provided in the annotations of the Radiometric Calibration Level 1 products, for the Shutter frames, all radiometric calibrations so far.
5.1.1.2 OLCI-B

As for OLCI-A, the variations of CCD temperature are very small (0.08°C peak-to-peak) and no trend can be identified.

Figure 3: long term monitoring of OLCI-B CCD temperatures using minimum value (top), time averaged values (middle), and maximum value (bottom) provided in the annotations of the Radiometric Calibration Level 1 products, for the Shutter frames, all radiometric calibrations so far except the first one (absolute orbit 167) for which the instrument was not yet thermally stable.
5.1.2  Signal to noise ratio

5.1.2.1  OLCI-A

OLCI signal to noise ratio (SNR) is monitored using Radiometric Calibration data acquired on the radiometric diffuser that provides a signal smoothly varying with time. After correction for the variation due to the variation of the illumination with illumination geometry during the 24 seconds of acquisitions, variability is assessed and SNR is derived, as the incoming radiance is known. SNR values obtained at the Calibration signal level are then downscaled to a typical clear sky ocean signal level, as defined in the mission requirements.

SNR computed for all radiometric calibration data is presented on Figure 4 as a function of band number. Stability with time is shown on Figure 5: SNR of band Oa01 (400nm, the most varying) is plotted against orbit number.

There is no significant evolution of this parameter over the mission and the ESA requirement is fulfilled for all bands.
Figure 4: OLCI-A Signal to Noise ratio as a function of the spectral band for the 5 cameras. These results have been computed from radiometric calibration data. All calibrations except first one (orbit 183) are present with the colours corresponding to the orbit number (see legend). The SNR is very stable with time: the curves for all orbits are almost superimposed. The dashed curve is the ESA requirement.
Figure 5: OLCI-A long-term stability of the SNR estimates from Calibration data, example of channel Oa01.

The mission averaged SNR figures are provided in Table 4, together with their radiance reference level. According to the OLCI SNR requirements, these figures are valid at these radiance levels and at Reduced Resolution (RR, 1.2 km). They can be scaled to other radiance levels assuming shot noise (CCD sensor noise) is the dominating term, i.e. radiometric noise can be considered Gaussian with its standard deviation varying as the square root of the signal: $SNR(L) = SNR(L_{ref}) \cdot \sqrt{\frac{L}{L_{ref}}}$. Following the same assumption, values at Full Resolution (300 m) can be derived from RR ones as 4 times smaller.
Table 4: OLCI-A SNR figures as derived from Radiometric Calibration data. Figures are given for each camera (time average and standard deviation), and for the whole instrument. The requirement and its reference radiance level are recalled (in mW.sr$^{-1}$.m$^{-2}$.nm$^{-1}$).

<table>
<thead>
<tr>
<th>nm</th>
<th>$L_{ref}$</th>
<th>SNR</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LU</td>
<td>RQT</td>
<td>avg</td>
<td>std</td>
<td>avg</td>
<td>std</td>
<td>avg</td>
<td>std</td>
</tr>
<tr>
<td>400.0</td>
<td>63.0</td>
<td>2188</td>
<td>2421</td>
<td>6.2</td>
<td>2398</td>
<td>6.2</td>
<td>2333</td>
<td>8.3</td>
</tr>
<tr>
<td>412.0</td>
<td>74.1</td>
<td>2061</td>
<td>2386</td>
<td>9.4</td>
<td>2403</td>
<td>7.4</td>
<td>2339</td>
<td>5.0</td>
</tr>
<tr>
<td>442.0</td>
<td>65.6</td>
<td>1811</td>
<td>2157</td>
<td>6.0</td>
<td>2196</td>
<td>6.1</td>
<td>2163</td>
<td>5.0</td>
</tr>
<tr>
<td>490.0</td>
<td>51.2</td>
<td>1541</td>
<td>1999</td>
<td>4.8</td>
<td>2036</td>
<td>4.7</td>
<td>1998</td>
<td>4.3</td>
</tr>
<tr>
<td>510.0</td>
<td>44.4</td>
<td>1488</td>
<td>1979</td>
<td>5.3</td>
<td>2014</td>
<td>4.9</td>
<td>1986</td>
<td>4.5</td>
</tr>
<tr>
<td>560.0</td>
<td>31.5</td>
<td>1280</td>
<td>1775</td>
<td>4.7</td>
<td>1802</td>
<td>4.1</td>
<td>1803</td>
<td>4.7</td>
</tr>
<tr>
<td>620.0</td>
<td>21.1</td>
<td>997</td>
<td>1591</td>
<td>4.1</td>
<td>1608</td>
<td>4.3</td>
<td>1624</td>
<td>3.1</td>
</tr>
<tr>
<td>665.0</td>
<td>16.4</td>
<td>883</td>
<td>1545</td>
<td>4.2</td>
<td>1557</td>
<td>4.5</td>
<td>1566</td>
<td>3.9</td>
</tr>
<tr>
<td>674.0</td>
<td>15.7</td>
<td>707</td>
<td>1328</td>
<td>3.4</td>
<td>1337</td>
<td>3.7</td>
<td>1350</td>
<td>2.8</td>
</tr>
<tr>
<td>681.0</td>
<td>15.1</td>
<td>745</td>
<td>1319</td>
<td>3.5</td>
<td>1325</td>
<td>3.3</td>
<td>1338</td>
<td>2.6</td>
</tr>
<tr>
<td>709.0</td>
<td>12.7</td>
<td>785</td>
<td>1420</td>
<td>4.2</td>
<td>1420</td>
<td>4.0</td>
<td>1435</td>
<td>3.3</td>
</tr>
<tr>
<td>754.0</td>
<td>10.3</td>
<td>605</td>
<td>1127</td>
<td>3.0</td>
<td>1121</td>
<td>2.8</td>
<td>1136</td>
<td>3.1</td>
</tr>
<tr>
<td>761.0</td>
<td>6.1</td>
<td>232</td>
<td>502</td>
<td>1.1</td>
<td>498</td>
<td>1.1</td>
<td>505</td>
<td>1.1</td>
</tr>
<tr>
<td>764.0</td>
<td>7.1</td>
<td>305</td>
<td>663</td>
<td>1.5</td>
<td>658</td>
<td>1.5</td>
<td>668</td>
<td>2.0</td>
</tr>
<tr>
<td>768.0</td>
<td>7.6</td>
<td>330</td>
<td>558</td>
<td>1.4</td>
<td>554</td>
<td>1.3</td>
<td>563</td>
<td>1.3</td>
</tr>
<tr>
<td>779.0</td>
<td>9.2</td>
<td>812</td>
<td>1516</td>
<td>4.6</td>
<td>1498</td>
<td>4.4</td>
<td>1527</td>
<td>5.1</td>
</tr>
<tr>
<td>865.0</td>
<td>6.2</td>
<td>666</td>
<td>1243</td>
<td>3.5</td>
<td>1213</td>
<td>3.4</td>
<td>1239</td>
<td>3.8</td>
</tr>
<tr>
<td>885.0</td>
<td>6.0</td>
<td>395</td>
<td>823</td>
<td>1.6</td>
<td>801</td>
<td>1.6</td>
<td>814</td>
<td>1.9</td>
</tr>
<tr>
<td>900.0</td>
<td>4.7</td>
<td>308</td>
<td>691</td>
<td>1.6</td>
<td>673</td>
<td>1.3</td>
<td>683</td>
<td>1.6</td>
</tr>
<tr>
<td>940.0</td>
<td>2.4</td>
<td>203</td>
<td>534</td>
<td>1.2</td>
<td>522</td>
<td>1.2</td>
<td>525</td>
<td>1.0</td>
</tr>
<tr>
<td>1020.0</td>
<td>3.9</td>
<td>152</td>
<td>345</td>
<td>0.9</td>
<td>337</td>
<td>0.8</td>
<td>348</td>
<td>0.7</td>
</tr>
</tbody>
</table>
5.1.2.2 OLCI-B

As for OLCI-A there is no significant evolution of the SNR over the mission and the ESA requirement is fulfilled for all bands.

![OLCI-B SNR plots](image)

Figure 6: OLCI-B Signal to Noise ratio as a function of the spectral band for the 5 cameras. These results have been computed from radiometric calibration data. All calibrations except first one (orbit 167) are presented with the colours corresponding to the orbit number (see legend). The SNR is very stable with time: the curves for all orbits are almost superimposed. The dashed curve is the ESA requirement.

As for OLCI-A, the mission averaged SNR figures are provided in Table 5 below, together with their radiance reference level.
**Table 5: OLCI-B SNR figures as derived from Radiometric Calibration data. Figures are given for each camera (time average and standard deviation), and for the whole instrument. The requirement and its reference radiance level are recalled (in mW.sr\(^{-1}.m^2.nm\(^{-1}\)).**

<table>
<thead>
<tr>
<th>L(_{\text{ref}})</th>
<th>SNR</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>nm</td>
<td>LU</td>
<td>RQT</td>
<td>avg</td>
<td>std</td>
<td>avg</td>
<td>std</td>
<td>avg</td>
</tr>
<tr>
<td>400.000</td>
<td>63.0</td>
<td>2188</td>
<td>2456</td>
<td>18.8</td>
<td>2296</td>
<td>16.6</td>
<td>2419</td>
</tr>
<tr>
<td>412.000</td>
<td>74.1</td>
<td>2061</td>
<td>2654</td>
<td>7.0</td>
<td>2569</td>
<td>6.4</td>
<td>2543</td>
</tr>
<tr>
<td>442.000</td>
<td>65.6</td>
<td>1811</td>
<td>2323</td>
<td>6.6</td>
<td>2316</td>
<td>6.2</td>
<td>2299</td>
</tr>
<tr>
<td>490.000</td>
<td>51.2</td>
<td>1541</td>
<td>1966</td>
<td>4.9</td>
<td>1990</td>
<td>5.6</td>
<td>1971</td>
</tr>
<tr>
<td>510.000</td>
<td>44.4</td>
<td>1488</td>
<td>1939</td>
<td>4.9</td>
<td>1968</td>
<td>6.1</td>
<td>1943</td>
</tr>
<tr>
<td>560.000</td>
<td>31.5</td>
<td>1280</td>
<td>1813</td>
<td>4.7</td>
<td>1848</td>
<td>4.9</td>
<td>1829</td>
</tr>
<tr>
<td>620.000</td>
<td>21.1</td>
<td>997</td>
<td>1572</td>
<td>4.2</td>
<td>1626</td>
<td>4.6</td>
<td>1624</td>
</tr>
<tr>
<td>665.000</td>
<td>16.4</td>
<td>883</td>
<td>1513</td>
<td>4.2</td>
<td>1579</td>
<td>3.8</td>
<td>1573</td>
</tr>
<tr>
<td>674.000</td>
<td>15.7</td>
<td>707</td>
<td>1300</td>
<td>3.9</td>
<td>1358</td>
<td>3.5</td>
<td>1353</td>
</tr>
<tr>
<td>681.000</td>
<td>15.1</td>
<td>745</td>
<td>1293</td>
<td>3.5</td>
<td>1347</td>
<td>3.2</td>
<td>1343</td>
</tr>
<tr>
<td>709.000</td>
<td>12.7</td>
<td>785</td>
<td>1390</td>
<td>3.9</td>
<td>1447</td>
<td>4.0</td>
<td>1443</td>
</tr>
<tr>
<td>754.000</td>
<td>10.3</td>
<td>605</td>
<td>1096</td>
<td>3.6</td>
<td>1143</td>
<td>3.6</td>
<td>1142</td>
</tr>
<tr>
<td>761.000</td>
<td>6.1</td>
<td>232</td>
<td>488</td>
<td>1.2</td>
<td>509</td>
<td>1.2</td>
<td>509</td>
</tr>
<tr>
<td>764.000</td>
<td>7.1</td>
<td>305</td>
<td>643</td>
<td>1.6</td>
<td>673</td>
<td>2.0</td>
<td>672</td>
</tr>
<tr>
<td>768.000</td>
<td>7.6</td>
<td>330</td>
<td>541</td>
<td>1.5</td>
<td>568</td>
<td>1.4</td>
<td>564</td>
</tr>
<tr>
<td>779.000</td>
<td>9.2</td>
<td>812</td>
<td>1467</td>
<td>4.1</td>
<td>1535</td>
<td>4.7</td>
<td>1527</td>
</tr>
<tr>
<td>865.000</td>
<td>6.2</td>
<td>666</td>
<td>1221</td>
<td>3.5</td>
<td>1287</td>
<td>3.8</td>
<td>1258</td>
</tr>
<tr>
<td>885.000</td>
<td>6.0</td>
<td>395</td>
<td>808</td>
<td>2.2</td>
<td>848</td>
<td>1.9</td>
<td>834</td>
</tr>
<tr>
<td>900.000</td>
<td>4.7</td>
<td>308</td>
<td>679</td>
<td>1.5</td>
<td>714</td>
<td>1.9</td>
<td>704</td>
</tr>
<tr>
<td>940.000</td>
<td>2.4</td>
<td>203</td>
<td>527</td>
<td>1.2</td>
<td>549</td>
<td>1.5</td>
<td>551</td>
</tr>
<tr>
<td>1020.000</td>
<td>3.9</td>
<td>152</td>
<td>336</td>
<td>0.8</td>
<td>358</td>
<td>1.2</td>
<td>358</td>
</tr>
</tbody>
</table>
5.1.3 Spectral Calibration

5.1.3.1 OLCI-A

OLCI’s spectral characteristics are regularly monitored in-flight by different spectral campaigns, which are shortly outlined in the following. A detailed description is given in S3-TN-ESA-OL-660. The procedures use the programming capability of OLCI to define 45 bands around stable spectral features, to characterize the spectral dispersion of each camera system with respect to the spectral dimension and the spatial (across track) dimension. Simulations of OLCI measurements in the 45 bands are optimized for best agreement with the spectral features, as a function of individual bandwidth and band centre wavelength. Depending on the used spectral feature the achieved accuracy for the centre wavelength is in the order of 0.1-0.2 nm, the precision (repeatability) is better than 0.05 nm.

Three different calibration sequences S0* are used regularly:

❖ S09: The 45 bands are grouped around the atmospheric oxygen absorption band at 770 nm and around distinct solar Fraunhofer lines at 485 nm, 656 nm and 854 nm. To increase the feature stability, the same few hundred frames are acquired at the same position in the orbital cycle (relative orbit number 107), over Libyan Desert. Since May 2016, 22 S09 campaigns have been performed.

❖ S02/S03: The 45 bands are grouped around three spectral features of the on-board spectral diffuser at 405 nm, 520 nm and 800 nm. 500 frames are acquired on the white diffuser (S02) as reference and on the spectral (so called pink) diffuser (S03). Since thermal stabilisation of the instrument, 22 S02/S03 campaigns have been performed, starting in April 2016.

❖ S02 solar: Solely the white diffuser data is used to identify and utilize solar Fraunhofer lines and to provide therewith a spectral characterization independent from the on-board spectral diffuser.

The spectral campaigns performed during and after the commissioning phase reveal a high agreement of the in-flight characterisation with the instrument spectral model derived from the pre-flight spectral calibration. The resulting differences of the centre wavelengths of the nominal OLCI bands between pre- and in-flight calibration are smaller than 0.1 nm, despite of band 1 and 21, where differences <= 0.2 nm have been detected (see in Figure 7 the example of Spectral Diffuser Line #1, camera 3).
Figure 7: OLCI-A across track spectral calibration from all S02/S03 sequences since the beginning of the mission. Upper top plot is spectral line 1; middle plot is spectral line 2 and bottom plot spectral line 3. On-ground spectral characterisation is in red and nominal spectral assignment is in red dotted line.

A long-term temporal evolution can be observed since the first in-flight characterisation. This is shown in Figure 8 (S02/S03) and Figure 9 (S09), where the camera mean spectral distance to its value since respectively orbit 881 (April 2016) and orbit 1107 (May 2016) is plotted.

The evolution of the centre wavelength is different for each camera but approximately the same for all wavelengths. Since the end of the commissioning phase (June 2016, ~ orbit 1800) the observed changes are smaller than 0.15 nm (0.23 nm for camera 5).

We see that the long-term evolution of the spectral calibration obtained with sequence S09 is in good agreement with the one obtained with sequence S02/S03. Both show the same trends: very small interannual variations and the even smaller drift (0.01nm / 4yrs) towards shorter wavelengths.
Figure 8: OLCI-A camera averaged spectral calibration evolution as a function of absolute orbit number (all spectral S02/S03 calibrations since the beginning of the mission are included except the very first one (1st March 2016, orbit 195)). The data are normalized with the first Spectral Calibration of the plot, which is from 18 April 2016 (orbit 881). The last spectral Calibration is from 07 January 2023 (orbit 35886/35887).

Figure 9: OLCI-A line-averaged spectral calibration relative to the one acquired on 4th May 2016 (orbit 1107), as a function of time derived from all S09 sequences. The last calibration is from 07 January 2023 (orbit 35885). For each camera, the spectral evolution derived from spectral lines at 485 nm, 656 nm, 770 nm and 854 nm have been averaged.
5.1.3.2 OLCI-B

ACT profiles of absolute spectral calibration obtained with all S02/S03 sequences, including comparison with on-ground characterisation, are plotted in Figure 10, showing the very good agreement between pre-flight and in-flight spectral calibrations. Differences are roughly < 0.2 nm except for line 3 camera 2, which is < 0.3 nm.

![Figure 10: OLCI-B across track spectral calibration from all S02/S03 sequences since the beginning of the mission. Upper top plot is spectral line 1; Middle plot is spectral line 2 and bottom plot spectral line 3. On-ground spectral characterisation is in red and nominal spectral assignment is in red dotted line.](image)

For OLCI-B there has been 17 S02/S03 campaigns and 35 S09.

Figure 11 shows the temporal evolution of the spectral calibration obtained with all S02/S03 sequences since the beginning of the mission. As for OLCI-A a small drift is observed. For OLCI-B, this drift is positive for camera 1, 2, 4 and 5 and negative for camera 3.

Evolution derived from the S09 calibration sequence (spectral calibration using O2 absorption and Fraunhofer lines) is presented in Figure 12. As for OLCI-A, we see that the long-term evolution of the spectral calibration obtained with sequence S09 is in good agreement with the one obtained with sequence S02/S03.
Figure 11: OLCI-B camera averaged spectral calibration evolution as a function of absolute orbit number (all spectral S02/S03 calibrations since the beginning of the mission are included). The data are normalized with the first Spectral Calibration. The first (reference) calibration is from 8 May 2018 (orbit 182), the last 17 Jan. 2023 (orbit 24635/24636).

Figure 12: OLCI-B camera averaged spectral calibration evolution as a function of absolute orbit number from S09 calibrations since the beginning of the mission. The first (reference) calibration is from 8 May 2018, the last is from 17 Jan. 2023 (orbit 24634). For each camera, the spectral evolution derived from spectral lines at 485 nm, 656 nm, 770 nm and 854 nm have been averaged. The data are normalized with the first Spectral Calibration. For better comparability, the orbit range (abscissa) is equal to that of OLCI-A (Figure 8).
5.1.4 Radiometric stability

5.1.4.1 OLCI-A

The stability with time of the instrument sensitivity is monitored through the radiometric calibration processing results: time series of radiometric gains normalised to a given date are analysed. This is done at the full spatial resolution before being summarised by spatial averaging over each camera: if there is some variability of the sensitivity evolution for a given channel inside a given camera, it remains limited with respect to camera-to-camera variability.

The overall instrument evolution (since channel programming change, 25/04/2016 to 26/01/2023) is shown on Figure 13: a maximum of about 3.1% is reached at 400 nm, with a high inter-camera variability, while other bands show much lower values, within ±1% (-1.8% for band Oa2 camera 1). The spectral behaviour of the 5 cameras is very similar, to the exception of camera 1 at the blue edge (bands Oa1 and Oa2, 400 & 412 nm), and camera 5 in the red to NIR spectral range.

![Figure 13: OLCI-A camera-averaged instrument evolution since channel programming change (25/04/2016) and up to most recent calibration (26/01/2023) versus wavelength.](image)

Time series of sensitivity evolution are shown on Figure 14 one plot per camera, as a function of elapsed time since launch. It shows that, if a significant evolution occurred during the early mission, the trends tend to stabilize, with the notable exception of band 1 in particular for camera 4. An example of an evolution surface for channel Oa2 (412 nm) is given below for Camera 1 (Figure 15), justifying the use of spatial averages for long-term monitoring.
Figure 14: OLCI-A camera averaged gain relative evolution with respect to calibration of 25/04/2016 (change of OLCI channel settings), as a function of elapsed time since launch; one curve for each band (see colour code on plots), one plot for each module. The diffuser ageing has been taken into account. Early mission data (16 Feb. to 25 April) is not available due to missing information required for accurate gain computation.
5.1.4.2 OLCI-B

The overall instrument evolution (18/06/2018 to 30/01/2023) is shown on Figure 16: a maximum of about 4.5% is reached at 400 nm, while other bands show lower values, within ± 1.5%. The spectral behaviour of the 5 cameras is very similar, to the exception of camera 3 at both edges (bands Oa1 and Oa21, 400 & 1020 nm).
Figure 16: OLCI-B camera-averaged instrument evolution since channel programming change (18/06/2018) and up to most recent calibration (30/01/2023) versus wavelength.

Time series of sensitivity evolution are shown on Figure 17, one plot per camera, as a function of elapsed time since launch. It shows that, if a significant evolution occurred during the early mission, the trends tend to stabilize.
5.1.5 Ageing of radiometric diffuser

5.1.5.1 OLCI-A

The ageing of the nominal radiometric solar diffuser is monitored using a second, or reference, radiometric diffuser. The relative darkening of the solar diffuser, expected to be measurable after significant cumulated exposure to UV light, is assessed at every channel through the evolution with time.
of the relative response of the nominal diffuser with respect to that of the reference one acquired under almost identical illumination conditions one orbit after the nominal one; the first pair of measurements is used as the reference point. Ageing is first assessed at every spatial pixel and then averaged over the field-of-view (FOV) as independent of the instrument itself.

FOV-averaged ageing as a function of wavelength is represented in Figure 18 for all available ageing acquisition (29 so far, excluding the first sequence used as the reference). As expected, ageing is rather low (<0.75% after about 7 years) and stronger for the ‘bluest’ spectral bands (short wavelengths). At present, ageing is clearly visible for wavelengths up to about 650 nm.

Figure 18: OLCI-A Diffuser 1 ageing as a function of wavelength (or spectral band). Ageing is clearly visible in spectral band #1 to #6.

A model of the nominal diffuser ageing is derived by fitting the measured ageing against cumulated exposure to light, so that it can be used to accurately predict (or model) the nominal diffuser reflectance at any time. This model is used to derive the OLCI Radiometric Gain Model (see section 5.2.1.2). The slope of this ageing model (% of reflectance loss per exposure) as a function of wavelength is presented in 24 for eighteen consecutive estimations (during orbit cycles 20, 24, 27, 29, 33, 38, 40, 43, 47, 52, 54, 56, 58, 60, 65, 70, 74, 80, 01/2022, 03/2022, 07/2022, 11/2022 and 01/2023 i.e. between July 2017 and January 2023). It shows that the stability is excellent.
Figure 19: Slope of ageing fit (% of loss per exposure) vs wavelengths, using all the available ageing sequence at the time of the most recent monthly Data Quality Report with an ageing measurement (January 2023 = red curve), and at the time of the 23 previous cycles with an aging sequence (see legend above the curves).

5.1.5.2 OLCI-B

OLCI-B FOV-averaged ageing as a function of wavelength is represented in Figure 20 for all available ageing acquisition (21 so far, excluding the first sequence used as the reference). The ageing is clearly visible in spectral band Oa01 to Oa05, with the expected spectral shape and order of magnitude. However, we also observe some ageing in bands Oa06 to Oa11; such an unexpected behaviour preventing further use of the nominal ageing assessment method, an alternative one is reported below and used operationally.
As for OLCI-A, the OLCI-B Diffuser Ageing has been modelled as a function of cumulated exposure time (i.e. number of acquisition sequence on nominal diffuser, regardless of the band setting). The OLCI-A modelling methodology has been applied to OLCI-B. The results of this modelling, iterated at each new Ageing Sequence acquisition, expressed as the rate of ageing (% of loss per exposure) as a function of wavelength is presented in Figure 21. The unexpected bump near 650-700 nm mentioned in previous annual reports seems to decrease with time (i.e with the quantity of data used for modelling the ageing) which is a good point since there is no expected significant ageing at these wavelengths.
Figure 21: OLCI-B: Slope of ageing fit (% of loss per exposure) vs wavelengths, using all the available ageing sequence at the time of the current cycle (red curve) and at the time of previous cycle for which an ageing sequence was measured (see legend within the figure).

It has been shown that the unexpected behaviour of the ageing rates is linked to differences in the BRDF of the two diffusers. An upgrade of the ageing assessment method is based on the grouping of Ageing Calibrations sequences by similar illumination geometry (which are reproduced year after year) that are analysed independently. This upgrade allowed to retrieve three fairly consistent independent Ageing assessments, that also compare better with OLCI-A values. These estimates are lower than previous ones, derived from Yaw Manoeuvres (YM, see Annual Performance Report for 2020) with an impact of up to about 0.2% on radiances when used to derive the Gain Model (see section 5.2.2.2).
Figure 22: Ageing rates (in % per exposure) as derived using nominal methods (at cycles 28, 46 and 62), and successive upgrades (Yaw Manoeuvres and Equal SZA at 3 different dates).

5.2 L1 products performances

5.2.1 Geometric Performance

Regular monitoring of the geolocation performance by correlation with GCP (Ground Control Points) imagettes using the so-called GeoCal Tool is done continuously.

5.2.1.1 OLCI-A

The good performance of OLCI-A georeferencing since the introduction of the upgraded Geometric Calibration on 14/03/2018 is confirmed. It has however significantly improved after its last full revision of GCMs (Geometric Calibration Models, or platform to instrument alignment quaternions) and IPPVMs (Instrument Pixels Pointing Vectors) both derived using the GeoCal Tool and put in production on 30/07/2019. The following figures show time series of the overall RMS performance (Figure 23, requirement criterion) and of the across-track and along-track biases for each camera (Figure 24 to Figure 28). Figure 29 and Figure 30 address the monitoring of the performance homogeneity within the field of view: georeferencing errors in each direction at camera transitions (difference between last pixel of
camera N and first pixel of camera N+1) and within a given camera (maximum bias minus minimum inside each camera).

The performance improvement since the 30/07/2019 is significant on most figures: the global RMS value decreases form around 0.35 to about 0.2 and remains below 0.3 since then (Figure 23), the across-track biases decrease significantly for all cameras (Figure 24 to Figure 28), the along-track bias reduces where it was significant (camera 3, Figure 26) and the field of view homogeneity improves drastically (Figure 29 and Figure 30). It is also worth to mention a reduction of the dispersion – distance between the ± 1 sigma lines – in Figure 24 to Figure 28). Along-track biases of cameras 3 to 5 are however still slightly drifting, resulting in slowly degrading RMS performance (Figure 23), but this is closely monitored so that appropriate actions can be taken.

![Image](image.png)

**Figure 23**: overall OLCI-A georeferencing RMS performance time series over the whole monitoring period (left) and number of validated control points corresponding to the performance time series (right).

![Image](image.png)

**Figure 24**: across-track (left) and along-track (right) georeferencing biases time series for Camera 1 (starting 01/03/2018).
Figure 25: same as Figure 24 for Camera 2.

Figure 26: same as Figure 24 for Camera 3.

Figure 27: same as Figure 24 for Camera 4.
Figure 28: same as Figure 24 for Camera 5.

Figure 29: OLCI-A spatial across-track misregistration at each camera transition (left) and maximum amplitude of the across-track error within each camera (left).
5.2.1.2 OLCI-B

The performance of OLCI-B georeferencing is within requirements since the introduction of the 3rd Geometric Calibration on 12/12/2018. Significant persistent along-track depointing drifts required several re-calibration since the: a major upgrade was introduced on 30/07/2019, followed by further adjustments, the most recent being that of 29/07/2021. The following figures show time series of the overall RMS performance (Figure 31, requirement criterion), the across-track and along-track biases for each camera (Figure 32 to Figure 36), as well as pointing homogeneity across the field of view (bias differences at camera interfaces and bias amplitudes within each camera, Figure 37 and Figure 38).

As for OLCI-A, despite compliance to the RMS requirement of 0.5 pixel, OLCI-B showed significant heterogeneity of the performance within the field of view, with discrepancies at camera transitions of up to 1 pixel. Introduction of upgraded pointing vectors (first occurrence 30/07/2019) greatly improved many performance indicators: the global RMS value decreases from around 0.4 to about 0.3 (Figure 31), the across-track biases decrease significantly for all cameras (Figure 32 to Figure 36) and the field of view homogeneity improves drastically (Figure 37 and Figure 38, but also reduction of the dispersion – distance between the ± 1 sigma lines – in Figure 32 to Figure 36).

The global RMS performance as well as the along and across-track average biases are quite stable since then, however in-FOV across-track pointing homogeneity slowly degrades continuously (Figure 37) and frequent re-calibrations were necessary to maintain the performance.
Figure 31: overall OLCI-B georeferencing RMS performance time series (left) and number of validated control points corresponding to the performance time series (right) over the whole monitoring period.

Figure 32: across-track (left) and along-track (right) georeferencing biases time series for Camera 1.

Figure 33: same as Figure 32 for Camera 2.
Figure 34: same as Figure 32 for Camera 3.

Figure 35: same as Figure 32 for Camera 4.

Figure 36: same as Figure 32 for Camera 5.
Figure 37: OLCI-B spatial across-track misregistration at each camera transition (left) and maximum amplitude of the across-track error within each camera (left).

Figure 38: OLCI-B spatial along-track misregistration at each camera transition (left) and maximum amplitude of the along-track error within each camera (left).
5.2.2 Radiometric Gain Model Performance

5.2.2.1 OLCI-A

OLCI radiometric Calibration is based on its on-board calibration system: a carefully characterised solar diffuser is used as a secondary radiometric standard to derive instantaneous radiometric gains from diffuser measurements and computation of the incoming radiance, by use of diffuser characterisation, illumination and viewing geometry as well as spectral response functions.

OLCI Level 1 data processing to calibrate measured radiances using a Radiometric Gain Model (RGM) includes a long term drift correction, in order to avoid radiometric discontinuities between successive gain estimates as well as simplifying maintenance of operational processing configuration. The model is expressed as a bounded exponential time evolution applied onto the gain at a reference date. The time evolution model is fitted, on a per band and per pixel basis, on the evolution data presented above (section 5.1.4.1); the Gain at the reference date is obtained by time averaging after correction of the evolution. Diffuser ageing (see section 5.1.5) is of course accounted for during this process.

Consequently, the model is always used in extrapolation for routine production, as derived from already acquired data; it can only be used in interpolation for data reprocessing. Its performance is thus continuously monitored against new radiometric calibration, regularly acquired. The current operational RGM has been derived from data spanning 23/10/2018 to 30/04/2022 and put in operations the 19/07/2022 (processing baseline OL__L1_002.23.00). It includes the correction of the diffuser ageing for the six bluest bands (Oa1 to Oa6) for which it is clearly measurable.

The model RMS performance over the complete dataset (including the 17 calibrations in extrapolation over about 9 months) remains better than 0.1% – except for channels Oa1 (400 nm) that reaches 0.16% near orbit 33000 – when averaged over the whole field of view (Figure 39) even if a small drift of the model with respect to most recent data is already visible in most channels.
Figure 39: RMS performance of the Gain Model of current Processing Baseline as a function of orbit.

More details are provided on Figure 40 on which per camera mean and standard deviation of Model over Data ratios are plotted against wavelength for each orbit. Conclusions are however the same with performances within 0.1% (1-σ) but for Oa1 and Oa21, the former reaching 0.15% in cameras 4, while the latter has a larger dispersion (up to 0.2%) in camera 5, due to a group of pixels with an anomalous behaviour that cannot be fully captured by the model mathematical expression.
Figure 40: For the 5 cameras: Evolution model performance, as camera-average and standard deviation of ratio of Model over Data vs. wavelength, for each orbit of the test dataset, including 17 calibrations in extrapolation, with a colour code for each calibration from blue (oldest) to red (most recent).
5.2.2.2  OLCI-B

Instrument response and degradation modelling for OLCI-B, including the use of the in-flight BRDF model (based on 11th December 2018 Yaw Manoeuvres), has been refreshed and deployed at PDGS on 19th July 2022 (Processing Baseline OL__L1_002.23.00). The model has been derived from a more recent Radiometric Calibration dataset (from 13/04/2019 to 29/04/2022). It includes the correction of the diffuser ageing for the five bluest bands (Oa1 to Oa5) for which it is clearly measurable. The model performance over the complete dataset (including the 17 calibrations in extrapolation over about 9 months) is illustrated in Figure 41. It remains better than 0.1% when averaged over the whole field of view for most bands with peaks up to 0.13% for Oa01 and Oa02. A slight drift of the model with respect to the most recent data seems to appear for all bands.

![Global Fit RMS Performances, whole FOV, vs. time](image)

Figure 41: RMS performance of the OLCI-B Gain Model of the current processing baseline as a function of orbit.

More details are provided on Figure 42 on which per camera mean and standard deviation of Model over Data ratios are plotted against wavelength for each orbit.
Figure 42: For the 5 cameras: Evolution model performance, as camera-average and standard deviation of ratio of Model over Data vs. wavelength, for each orbit of the test dataset, including 10 calibrations in extrapolation, with a colour code for each calibration from blue (oldest) to red (most recent).
5.2.3 Radiometric Validation

Radiometric Validation is performed at the Optical MPC using three indirect methods, comparing simulated TOA radiances to that measured by the OLCI instrument.

- The “Rayleigh” method: measurement of the Rayleigh atmospheric backscattering over open ocean sites in clear sky off-glint conditions with low aerosol load to provide absolute calibration in the blue-to-red spectral domain.

- The “Glint” method: using the specular reflection of the sun (i.e. sun glint) on the open ocean surface and its known spectral dependency to assess inter-band calibration in the red-to-NIR spectral range. Its results are then scaled to an absolute result at a reference band, 665 nm, taken from the PICS method for DIMITRI and from the Rayleigh method for OSCAR.

- The PICS method: measurement over well characterized, temporally stable desert areas (Pseudo-Invariant Calibration Sites or PICS) to provide absolute calibration over the whole spectral domain. This method also allows cross-mission intercomparison with other sensors providing comparable spectral channels (e.g. Aqua/MODIS, S2A/MSI and MERIS/3REP).

The first two methods are undertaken by two different implementations: DIMITRI operated by ARGANS, and OSCAR operated by VITO. The third method is only implemented in DIMITRI.

Despite their discrepancies, more or less within their claimed accuracies, all methods do point out an excess of brightness for OLCI-A radiances (Figure 43, Figure 44, Figure 49, and Table 9). Results are in pretty close agreement (except DIMITRI Rayleigh, a bit higher) around 2-3% between 510 and 900 nm. Biases are a bit worse in the blue, but the different methods (Rayleigh DIMITRI, Rayleigh OSCAR, and PICS) do not agree in that spectral range: Rayleigh DIMITRI gives about 7-8%, Rayleigh OSCAR 5-6% while PICS remains around 2%. The Rayleigh method is however suspected to underestimate the simulated signal in the blue region whatever the sensor and the implementation, and in addition shows a much higher year-to-year variability, so that the 2-3% estimate of the PICS method is likely more reliable. Results for 1020 nm are much worse (6 to 8%, depending on the reference band).

The same figures for OLCI-B show current performance within the 2% requirement for all bands from 510 nm (Oa04) to 940 nm (Oa20) with remarkable agreement for all methods but DIMITRI Rayleigh. As for OLCI-A, the two Rayleigh methods indicate excess of brightness for the 4 bluest channels, between 2 and 5 %, while the PICS results provide very good performance estimates. Here again, results for 1020 nm are much worse (4 to 6%, depending on the reference band).
5.2.3.1 DIMITRI results

The time-series from the PICS method over the operational products display a good consistency over all the used CalVal sites (Figure 44 and Figure 45) and highlights a good stability of both sensors (OLCI-A and OLCI-B) over the analysed period.

Figure 43: comparison of OSCAR and DIMITRI results for the various methods.
Figure 44: Time-series of the elementary ratios (observed/simulated) signal from S3A/OLCI for (top to bottom) Band Oa03 and band Oa18 respectively, over Six PICS Cal/Val sites. Dashed-green and orange lines indicate the 2% and 5% biases respectively. Error bars indicate the desert methodology uncertainty.
Figure 45: Time-series of the elementary ratios (observed/simulated) signal from S3B/OLCI for (top to bottom) Band Oa03 and band Oa18 respectively over Six PICS Cal/Val sites. Dashed-green and orange lines indicate the 2% and 5% biases respectively. Error bars indicate the desert methodology uncertainty.

The synthesis of the results shows a good consistency over Rayleigh, Glint and PICS methods (Table 6, Table 7 and Figure 46) from OLCI-A and OLCI-B over the period January - December 2022.
Table 6: Synthesis of the DIMITRI results: estimated gain values for S3A/OLCI from Glint, Rayleigh and PICS over the period January 2022-December 2022.

<table>
<thead>
<tr>
<th>S3A-OLCI Bands</th>
<th>Rayleigh Over Jan’22-Dec’22</th>
<th>Glint Over Jan’22-Dec’22</th>
<th>PICS Over Jan’22-Dec’22</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Waveleength (nm)</td>
<td>Rayleigh Gain Coefficient</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Oa01</td>
<td>400</td>
<td>1.073</td>
<td>0.029</td>
</tr>
<tr>
<td>Oa02</td>
<td>412</td>
<td>1.082</td>
<td>0.030</td>
</tr>
<tr>
<td>Oa03</td>
<td>443</td>
<td>1.074</td>
<td>0.027</td>
</tr>
<tr>
<td>Oa04</td>
<td>490</td>
<td>1.081</td>
<td>0.027</td>
</tr>
<tr>
<td>Oa05</td>
<td>510</td>
<td>1.057</td>
<td>0.022</td>
</tr>
<tr>
<td>Oa06</td>
<td>560</td>
<td>1.040</td>
<td>0.020</td>
</tr>
<tr>
<td>Oa07</td>
<td>620</td>
<td>1.034</td>
<td>0.015</td>
</tr>
<tr>
<td>Oa08</td>
<td>665</td>
<td>1.036</td>
<td>0.014</td>
</tr>
<tr>
<td>Oa09</td>
<td>674</td>
<td>1.042</td>
<td>0.014</td>
</tr>
<tr>
<td>Oa10</td>
<td>681</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa11</td>
<td>709</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa12</td>
<td>754</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa13</td>
<td>761</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa14</td>
<td>764</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa15</td>
<td>768</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa16</td>
<td>779</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa17</td>
<td>865</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa18</td>
<td>885</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa19</td>
<td>900</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa20</td>
<td>940</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa21</td>
<td>1020</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 7: Synthesis of the DIMITRI results: estimated gain values for S3B/OLCI from Glint. Rayleigh and PICS over the period January 2022- December 2022.

<table>
<thead>
<tr>
<th>S3B-OLCI Bands</th>
<th>Wave length (nm)</th>
<th>Rayleigh Gain Coefficient</th>
<th>Standard deviation</th>
<th>Glint Gain Coefficient</th>
<th>Standard deviation</th>
<th>PICS Gain Coefficient</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oa01</td>
<td>400</td>
<td>1.035</td>
<td>0.032</td>
<td>NA</td>
<td>NA</td>
<td>1.016</td>
<td>0.047</td>
</tr>
<tr>
<td>Oa02</td>
<td>412</td>
<td>1.048</td>
<td>0.035</td>
<td>NA</td>
<td>NA</td>
<td>0.997</td>
<td>0.025</td>
</tr>
<tr>
<td>Oa03</td>
<td>443</td>
<td>1.036</td>
<td>0.031</td>
<td>NA</td>
<td>NA</td>
<td>1.002</td>
<td>0.033</td>
</tr>
<tr>
<td>Oa04</td>
<td>490</td>
<td>1.046</td>
<td>0.032</td>
<td>NA</td>
<td>NA</td>
<td>1.004</td>
<td>0.046</td>
</tr>
<tr>
<td>Oa05</td>
<td>510</td>
<td>1.042</td>
<td>0.030</td>
<td>1.019</td>
<td>0.019</td>
<td>1.010</td>
<td>0.042</td>
</tr>
<tr>
<td>Oa06</td>
<td>560</td>
<td>1.035</td>
<td>0.027</td>
<td>1.012</td>
<td>0.010</td>
<td>1.007</td>
<td>0.036</td>
</tr>
<tr>
<td>Oa07</td>
<td>620</td>
<td>1.030</td>
<td>0.025</td>
<td>1.009</td>
<td>0.003</td>
<td>1.004</td>
<td>0.020</td>
</tr>
<tr>
<td>Oa08</td>
<td>665</td>
<td>1.027</td>
<td>0.023</td>
<td>1.012</td>
<td>0.000</td>
<td>1.012</td>
<td>0.024</td>
</tr>
<tr>
<td>Oa09</td>
<td>674</td>
<td>1.030</td>
<td>0.024</td>
<td>1.018</td>
<td>0.002</td>
<td>1.011</td>
<td>0.018</td>
</tr>
<tr>
<td>Oa10</td>
<td>681</td>
<td>NA</td>
<td>NA</td>
<td>1.016</td>
<td>0.002</td>
<td>1.012</td>
<td>0.020</td>
</tr>
<tr>
<td>Oa11</td>
<td>709</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa12</td>
<td>754</td>
<td>NA</td>
<td>NA</td>
<td>1.012</td>
<td>0.008</td>
<td>1.010</td>
<td>0.018</td>
</tr>
<tr>
<td>Oa13</td>
<td>761</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa14</td>
<td>764</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa15</td>
<td>768</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa16</td>
<td>779</td>
<td>NA</td>
<td>NA</td>
<td>1.000</td>
<td>0.011</td>
<td>1.007</td>
<td>0.018</td>
</tr>
<tr>
<td>Oa17</td>
<td>865</td>
<td>NA</td>
<td>NA</td>
<td>1.006</td>
<td>0.018</td>
<td>1.010</td>
<td>0.017</td>
</tr>
<tr>
<td>Oa18</td>
<td>885</td>
<td>NA</td>
<td>NA</td>
<td>1.000</td>
<td>0.021</td>
<td>1.014</td>
<td>0.016</td>
</tr>
<tr>
<td>Oa19</td>
<td>900</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.992</td>
<td>0.029</td>
</tr>
<tr>
<td>Oa20</td>
<td>940</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa21</td>
<td>1020</td>
<td>NA</td>
<td>NA</td>
<td>1.054</td>
<td>0.037</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Figure 46: The estimated gain values for (top) S3A/OLCI and (bottom) S3B/OLCI from Glint, Rayleigh and PICS methods as a function of wavelength. We use the gain value of Oa8 from PICS method as reference gain for Sunglint method. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the methods uncertainties.

Cross-mission Intercomparison with MSI-A, MSI-B has been performed until November 2022, and with SLSTR-A and SLSTR-B until December 2022. Figure 47 shows the estimated gain over different time-series for different sensors over PICS. The spectral bands with significant absorption from water vapor and O2 are excluded. OLCI-A seems to have higher gain wrt the other sensors, and about 1-2% higher gain wrt to OLCI-B over VNIR spectral range.
Figure 47: Ratio of observed TOA reflectance to simulated one for (black) MERIS, (pale-green) S2A/MSI, (white) S2B/MSI, (blue) S3A/OLCI, (green) S3B/OLCI, (red) S3A/SLSTR-NADIR, and (cyan) S3B/SLSTR-NADIR averaged over the six PICS test sites as a function of wavelength.

5.2.3.2 OSCAR results

The OSCAR Rayleigh and Glint methods have been applied to the S3A and S3B S3ETRAC data from the 6 oceanic calibration sites listed in Table 8. The OSCAR Rayleigh method has been improved by the use of a new chlorophyll climatology, described below.

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Ocean</th>
<th>North Latitude</th>
<th>South Latitude</th>
<th>East Longitude</th>
<th>West Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>PacSE</td>
<td>South-East of Pacific</td>
<td>-20.7</td>
<td>-44.9</td>
<td>-89</td>
<td>-130.2</td>
</tr>
<tr>
<td>PacNW</td>
<td>North-West of Pacific</td>
<td>22.7</td>
<td>10</td>
<td>165.6</td>
<td>139.5</td>
</tr>
<tr>
<td>PacN</td>
<td>North of Pacific</td>
<td>23.5</td>
<td>15</td>
<td>200.6</td>
<td>179.4</td>
</tr>
<tr>
<td>AtlN</td>
<td>North of Atlantic</td>
<td>27</td>
<td>17</td>
<td>-44.2</td>
<td>-62.5</td>
</tr>
<tr>
<td>AtlS</td>
<td>South of Atlantic</td>
<td>-9.9</td>
<td>-19.9</td>
<td>-11</td>
<td>-32.3</td>
</tr>
<tr>
<td>IndS</td>
<td>South of Indian</td>
<td>-21.2</td>
<td>-29.9</td>
<td>100.1</td>
<td>89.5</td>
</tr>
</tbody>
</table>
Updates to the OSCAR Rayleigh method

A new CHL climatology (Figure 48) has been derived from the CMEMS GlobColour chlorophyll products which are publicly available on the CMEMS web portal. The climatology has been derived from the CMEMS OLCI monthly CHL products with a 4 km spatial resolution (i.e., dataset-oc-glo-chl-olci-a-l4-av_4km_monthly-rt-v02 products) considering the years 2017, 2018 and 2019. To impact of the climatology on the OSCAR Rayleigh results was assessed by reprocessing S3ETRAC data from the year 2019 with the new climatology. Overall, the new OLCI derived CHL climatology had a small effect on the OSCAR Rayleigh results with a slight decrease in the calibration results (i.e. smaller bias between modelled and measured values).

![Figure 48. New CHL climatology for the Rayleigh calibration sites based CMEMS GlobColour products.](image)

OSCAR Rayleigh results

The OSCAR Rayleigh have been applied to the OLCI-A and OLCI-B S3ETRAC data from the 6 oceanic calibration sites using the new chlorophyll climatology. In Figure 49, the average OSCAR OLCI-A and OLCI-B Rayleigh results for the year 2022 are given. This average is obtained from 344 OLCI-A and 362 OLCI-B scenes from 2022 with valid results. A bias is observed between OLCI-A and OLCI-B, with OLCI-A being about 2 % brighter than OLCI-B in blue bands (i.e. Oa1 to Oa3). This bias seems to decrease with wavelength to about 1% in green bands and about 0.7% in red bands.

In Figure 50 and Figure 51 the average results of 2022 are compared with the average results of 2021 for respectively OLCI-A and OLCI-B. The results are very consistent between the years, both for OLCI-A and OLCI-B.
Figure 49: OSCAR Rayleigh S3A and S3B Calibration results for 2022 as a function of wavelength.

Figure 50: OSCAR Rayleigh OLCI-A Calibration results 2021 and 2022 as a function of wavelength.
Figure 51: OSCAR Rayleigh OLCI-B Calibration results for 2021 and 2022 as a function of wavelength.

OSCAR Glitter results

In Figure 52, the average OSCAR OLCI-A and OLCI-B Glitter results, adapted to the Rayleigh result at 665 nm, are given for the year 2022. Similarly, as for the Rayleigh results, a bias is observed between OLCI-A and OLCI-B, with OLCI-A being slightly brighter than OLCI-B with a bias decreasing with wavelength. Inter-band differences are small (< 1%) and well within the requirements except for the bands Oa21 (i.e. 1020 nm) and Oa4.
Synthesis OSCAR Results

The synthesis of the OSCAR results is given in Table 9 below. This table shows a good consistency between the Rayleigh and Glitter results.

Figure S2: OSCAR Glitter S3A and S3B Calibration results (adapted to Rayleigh result at 665 nm) for 2022 as a function of wavelength.
## Table 9. OSCAR Rayleigh and Glitter calibration results for S3A and S3B for 2022

<table>
<thead>
<tr>
<th>OLCI band</th>
<th>Wavelength (nm)</th>
<th>Oscar Rayleigh S3A 2022</th>
<th>Oscar Glint S3A 2022</th>
<th>Oscar Rayleigh S3B 2022</th>
<th>Oscar Glint S3B 2022</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>avg</td>
<td>stdev</td>
<td>avg</td>
<td>stdev</td>
<td>avg</td>
</tr>
<tr>
<td>Oa01</td>
<td>400</td>
<td>1.049</td>
<td>0.029</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa02</td>
<td>412</td>
<td>1.059</td>
<td>0.031</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa03</td>
<td>443</td>
<td>1.051</td>
<td>0.028</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa04</td>
<td>490</td>
<td>1.048</td>
<td>0.017</td>
<td>1.041</td>
<td>0.007</td>
</tr>
<tr>
<td>Oa05</td>
<td>510</td>
<td>1.026</td>
<td>0.010</td>
<td>1.020</td>
<td>0.005</td>
</tr>
<tr>
<td>Oa06</td>
<td>560</td>
<td>1.016</td>
<td>0.009</td>
<td>1.013</td>
<td>0.003</td>
</tr>
<tr>
<td>Oa07</td>
<td>620</td>
<td>1.012</td>
<td>0.006</td>
<td>1.009</td>
<td>0.002</td>
</tr>
<tr>
<td>Oa08</td>
<td>665</td>
<td>1.016</td>
<td>0.005</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa09</td>
<td>674</td>
<td>1.018</td>
<td>0.005</td>
<td>1.019</td>
<td>0.001</td>
</tr>
<tr>
<td>Oa10</td>
<td>681</td>
<td>1.015</td>
<td>0.005</td>
<td>1.017</td>
<td>0.001</td>
</tr>
<tr>
<td>Oa11</td>
<td>709</td>
<td>0.999</td>
<td>0.007</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa12</td>
<td>754</td>
<td>1.010</td>
<td>0.002</td>
<td>1.015</td>
<td>0.002</td>
</tr>
<tr>
<td>Oa13</td>
<td>761.25</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa14</td>
<td>764.375</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa15</td>
<td>767.5</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa16</td>
<td>778.75</td>
<td>NA</td>
<td>NA</td>
<td>1.004</td>
<td>0.002</td>
</tr>
<tr>
<td>Oa17</td>
<td>865</td>
<td>NA</td>
<td>NA</td>
<td>1.012</td>
<td>0.004</td>
</tr>
<tr>
<td>Oa18</td>
<td>885</td>
<td>NA</td>
<td>NA</td>
<td>1.008</td>
<td>0.005</td>
</tr>
<tr>
<td>Oa19</td>
<td>900</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa20</td>
<td>940</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Oa21</td>
<td>1020</td>
<td>NA</td>
<td>NA</td>
<td>1.038</td>
<td>0.005</td>
</tr>
</tbody>
</table>

*OSCAR Rayleigh results for band Oa01 have to be considered with care due to larger uncertainty in the radiative transfer calculation*
5.3 L2 product performances

5.3.1 Level 2 cloud screening

5.3.1.1 Introduction

Providing clear sky conditions for production of Sentinel-3 OLCI Level 2 products is essential to ensure a good and reliable Level 2 product quality for the users. After issues with the cloud screening in the initial processing baseline, a big effort was made by the former Sentinel-3 MPC to improve the level 2 cloud flagging algorithms. Since 2017 a new cloud flagging is implemented in the current operational and reprocessed products. This had been extensively validated during 2019 and is documented in the Annual Report 2018. In brief summary, the overall accuracy is 86%, and the user accuracy for clear sky conditions which is the most relevant criteria for users, is 92.1%.

The work in the year 2022 focussed on improving the quality control of the cloud screening by introducing an automated validation approach and improving the limitations which are still in the cloud screening. The achievements are summarized below. Since these improvements do not depend on the platform (S-3A or B) we do not differentiate.

5.3.1.2 Cloud screening neural network (NN)

The development of a clear NN was on-going throughout 2022 but with a bit reduced effort, since more focus with set on first improving the current NN structure by retraining with a new training dataset. The latter is still on-going and hopefully will provide an updated NN in 2023. The effort of updating the old NN structure was made, due to identified semi-systematic errors along coastlines caused by the NN. The error leads to coastlines being flagged as cloud quite regularly but not always. Since implementing a new NN structure in the IPF is more complex than replacing the current NN by a newly trained one, the latter route was decided to be followed.

5.3.1.3 Cloud Shadow

Developments 2022:

During 2022 it was concluded within the MPC that implementing the cloud shadow operator into the current IPF is very complex. Due to spatial requirements of the algorithm for neighbouring pixels that exceed the currently available area during processing in the IPF, an implementation is not straightforward and would require bigger adaptations of the processing framework. Since a restructuring of the whole IPF framework is planned in the future, the implementation of the cloud shadow was postponed after the change.

5.3.1.4 Quality assessment of the L2 cloud screening using ground-based sky camera observations.

For the OPT-MPC a more systematic QC of the cloud masking was proposed, to better fulfil the needs of the DQR, APR and in general for a good data quality monitoring. One of the approaches besides the prior used PixBox (expert pixel collection) and visual scene analysis, was Quality Control (QC) using ground-based sky cameras.

Starting in July 2022 this approach was used on a monthly basis to report the cloud screening accuracy and identify potential anomalies. In this section we will give a brief overview of the methodology and then show results for the reporting period (2022).
Note: The sky camera approach is still under on-going further development and thus is referred to as “prototype”.

5.3.1.4.1 Sky Camera based validation approach.

In the aftermath of the Cloud Mask Intercomparison eXercise (CMIX), the idea of a network of low-cost ground-based sky cameras (stereo) was born, to provide an objective and sensor independent source for cloud mask validation. The sky camera design was developed by NASA GSFC and University of Maryland. Within the framework of ESA’s Quality assurance framework for earth observation (QA4EO), a pair of sky cameras (SC) was installed at La Sapienza University in Rome. The objective of the project was to analyze the usage of ground-based sky cameras, as an independent validation source for satellite cloud masking algorithms. The scope of this work was to prototype algorithms and methods to process sky camera data and compare them with satellite-based cloud masks.

5.3.1.4.1.1 INSTRUMENTATION SETUP:

❖ A set of two cameras (stereo pair) was setup at La Sapienza University in Rome (see Figure 53).
❖ The cameras use a Raspberry Pi 4 and the Omnivision OV5647 sensor. The field of view is 194 (horizontal) and 142 (vertical). Distance between cameras is around 260 meters. Currently, the cameras are collecting data every minute between 08:00 and 14:00 UTC.

5.3.1.4.1.2 PRE-PROCESSING OF THE SC DATA:

Pre-processing of sky camera data is needed to better match the satellite observations

- **Crop**: Reduce geometric distortion (increasing outside of the centre).
- **Rotate**: The SCs are installed looking a bit northwest.
- **Flip**: The SC is looking from the ground up and the satellite does the opposite.

5.3.1.4.1.3 AUTOMATED CLASSIFICATION OF THE SC DATA:

Since the SC data cannot be classified manually in an operational setup, a classification method for the SC data was needed. Two Random Forest classifiers (one for each SC) have been trained and validated against manual classifications.

Validation of the RF classifier shows high accuracy between 93% and 96% Overall Accuracy (OA), as shown in Figure 54 and Figure 55.

![Sky Camera 1 manual classification vs. SkyCam 1 auto classification](image)

Scotts PI: 0.868
Krippendorfs alpha: 0.869
Cohens kappa: 0.868

*Figure 54: Validation results of RF classifier for SC1*
5.3.1.4.1.4 TEST VALIDATION FOR OLCI – AUTOMATIC SC CLASSIFICATION

A test was made with all Sentinel-3 OLCI L2 data between 01.01.2021 and 31.12.2021. 282 matchups have been identified between the SC sites and S3 OLCI overpasses. One OLCI pixel over SC site is validated against cloud fraction in a defined window of SC. The SC cloud is defined as 50% cloud cover in 500x500 pixel window. This definition was chosen as a starting point and needs to be adapted. The results (Figure 56) show comparably low OA of 78%. This is due to the skewed reference data (189 clear vs. 93 cloud). Calculation of balanced overall accuracy (BOA) can correct for this and leads to a BOA of 82%.

Figure 55: Validation results of RF classifier for SC2
Optical MPC
Sentinel-3 optical Annual Performance Report
Year 2022

Ref.: OMPC.ACR.APR.002
Issue: 1.1
Date: 17/03/2023
Page: 65

Sky Camera validation over Rome
Using 2021 LFR data

SC 1 automatic classification vs. OLCI L2 LFR Cloud & Ambiguous & Margin
Sky Camera 1

<table>
<thead>
<tr>
<th>Class</th>
<th>Clear</th>
<th>Cloud</th>
<th>Sum</th>
<th>U A</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEAR</td>
<td>136</td>
<td>7</td>
<td>143</td>
<td>95.1</td>
<td>4.9</td>
</tr>
<tr>
<td>CLOUD</td>
<td>53</td>
<td>86</td>
<td>139</td>
<td>61.9</td>
<td>38.1</td>
</tr>
<tr>
<td>Sum</td>
<td>189</td>
<td>93</td>
<td>282</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PA: 72.0
EA: 28.0
OA: 78.72
BOA: 82.25

Scotts Pi: 0.56
Krippendorfs alpha: 0.561
Cohens kappa: 0.572

Figure 56: 2021 test validation results for OLCI FR using reference data from SC1

These numbers are quite comparable with the validation results of OLCI PixBox validation (2021) over land surfaces, as shown in Figure 57.

Sky Camera validation over Rome
Using 2021 LFR data

PixBox validation 2021
using 2018 data over land surfaces

OLCI A+B FR IdePix cloud val. - land surfaces
In-Situ Database

<table>
<thead>
<tr>
<th>Class</th>
<th>Clear</th>
<th>Cloud</th>
<th>Sum</th>
<th>U A</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEAR</td>
<td>3442</td>
<td>443</td>
<td>3885</td>
<td>88.6</td>
<td>11.4</td>
</tr>
<tr>
<td>CLOUD</td>
<td>1039</td>
<td>3183</td>
<td>4232</td>
<td>75.4</td>
<td>24.6</td>
</tr>
<tr>
<td>Sum</td>
<td>4481</td>
<td>3626</td>
<td>8107</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PA: 76.6
EA: 23.2
OA: 81.72
BOA: 82.3

Scotts Pi: 0.634
Krippendorfs alpha: 0.634
Cohens kappa: 0.635

Figure 57: Comparison between PixBox and Sky Camera validation

Again, the skewed distribution of SC classification hinders the comparison a bit. Nevertheless, the BOA is quite comparable.
5.3.1.4.2 Sky Camera based validation – prototype results the year 2022

Figure 58 and Figure 59 show the prototype validation results for the year 2022 using sky camera 1 (Marconi) over the Rome site. Comparing the results shown in Figure 58, including the cloud margin with those of 2021, shown in Figure 57, lead to very similar results. With a difference on OA of below 0.5%. This shows that the method is very stable.

![Confusion Matrix](image)

**Figure 58: Confusion matrix showing validation results for OLCI L2 cloud screening including margin against SC1 automated classification for the year 2022.**

As we have shown in the monthly DQRs, including the margin in the analysis can often be a bit misleading and thus exclusion of the margin for the analysis is preferred. Especially since the margin is a geometric feature and no spectral cloud identification and thus is more a precaution measure for stricter cloud masking requirements of users.

When removing the margin from the analysis, we can compare the results (see Figure 59) to those of the PixBox collection made in 2021 using 2018 data (see Figure 57). Now the results are also quite comparable. All these analysis show that the cloud screening is very stable over the years and stable throughout the reporting period.
### Figure 59: Confusion matrix showing validation results for OLCI L2 cloud screening excluding margin against SC1 automated classification for the year 2022.

<table>
<thead>
<tr>
<th>Class</th>
<th>Clear</th>
<th>Cloud</th>
<th>Sum</th>
<th>U</th>
<th>A</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEAR</td>
<td>250</td>
<td>22</td>
<td>272</td>
<td>91.9</td>
<td>8.1</td>
<td></td>
</tr>
<tr>
<td>CLOUD</td>
<td>36</td>
<td>87</td>
<td>123</td>
<td>70.7</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>286</td>
<td>109</td>
<td>395</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scott's Pi: 0.646  
Krippendorf's alpha: 0.646  
Cohen's kappa: 0.646
Figure 60: Validation results for OLCI A+B separated between harmonized and non-harmonized data, as well as all surfaces, as well as land and water surfaces
**Figure 61: Validation results for OLCI A separated between harmonized and non-harmonized data, as well as all surfaces, as well as land and water surfaces**
Figure 62: Validation results for OLCI B separated between harmonized and non-harmonized data, as well as all surfaces, as well as land and water surfaces
The following can be concluded:

The analysis has shown that the OLCI-A trained cloud screening NN can be used without any constraint for OLCI-B. It has also shown that applying flat-fielding and harmonization before using the NN is not beneficial, but disadvantageous. Therefore, applying FFH is not recommended.

It was also shown that the cloud screening NN works slightly better for OLCI-B compared to OLCI-A and performance over water surfaces is a lot better compared to land surfaces.

Roughly 2/3 of the OLCI-A products have been acquired before the improvements of the radiometric gain model. This might have had an influence on the NN performance for OLCI-A and could explain the better performance of OLCI-B. Unfortunately, the dataset was too small to make robust analysis.

5.3.2 Integrated Water Vapour (IWV)

The OLCI L2 IWV processor distinguishes between ocean and land surfaces and works very differently above the respective surfaces. The algorithm above water shows some serious flaws and is therefore not further investigated. A new version is under development, that will be integrated into an atmospheric branch. However, despite of a small systematic overestimation, the water vapour above land works very well and stable, and it is a good indicator for a system monitoring.

OLCI’s IWV above land surface is validated using the following ground truth data:

1. Global GNSS data, with a focus to north America (SUOMI NET, Ware et al. 2000)
3. GRUAN radiosonde observations IWV (Immler et al 2010, Bodeker 2015). In 2022 we discontinued the comparisons, since the number of matchups is very low (just a few).
4. AERONET (Holben et al 1998), using atmospheric transmission measurements at 0.9 µm.

All L2 product types have been validated: full resolution and reduced resolution, near real time and non time critical, Ocean Colour (wrr, wfr) and Land Colour (lrr, lfr). The found results for all product types are identical, as expected, since the used processor is the same. The following quantitative comparisons are hence restricted to wrr NT (Ocean Colour Product, reduced resolution, non time critical). Since the ocean colour product and the land colour product provide water vapour above land and water surfaces, the comparison is comprehensive. OLCI A data partly belong to reprocessed data if processed before Nov/2017. The ocean colour products from OLCI A have been taken from Eumetsats’ rolling archive CODA (Copernicus Online Data Access, https://coda.eumetsat.int/#/home) or reprocessed OLCI A CODAREP (https://codarep.eumetsat.int/#/home) websites. All OLCI B data is from Eumetsats’ CODA. From April 2022 the data acquisition from EUMETSAT CODA was partly malfunctioning (the lat lon based orbit selection did not work for RR orbits) and discontinued in September 2022. Since October 2022 we are using EUMETSATs datastore (collection id: EO:EUM:DAT:0410). Data from Apr 2022 on belongs to that data source.
5.3.2.1 Integrated water vapour above land

5.3.2.1.1 Validation of OLCI A IWV using GNSS

700,000 potential matchups within the period of June 2016 to January 2023 have been analysed. The scenes cover high and low elevations, however, the majority of the used SUOMI-NET ground stations are in North and Central America. Only OLCI measurements are taken for the validation which are above land and are cloud-free in an area of about 10x10 km$^2$ around the GNSS stations. For the cloud detection, the standard L2 cloud-mask has been applied (including the cloud ambiguous and cloud margin flags). The comparison of OLCI and GNSS shows a very high agreement (Figure 63). The correlation between both quantities is 0.98. The root-mean-squared-difference is 2.1 kg/m$^2$. The systematic overestimation by OLCI is 12%. The bias corrected rmsd is 1.3 kg/m$^2$. Interesting is the strong seasonal pattern of the bias. It is also partly visible in the systematic overestimation swinging between 5 and 12%. This clearly belongs to the seasonality of water vapour in North America, with lower (better) values during winter. This could be an artefact of the retrieval inherent spectral extrapolation of the surface reflectance from window bands to the absorption band, but also a deficit in the cloud detection during winter, that by itself would produce a dry bias and thus may lead to an increased wet bias.
Figure 63: Upper left: Scatter plot of the IWV products, derived from OLCI A above land and from SUOMI NET GNSS measurements. Upper right: Histogram of the difference between OLCI and GNSS (blue: original OLCI, orange: bias corrected OLCI). Lower left: Temporal evolution of different quality measures (from top to bottom: systematic deviation factor, bias, root mean squared difference (with and without bias correction), explained variance (number in boxes are the numbers of matchups)). Lower right: Positions of the GNSS stations (grey: no valid matchup).

5.3.2.1.2 Validation of OLCI A IWV using passive microwave radiometer at ARM sites

Microwave radiometer measurements at the Atmospheric Radiation Measurement (ARM) Climate Research Facility of the US Department of Energy provides the ground truth with the highest accuracy (0.6 to 0.8 kg/m²). Currently 3 ARM sites are operated continuously, only the SGP (southern great plains) site provided a significant amount of cloud free measurements. 4000 potential matchups within the period of June 2016 to January 2023 have been analysed. Only OLCI measurements are taken for the validation which are above land and are cloud-free in an area of about 10x10 km² around SGP.

Since 2021 we are using the mwrls (microwave line of sight retrieval): sgpmwrlsC1.b1 as the reference water vapor data (and we have reprocessed the full matchup data base starting from 2016/18 to guarantee temporal stability). mwrls provides ground truth with a high precision of (0.8 kg/m²).
For the cloud detection, the standard L2 cloud-mask has been applied (including the cloud ambiguous and cloud margin flags), resulting in 259 valid matchups. The comparison shows a very high agreement (Figure 64). The correlation between both quantities is 0.99. The root-mean-squared-difference is 1.4 kg/m². The systematic overestimation by OLCI is 9%. The bias corrected $rmsd$ is 0.9 kg/m², close to the uncertainty of the MWR. The investigation of the temporal evolution shows the same seasonal pattern as the GNSS comparisons, again belonging to the same seasonality of water vapour, snow cover and cloud cover in North America.

![Image](image_url)

**Figure 64:** Upper left: Scatter plot of the IWV products, derived from OLCI A above land and from AMR MWR. Upper right: Histogram of the difference between OLCI and ARM (blue: original OLCI, orange: bias corrected OLCI). Lower left: Temporal evolution of different quality measures (from top to bottom: systematic deviation factor, bias, root mean squared difference (with and without bias correction), explained variance (number in boxes are the numbers of matchups)). Lower right: Position of ARM SGP.
5.3.2.1.3  Validation of OLCI A IWV using AERONET observations

AERONET observations, regardless not primary made for water vapour, allow the direct estimation of the total column of water vapour by measuring the extinction of the direct solar irradiance at 900 nm. The used operational algorithm is quite simple and eventually relies on a logarithmic fit (incl. quadratic corrections). We are using AERONET for the IWV comparison, since AERONET data are much better globally distributed, than ARM and SUOMINET. We use AERONET V3 level 1.5. Level 1.5 is not stringently quality controlled, but the retrieval algorithm is exactly the same as for level 2 and the data is published without delay.

Only OLCI measurements are used for the validation which are cloud-free (according to the standard cloud flags: cloud, cloud margin and cloud ambiguous) in an area of about 10x10 km$^2$ around the AERONET acquisition. From the 257000 potential matchups within the period of June 2016 to January 2023, 47500 valid matchups could be used. (Figure 65). The correlation between both quantities is 0.97. The root-mean-square-difference is 3.5 kg/m$^2$. The systematic overestimation by OLCI is 19%. The bias corrected rmsd is 1.6 kg/m$^2$. The systematic deviation between OLCI and AERONET of 19% is significantly larger than the one found for GNSS and ARM (~10%). We think that this stems from a dry bias of AERONET and accordingly deficits in its operational algorithm, but we have not investigated it deeper.

*Figure 65: Upper left: Scatter plot of the IWV products, derived from OLCI A above land and from AERONET. Upper right: Histogram of the difference between OLCI and AERONET (blue: original OLCI, orange: bias corrected OLCI). Lower: Positions of the used AERONET stations (grey: no valid matchup).*
5.3.2.1.4 Validation of OLCI B IWV

Within the period of June 2018 to January 2023 several 100000 scenes have been analysed yet. 24000 of them are valid for SUOMI-NET CONUS ground stations in North and Central America, 154 for ARM MWR and 29500 for AERONET.

As for OLCI A, only measurements are taken for the validation which are above land and are cloud-free in an area of about 10x10 km$^2$ around the corresponding stations. For the cloud detection, the standard L2 cloud-mask has been applied (including the cloud ambiguous and cloud margin flags). The comparison for OLCI B shows almost identical results as for OLCI A (Figure 66). The same is true for the seasonal patterns, when comparing with SUOMI-NET.
5.3.2.2 Integrated water vapour above water

5.3.2.2.1 Quantitative validation using GNSS

OLCIs IWV above water surfaces has been quantitatively validated via global GNSS measurements too, however with few additional assumptions:

Since the GNSS stations are usually not directly above water, the closest water pixel (within 1km) is used for the satellite measurement.

No height correction has been applied to account for the potentially elevated GNSS station.
For OLCI-A, 83 matchups remain after filtering (Figure 67). They show a large bias 10 kg/m² and a large scatter (>6 kg/m²). For OLCI-B the number of valid matchups is smaller (49), but all indications point to similar systematic deviations and retrieval noise. This is in accordance with the visual inspection.

Figure 67: Upper: Scatter plot of the IWV products, derived from OLCI (A left, B right) above ocean and from SUOMI NET GNSS measurements. Lower: Positions of the GNSS (A: left, B: right).

5.3.2.2 Validation by AERONET IWV Retrievals – Ocean

OLCIs IWV above water surfaces has been quantitatively validated via global AERONET-OC measurements. All filters are as for land matchups. The remaining 3100 (OLCI-A) and 2000 (OLCI-B) matchups show a large bias of about 9 kg/m², a large scatter (>6 kg/m²) and a systematic overestimation of about 20% (Figure 68). This is in accordance with the visual inspection and with the GNNS matchups (Figure 67) over oceans.
5.3.2.3 Summary

The continuous validation of the OLCI A IWV over land product using 3 different sources of ground truth shows consistently, that the product is of high quality (bias corrected root mean squared distance of down to 1.5 –0.8 kg/m²) and very stable. However, there is a systematic overestimation of 9% to 13%. An equivalent validation of OLCI B shows the same results, no systematic differences between OLCI A and B have been found. The validation with Suominet shows seasonal patterns of the overestimation with better values during winter seasons.

Retrievals above ocean show an overestimation in transition zones between glint and off glint (not shown herein). This is a clear deficit of the description of the scattering-absorption interaction. Further the IWV has a large wet bias over ocean.

5.3.2.4 References


5.3.3 OLCI Global Vegetation Index (OGVI), a.k.a. GIFAPAR, and OLCI Terrestrial Chlorophyll Index (OTCI)

This section presents the performance of two Level 2 products routinely generated from OLCI: the Green Instantaneous Fraction of Absorbed Photosynthetically Available Radiation (GIFAPAR, formerly known as the OLCI Global Vegetation Index (OGVI)) and the Terrestrial Chlorophyll Index (OTCI, a proxy of canopy chlorophyll content (CCC)). The performance evaluation activities over the past year involve indirect verification efforts, direct validation against GBOV in-situ data, and an investigation into the inter-annual variability of the land products.

5.3.3.1 Indirect verification

The indirect verification consists of examining the annual evolution of OTCI and OGVI as well as the comparison with archive Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyll Index (MTCI) and MERIS Green Vegetation Index (MGVI). The MERIS archive is also referred to as MERIS climatology. The verification is carried out using 3x3 pixel extractions from >50 European Space Agency (ESA) core and Committee for Earth Observation Satellites (CEOS) Land Product Validation group (LPV) sites, including a range of latitudes and land cover types (Table 10). The sites are scattered across 15 countries and are part of existing networks. Figure 69 to Figure 72 show the results of indirect verification on five unique land cover types: cropland, deciduous broadleaved forest, deciduous shrub, broadleaved evergreen forest, and needle-leaved evergreen forest. The results cover the period up to May 2022.
performance statistics between the monthly average OLCI and MERIS land products for every site are shown in Table 11. Generally, there is a good agreement between the land products with strong R² values and biases around 0. Similar seasonal trajectories and timings are shown in the extractions from both products at the following sites reviewed: BE-Brasschaat, DE-Hainich and FR-EsteesMons (Figure 69 to Figure 71). The monthly mean extractions from all sites are shown in Figure 72. OTCI from S3A shows a strong agreement with the MERIS archive, R²=0.93, NRMSD<0.08 with a low bias, -0.02. OGVI similarly shows a strong agreement with the MERIS archive, R²= 0.93, NRMSD<0.15, with a slightly higher bias of 0.06.
Table 10: Validation sites analysed in report S3A 78/S3B 59. Land cover data from GLC2000 grouped according to the International Geosphere-Biosphere Programme (IGBP) designations.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Country</th>
<th>Network</th>
<th>Lat</th>
<th>Lon</th>
<th>Land cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU-Cape-Tribulation</td>
<td>Australia</td>
<td>TERN-SuperSites, OzFlux</td>
<td>-16.106</td>
<td>145.378</td>
<td>EBF</td>
</tr>
<tr>
<td>AU-Cumberland</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-33.615</td>
<td>150.723</td>
<td>EBF</td>
</tr>
<tr>
<td>AU-Great-Western</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-30.192</td>
<td>120.654</td>
<td>DBF</td>
</tr>
<tr>
<td>AU-Litchfield</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-13.18</td>
<td>130.79</td>
<td>DBF</td>
</tr>
<tr>
<td>AU-Robson-Creek</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-17.117</td>
<td>145.63</td>
<td>EBF</td>
</tr>
<tr>
<td>AU-Rushworth</td>
<td>Australia</td>
<td>TERN-AusCover</td>
<td>-36.753</td>
<td>144.966</td>
<td>DBF</td>
</tr>
<tr>
<td>AU-Tumbarumba</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-35.657</td>
<td>148.152</td>
<td>EBF</td>
</tr>
<tr>
<td>AU-Warra-Tall</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-43.095</td>
<td>146.654</td>
<td>EBF</td>
</tr>
<tr>
<td>AU-Watts-Creek</td>
<td>Australia</td>
<td>TERN-AusCover</td>
<td>-37.689</td>
<td>145.685</td>
<td>DBF</td>
</tr>
<tr>
<td>AU-Wombat</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-37.422</td>
<td>144.094</td>
<td>EBF</td>
</tr>
<tr>
<td>BE-Brasschaat</td>
<td>Belgium</td>
<td>ICOS</td>
<td>51.308</td>
<td>4.52</td>
<td>ENF</td>
</tr>
<tr>
<td>BE-Vielsalm</td>
<td>Belgium</td>
<td>ICOS</td>
<td>50.305</td>
<td>5.998</td>
<td>ENF</td>
</tr>
<tr>
<td>BR-Mata-Seca</td>
<td>Brazil</td>
<td>ENVIRONET</td>
<td>-14.88</td>
<td>-43.973</td>
<td>non-forest</td>
</tr>
<tr>
<td>CA-Mer-Bleue</td>
<td>Canada</td>
<td>National Capitol Comission</td>
<td>45.0</td>
<td>-75.493</td>
<td>non-forest</td>
</tr>
<tr>
<td>CR-Santa-Rosa</td>
<td>Costa Rica</td>
<td>ENVIRONET</td>
<td>10.842</td>
<td>-85.616</td>
<td>EBF</td>
</tr>
<tr>
<td>CZ-Bili-Kriz</td>
<td>Czechia</td>
<td>ICOS</td>
<td>49.502</td>
<td>18.537</td>
<td>EBF</td>
</tr>
<tr>
<td>DE-Haininch</td>
<td>Deutschland</td>
<td>ICOS Associated</td>
<td>51.079</td>
<td>10.453</td>
<td>DBF</td>
</tr>
<tr>
<td>DE-Hones-Holz</td>
<td>Deutschland</td>
<td>ICOS</td>
<td>52.085</td>
<td>11.222</td>
<td>DBF</td>
</tr>
<tr>
<td>DE-Selhausen</td>
<td>Deutschland</td>
<td>ICOS</td>
<td>50.866</td>
<td>6.447</td>
<td>cultivated</td>
</tr>
<tr>
<td>DE-Tharandt</td>
<td>Deutschland</td>
<td>ICOS</td>
<td>50.964</td>
<td>13.567</td>
<td>ENF</td>
</tr>
<tr>
<td>FR-Aurade</td>
<td>France</td>
<td>ICOS</td>
<td>43.55</td>
<td>1.106</td>
<td>cultivated</td>
</tr>
<tr>
<td>FR-Estrees-Mons</td>
<td>France</td>
<td>ICOS Associated</td>
<td>49.872</td>
<td>3.021</td>
<td>DBF</td>
</tr>
<tr>
<td>FR-Guayaflux</td>
<td>France</td>
<td>ICOS Associated</td>
<td>5.279</td>
<td>-52.925</td>
<td>DBF</td>
</tr>
<tr>
<td>FR-Hesse</td>
<td>France</td>
<td>ICOS</td>
<td>48.674</td>
<td>7.065</td>
<td>DBF</td>
</tr>
<tr>
<td>FR-Montiers</td>
<td>France</td>
<td>ICOS</td>
<td>48.538</td>
<td>5.312</td>
<td>DBF</td>
</tr>
<tr>
<td>FR-Puechabon</td>
<td>France</td>
<td>ICOS</td>
<td>43.741</td>
<td>3.596</td>
<td>ENF</td>
</tr>
<tr>
<td>IT-Casterporziano2</td>
<td>Italy</td>
<td>ICOS</td>
<td>41.704267</td>
<td>12.357293</td>
<td>DBF</td>
</tr>
<tr>
<td>IT-Collelongo</td>
<td>Italy</td>
<td>EFDC</td>
<td>41.849</td>
<td>13.588</td>
<td>DBF</td>
</tr>
<tr>
<td>IT-Lison</td>
<td>Italy</td>
<td>ICOS</td>
<td>45.74</td>
<td>12.75</td>
<td>cultivated</td>
</tr>
<tr>
<td>NE-Loobos</td>
<td>Netherlands</td>
<td>ICOS Associated</td>
<td>52.166</td>
<td>5.744</td>
<td>ENF</td>
</tr>
<tr>
<td>SE-Dahra</td>
<td>Senegal</td>
<td>KIT / UC</td>
<td>15.4</td>
<td>-15.4</td>
<td>cultivated</td>
</tr>
<tr>
<td>UK-Wytham-Woods</td>
<td>United Kingdom</td>
<td>ForestGeo - NPL</td>
<td>51.774</td>
<td>-1.338</td>
<td>DBF</td>
</tr>
<tr>
<td>US-Bartlett</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>44.064</td>
<td>-71.287</td>
<td>DBF</td>
</tr>
<tr>
<td>US-Central-Plains</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>40.816</td>
<td>-104.746</td>
<td>non-forest</td>
</tr>
<tr>
<td>US-Harvard</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>42.537</td>
<td>-72.173</td>
<td>DBF</td>
</tr>
<tr>
<td>US-Moab-Site</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>38.248</td>
<td>-109.388</td>
<td>non-forest</td>
</tr>
<tr>
<td>US-Mountain-Lake</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>37.378</td>
<td>-80.525</td>
<td>DBF</td>
</tr>
<tr>
<td>US-Oak-Rige</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>35.964</td>
<td>-84.283</td>
<td>DBF</td>
</tr>
<tr>
<td>US-Ordway-Swisher</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>29.689</td>
<td>-81.993</td>
<td>ENF</td>
</tr>
<tr>
<td>US-Smithsonian</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>38.893</td>
<td>-78.14</td>
<td>DBF</td>
</tr>
<tr>
<td>US-Steigerwald</td>
<td>United States</td>
<td>NEON</td>
<td>45.509</td>
<td>-89.586</td>
<td>DBF</td>
</tr>
<tr>
<td>US-Talladega</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>32.95</td>
<td>-87.393</td>
<td>ENF</td>
</tr>
</tbody>
</table>
Table 11: Comparison statistics between monthly S3A/B OLCI land products and MERIS archive data.

<table>
<thead>
<tr>
<th>Site Acronym</th>
<th>S3A</th>
<th></th>
<th></th>
<th></th>
<th>S3B</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OTCI vs MTCI</td>
<td></td>
<td></td>
<td></td>
<td>OTCI vs MTCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>R2</td>
<td>NRMSD</td>
<td>Bias</td>
<td></td>
<td>n</td>
<td>R2</td>
<td>NRMSD</td>
</tr>
<tr>
<td>AU-Cape-Tribulation</td>
<td>12</td>
<td>0.82</td>
<td>0.04</td>
<td>-0.11</td>
<td>12</td>
<td>0.29</td>
<td>0.04</td>
<td>0.16</td>
</tr>
<tr>
<td>AU-Cumberland</td>
<td>12</td>
<td>0.82</td>
<td>0.02</td>
<td>-0.01</td>
<td>12</td>
<td>0.64</td>
<td>0.07</td>
<td>0.09</td>
</tr>
<tr>
<td>AU-Great-Western</td>
<td>12</td>
<td>0.96</td>
<td>0.02</td>
<td>0.1</td>
<td>12</td>
<td>0.52</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>AU-Robson-Creek</td>
<td>12</td>
<td>0.96</td>
<td>0.04</td>
<td>-0.04</td>
<td>12</td>
<td>0.9</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>AU-Rushworth</td>
<td>12</td>
<td>0.83</td>
<td>0.04</td>
<td>0.13</td>
<td>12</td>
<td>0.56</td>
<td>0.04</td>
<td>0.09</td>
</tr>
<tr>
<td>AU-Tumbalumba</td>
<td>12</td>
<td>0.89</td>
<td>0.04</td>
<td>0.21</td>
<td>12</td>
<td>0.52</td>
<td>0.06</td>
<td>0.09</td>
</tr>
<tr>
<td>AU-Warram-Tall</td>
<td>12</td>
<td>0.77</td>
<td>0.06</td>
<td>-0.06</td>
<td>12</td>
<td>0.26</td>
<td>0.17</td>
<td>0.04</td>
</tr>
<tr>
<td>AU-Watts-Creek</td>
<td>12</td>
<td>0.69</td>
<td>0.05</td>
<td>0.03</td>
<td>12</td>
<td>0.03</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>AU-Wombat</td>
<td>12</td>
<td>0.91</td>
<td>0.03</td>
<td>0.11</td>
<td>12</td>
<td>0.49</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>BE-Brasschaat</td>
<td>11</td>
<td>0.99</td>
<td>0.03</td>
<td>-0.08</td>
<td>11</td>
<td>0.96</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>BE-Vielcarm</td>
<td>3</td>
<td>1</td>
<td>0.03</td>
<td>-0.33</td>
<td>3</td>
<td>0.08</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>CA-Mer-Bluie</td>
<td>10</td>
<td>0.95</td>
<td>0.05</td>
<td>0.01</td>
<td>10</td>
<td>0.99</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td>CZ-Bili-Kriz</td>
<td>10</td>
<td>0.95</td>
<td>0.02</td>
<td>0.04</td>
<td>10</td>
<td>0.96</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>DE-Hainich</td>
<td>10</td>
<td>0.99</td>
<td>0.07</td>
<td>-0.04</td>
<td>10</td>
<td>0.58</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td>DE-Hones-Holz</td>
<td>11</td>
<td>1</td>
<td>0.03</td>
<td>0.06</td>
<td>11</td>
<td>1</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>DE-Iselhausen</td>
<td>12</td>
<td>0.89</td>
<td>0.07</td>
<td>-0.03</td>
<td>12</td>
<td>0.55</td>
<td>0.21</td>
<td>0.06</td>
</tr>
<tr>
<td>DE-Tharandt</td>
<td>12</td>
<td>0.97</td>
<td>0.04</td>
<td>-0.03</td>
<td>12</td>
<td>0.96</td>
<td>0.09</td>
<td>0.08</td>
</tr>
<tr>
<td>FR-Aurade</td>
<td>12</td>
<td>0.79</td>
<td>0.11</td>
<td>0.07</td>
<td>12</td>
<td>0.84</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td>FR-Estrees-Mons</td>
<td>12</td>
<td>0.94</td>
<td>0.07</td>
<td>0.05</td>
<td>12</td>
<td>0.89</td>
<td>0.11</td>
<td>0.06</td>
</tr>
<tr>
<td>FR-Guayafax</td>
<td>12</td>
<td>0.7</td>
<td>0.04</td>
<td>-0.16</td>
<td>12</td>
<td>0.22</td>
<td>0.11</td>
<td>0.18</td>
</tr>
<tr>
<td>FR-Montiers</td>
<td>12</td>
<td>1</td>
<td>0.03</td>
<td>-0.11</td>
<td>12</td>
<td>0.58</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>FR-Puechabon</td>
<td>12</td>
<td>0.92</td>
<td>0.02</td>
<td>-0.06</td>
<td>12</td>
<td>0.96</td>
<td>0.03</td>
<td>0.08</td>
</tr>
<tr>
<td>IT-Castelporziano2</td>
<td>12</td>
<td>0.96</td>
<td>0.02</td>
<td>-0.12</td>
<td>12</td>
<td>0.85</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>IT-Collelongo</td>
<td>12</td>
<td>0.98</td>
<td>0.07</td>
<td>0.01</td>
<td>12</td>
<td>0.99</td>
<td>0.08</td>
<td>0.03</td>
</tr>
<tr>
<td>IT-Lison</td>
<td>12</td>
<td>0.93</td>
<td>0.03</td>
<td>-0.05</td>
<td>12</td>
<td>0.97</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>NE-Looeos</td>
<td>12</td>
<td>0.76</td>
<td>0.07</td>
<td>0.08</td>
<td>12</td>
<td>0.89</td>
<td>0.11</td>
<td>0.05</td>
</tr>
<tr>
<td>UK-Wytham-Woods</td>
<td>12</td>
<td>0.97</td>
<td>0.05</td>
<td>0.11</td>
<td>12</td>
<td>0.97</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td>US-Bartlett</td>
<td>12</td>
<td>0.96</td>
<td>0.04</td>
<td>-0.01</td>
<td>12</td>
<td>0.58</td>
<td>0.11</td>
<td>0.06</td>
</tr>
<tr>
<td>US-Central-Plains</td>
<td>12</td>
<td>0.69</td>
<td>0.03</td>
<td>-0.01</td>
<td>12</td>
<td>0.93</td>
<td>0.11</td>
<td>0.06</td>
</tr>
<tr>
<td>US-Harvard</td>
<td>12</td>
<td>0.99</td>
<td>0.04</td>
<td>-0.21</td>
<td>12</td>
<td>0.97</td>
<td>0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>US-Jornada</td>
<td>11</td>
<td>0.67</td>
<td>0.05</td>
<td>0.05</td>
<td>11</td>
<td>0.88</td>
<td>0.21</td>
<td>0.01</td>
</tr>
<tr>
<td>US-Moab-Site</td>
<td>12</td>
<td>0.7</td>
<td>0.02</td>
<td>0.06</td>
<td>12</td>
<td>0.15</td>
<td>0.22</td>
<td>0.01</td>
</tr>
<tr>
<td>US-Mountain-Lake</td>
<td>12</td>
<td>0.99</td>
<td>0.04</td>
<td>-0.18</td>
<td>12</td>
<td>1</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>US-Oak-Rige</td>
<td>12</td>
<td>1</td>
<td>0.02</td>
<td>0.07</td>
<td>12</td>
<td>0.59</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>US-Smithsonian</td>
<td>11</td>
<td>0.99</td>
<td>0.04</td>
<td>-0.21</td>
<td>11</td>
<td>0.99</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>US-Steigerwaldt</td>
<td>12</td>
<td>0.93</td>
<td>0.03</td>
<td>0.02</td>
<td>12</td>
<td>0.59</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>US-Talladega</td>
<td>12</td>
<td>0.98</td>
<td>0.02</td>
<td>-0.13</td>
<td>12</td>
<td>0.58</td>
<td>0.05</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Figure 69: Time-series OGV1 and OTCI and corresponding scatterplot of monthly mean for site BE-Brasschaat, Belgium, land cover Needle-leaved, evergreen. A and C represent S3A; B and D represent S3B.
Figure 70: Time-series OGVI and OTCI and corresponding scatterplot of monthly mean for site DE-Haininch, Deutschland, land cover Broadleaved, deciduous, closed. A and C represent S3A; B and D represent S3B.
Figure 71: Time-series OGVI and OTCI and corresponding scatterplot of monthly mean for site FR-EstreesMons, France, land cover Cultivated and managed areas. A and C represent S3A; B and D represent S3B.
5.3.3.2 Direct validation. Comparison with GBOV (Ground-Based Observations for Validation) v3

Ground-based measurements from the Copernicus GBOV service and OLCI GIFAPAR until 31st December 2021 have been validated against Ground-based measurements from the Copernicus GBOV service from the GLCS (Global Land Cover Service). The validation is conducted using the MERMAID extractions of Sentinel-3 (i.e., 81 for OLCI-A and 25 for OLCI-B) over thirteen established validation sites and GBOV fAPAR land product (LP4) at 300 m (WGS-84). The in-situ data are measured with DHP methodology providing FIPAR that is up-scaled at 300 m. The selected sites are distributed across various geographical locations representing different land cover types (Table 12).

The methodology involves 1) extracting the GBOV values of the Land Product (LP) at 300 m (WGS-84) overlapping the S3 pixels, 2) filtering OLCI-A/OLCI-B and GBOV data, and 3) plotting the temporal profiles. Sentinel data are filtered following the quality flags (i.e., GIFAPAR_FAIL, CLOUD, GIFAPAR_CLASS_BAD, GIFAPAR_CLASS_WS, GIFAPAR_CLASS_CSI, GIFAPAR_CLASS_BRIGHT, COSMETIC, SUSPECT, no valid values - 255). GBOV data are filtered according to the recommendations (https://gbov.acri.fr/userSupport/). Hence the following values have been removed: 1) values with input or output out of range, 2) values outside the season used to establish the calibration function (Min DOY and Max DOY Table) and 3) threshold <50% on the percentage of valid native spatial resolution pixels. Two protocols have been evaluated. The first one considers all the pixels extracted by MERMAID, and the second one performs the validation centred in the MERMAID extraction’s central point and the surrounding 3x3 window. The second method improves the validation because it reduces the heterogeneity of the validation site.

Figure 72: Comparison of OTCI-MTCI (a) and OGVI-MGVI (b). Points in the scatterplot represent the monthly mean of all available S3A and MERIS archive over 42 validation sites. Red and grey lines represent the modelled and 1:1 line respectively. The scatterplots are updated to include data up to May 2022.
### Table 12: GBOV validation sites analysed.

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Country</th>
<th>LAT</th>
<th>LON</th>
<th>IGBP</th>
<th>GBOV Dates</th>
<th>Min DOY MaxDOY</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA_BART</td>
<td>Bartlett Experimental</td>
<td>USA</td>
<td>44.0639</td>
<td>-71.2873</td>
<td>Mixed Forest</td>
<td>20140104 - 20211217</td>
<td>8 - 349</td>
</tr>
<tr>
<td></td>
<td>Forest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA_CPER</td>
<td>Central Plains Experimental Range</td>
<td>USA</td>
<td>40.81555</td>
<td>-104.7460</td>
<td>Grasslands</td>
<td>20140116 - 20211229</td>
<td>3 - 362</td>
</tr>
<tr>
<td>PRI_GUAN</td>
<td>Guanica Forest</td>
<td>Puerto Rico</td>
<td>17.9695</td>
<td>-66.8687</td>
<td>Evergreen Broadleaf</td>
<td>20150131 - 20211228</td>
<td>8 - 349</td>
</tr>
<tr>
<td>GER-HAIN</td>
<td>Hainich</td>
<td>Germany</td>
<td>51.0792</td>
<td>10.4522</td>
<td>Mixed Forest</td>
<td>20200102 - 20211221</td>
<td>8-349</td>
</tr>
<tr>
<td>USA_HARV</td>
<td>Harvard</td>
<td>USA</td>
<td>42.5378</td>
<td>-72.1715</td>
<td>Mixed Forest</td>
<td>20140104 - 20211228</td>
<td>8 - 349</td>
</tr>
<tr>
<td>USA_JORN</td>
<td>Jornada</td>
<td>USA</td>
<td>32.5907</td>
<td>-106.842</td>
<td>Open Shrubland</td>
<td>20160106 - 20211229</td>
<td>3 - 362</td>
</tr>
<tr>
<td>AUS_Litchfield</td>
<td>Litchfield</td>
<td>AUS</td>
<td>-13.18</td>
<td>130.79</td>
<td>Woody Savannas</td>
<td>20210308 – 20211215</td>
<td>3 - 362</td>
</tr>
<tr>
<td>USA_Moab</td>
<td>Moab</td>
<td>USA</td>
<td>38.2483</td>
<td>-109.388</td>
<td>Open Shrubland</td>
<td>20150108 – 20201220</td>
<td>3 - 362</td>
</tr>
<tr>
<td>USA_ORNL</td>
<td>Oak Ridge</td>
<td>USA</td>
<td>35.9641</td>
<td>-84.2826</td>
<td>Mixed Forest</td>
<td>20150117 – 20211230</td>
<td>8 - 349</td>
</tr>
<tr>
<td>USA_SCBI</td>
<td>Smithsonian Conservation Biology</td>
<td>USA</td>
<td>38.8902</td>
<td>-78.1395</td>
<td>Mixed Forest</td>
<td>20150105 – 20211224</td>
<td>8 – 349</td>
</tr>
<tr>
<td>USA_STEI</td>
<td>Steigerwaldt Land Services</td>
<td>USA</td>
<td>45.5089</td>
<td>-89.5864</td>
<td>Deciduous Broadleaf</td>
<td>20160123 – 20211229</td>
<td>8 - 349</td>
</tr>
<tr>
<td>USA_TALL</td>
<td>Talladega National Forest</td>
<td>USA</td>
<td>32.9504</td>
<td>-87.3933</td>
<td>Evergreen Needleleaf</td>
<td>20150131 – 202112283</td>
<td>8 - 349</td>
</tr>
<tr>
<td>AUS_TUMB</td>
<td>Tumbarumba</td>
<td>AUS</td>
<td>-35.6565</td>
<td>148.1516</td>
<td>Evergreen Broadleaf</td>
<td>20200103-20201230</td>
<td>8 - 349</td>
</tr>
</tbody>
</table>
Figure 73: Comparison of GBOV vs OLCI-A (left) and GBOV vs OLCI-B (right). Points in the scatterplot represent the monthly mean of all available S3A and GBOV data (3×3) over 13 validation sites. Red and black lines represent the modelled and 1:1 line, respectively.

OLCI-A

Figure 74 to Figure 80 show the variability of both in situ values from GBOV and their corresponding OLCI-A matchups in time. OLCI-A reproduces the temporal variations of GBOV values for almost all sites satisfactorily. However, GBOV products provide systematically higher values than satellite products in forest classes. On the contrary, for shrubland and grasslands, satellite products tend to present higher values than GBOV. When applying the 3×3 windows (red colours), results remove outliers and suspicious values compared with the validation considering all the S3A values (i.e., 81).

Figure 74: Time series of GBOV FIPAR LP4 (red/green) and OLCI-A GIFAPAR (grey) for sites Steigerwaldt (left) and Guanica (right)
Figure 75: same as Figure 74 for sites Talladega (Left) and Tumbarumba (right)

Figure 76: same as Figure 74 for sites Bartlett (Left) and Oak (right)
Figure 77: same as Figure 74 for sites Hainich (Left) and Harvard (right)

Figure 78: same as Figure 74 for site Smithsonian
Figure 79: same as Figure 74 for sites Moab (Left) and CPER (right)

Figure 80: same as Figure 74 for sites Jornada (Left) and Litchfield (right)

OLCI-B

Figure 81 to Figure 87 show the temporal profiles of GBOV LP4 against OLCI-B GIFAPAR. Overall, there are similar seasonal trajectories between OLCI-B and GBOV. The pattern is similar to OLCI-A, with higher values of GBOV than in OLCI-B for forest classes. Similarly, grassland and shrubland classes show lower values than OLCI-B. Furthermore, the number of GBOV values after filtering must be increased to perform the validation, mainly in grasslands and shrublands sites (i.e., Moab, Central Plains). Applying the 3x3 protocol removes the outliers, but only one validation point remains in some cases (i.e., Moab). Nevertheless, validation results in 2022 should improve due to a modification of the MERMAID extraction for S3B that is expected to increase the number of validation points (see data for 2022 in the time-series profiles).
Figure 81: Time series of GBOV FIPAR LP4 (red/green) and OLCI-A GIFAPAR (grey) for sites Steigerwaldt (left) and Guanica (right)
Figure 82: same as Figure 81 for sites Talladega (Left) and Tumbarumba (right)

Figure 83: same as Figure 81 for sites Bartlett (Left) and Oak (right)

Figure 84: same as Figure 81 for sites Hainich (Left) and Harvard (right)
Figure 85: same as Figure 81 for site Smithsonian

Figure 86: same as Figure 85 for sites Moab (Left) and CPER (right)

Figure 87: same as Figure 81 for sites Jornada (Left) and Litchfield (right)
5.3.3.3 Sentinel-3A and 3B biophysical variables inter comparison results

To check the consistency and seasonality of OLCI GIFAPAR, the scatter plots between OLCI-A and OLCI-B for equal dates have been performed by GBOV site (left) and by land cover type (right). A high correlation (>0.9) is found.

![Scatter plots between OLCI-A and OLCI-B of equal dates for GBOV sites (left) and Land Cover types (right)](image)

Figure 88: Scatter plots between OLCI-A and OLCI-B of equal dates for GBOV sites (left) and Land Cover types (right)

5.3.3.4 Sentinel-3A and 3B biophysical variables inter-annual variability results

Monthly mean GIFAPAR and OTCI were calculated for S3A and S3B from MERMAID extractions for all years where data was available and over the CEOS-LPV sites (Table 10). Annual temporal profiles have been created to assess variability in GIFAPAR and OTCI between years and to compare the agreement between S3A and S3B.

The interannual variability of GIFAPAR and OTCI is variable among the CEOS sites. Several sites, such as US-Bartlett and CA-Mer-Bleue, show relative consistency between years in terms of value, seasonal pattern and the timing of peaks (Figure 89 and Figure 90). On the contrary, sites such as AU-Russworth and AU-Warra Tall show more variability between years of fluctuating yearly peaks and valleys (Figure 91 and Figure 92). S3A and S3B appear similar, with comparable seasonal trends for GIFAPAR and OTCI and equivalent value ranges.
Figure 89: Time series of monthly mean GIFAPAR and OTCI for S3A and S3B for site US-Bartlett. Top-left and right represent S3A; Bottom-left and right represent S3B.
Figure 90: Time series of monthly mean GIFAPAR and OTCI for S3A and S3B for site CA-Mer-Bleue. Top-left and right represent S3A; Bottom-left and right represent S3B.
Figure 91: Time series of monthly mean OTCI and GIFAPAR for S3A and S3B for site AU-Rushworth. Top-left and right represent S3A; Bottom-left and right represent S3B.
Figure 92: Time series of monthly mean OTCI and GIFAPAR for S3A and S3B for site AU-Warra Tall. Top-left and right represent S3A; Bottom-left and right represent S3B.
6 Summary of performances – SLSTR

6.1 Instrument performances

6.1.1 Instrument temperatures

As a thermal infrared instrument, thermal stability and uniformity of the optical mechanical enclosure, OME is critical to the radiometric calibration. During normal operations, temperatures have remained generally stable and consistent during the reporting period, with gradual changes due to the seasonal cycle, which are consistent with the previous years of operations for SLSTR-A. The exceptions are when the normal mode was disrupted by instrument operations or anomalies – for example, the anomaly and subsequent decontamination of 1-4th August for SLSTR-A, or the anomaly of 20th-22nd June for SLSTR-B.

Figure 93: Baffle temperature trends for SLSTR-A (left) and SLSTR-B (right) from 1st Jan 2022 to end of Jan 2023. The vertical dashed lines indicate the start and end of each cycle.
Figure 94: OME temperature trends for SLSTR-A (left) and SLSTR-B (right) from Jan 2022 to end of Jan 2023, showing the paraboloid stops and flip baffle (top two plots) and optical bench and scanner and flip assembly (lower two plots). The vertical dashed lines indicate the start and end of each cycle.
6.1.2 Detector Temperatures

The cooler is performing well, maintaining the IR detectors between 85 and 89 K – see Figure 95 and Figure 96. The IR FPA is affected by water ice contamination as is common for instruments with cryogenic optics, and was observed for all ATSR instruments. This affects the heat load on the IR FPA, which requires the cooler to run at increased drive levels and also affects the optical throughput of the channels. Therefore, periodic decontamination cycles are needed to remove the water ice from the cold surfaces. Unplanned decontaminations were performed on SLSTR-A in July and SLSTR-B in June as a result of instrument anomalies.
Figure 95: SLSTR-A detector temperatures for each channel from 1st Jan 2022 to end of Jan 2023. The discontinuity for the infrared channels occurs where the FPA was heated for decontamination in July. The vertical dashed lines indicate the start and end of each cycle.

Figure 96: SLSTR-B detector temperatures for each channel from Jan 2022 to end of Jan 2023. The discontinuity in the IR channels in June 2022 is a result of the anomaly and subsequent decontamination. The vertical dashed lines indicate the start and end of each cycle.
6.1.3 Scanner performance

The scanners have performed consistently since launch, operating within required limits. The scanners are controlled and monitored by absolute encoders mounted on the drive shafts. Scan jitter statistics for SLSTR-A are shown in Figure 98 and for SLSTR-B in Figure 100 with respect to the linear control law within each orbit.

6.1.3.1 SLSTR-A scanner performance

Overall, the SLSTR-A scanner and flip statistics are as good or better than the previous years, showing that the mirror mechanisms are performing well.

![Figure 97: SLSTR-A histogram of standard deviation with respect to the linear control law of the scanners and flip mirror in each orbit for the year 2022 for nadir view (left) and oblique view (right).]
Figure 98: SLSTR-A scanner and flip jitter from January 2022 to January 2023, showing mean (red) and stddev (blue) compared to the expected one for the nadir view (left) and oblique view (right). The vertical dashed lines indicate the start and end of each cycle.

6.1.3.2 SLSTR-B scanner performance

The scanner statistics for SLSTR-B are shown in Figure 99 and are generally consistent with the previous years. The SLSTR-B scanner and flip mean and standard deviations from their expected positions are broadly consistent with SLSTR-1. However we note that the oblique scanner has a slightly larger mean deviation of ~3”, and that the flip scanner has a bimodal distribution. The different behaviour of the flip mirror for SLSTR-B had been observed in the early years of the mission. There is no clear deterioration trend during the mission lifetime and no impact on image quality has been detected.
6.1.4 Black-Bodies

The blackbodies have functioned well over the reporting period. The heated blackbody (+YBB) is being maintained by the heaters approximately 37-38 K above the cool blackbody (-YBB). The long-term trends show no discernible degradation in the performance of the heaters.

Figure 99: SLSTR-B histogram of standard deviation with respect to the linear control law of the scanners and flip mirror for the year 2022.

Figure 100: SLSTR-B scanner and flip jitter, showing mean (red) and stddev (blue) compared to the expected one for the oblique view for March 2022 to February 2023. The vertical dashed lines indicate the start and end of each cycle.
6.1.4.1 SLSTR-A Black-Bodies

Figure 101 shows the blackbody temperatures and baseplate gradients for SLSTR-A. During December, the heated BB increased in temperature as the satellite approached perihelion. The peak temperature has been increasing year-by-year but the trend is levelling out. In order to ensure that the peak temperature remains below 305 K to avoid saturation of S7, the commanded heater power was decreased by one bit on 28th September 2021. This decrease immediately caused a decrease in the hot blackbody temperature of ~0.3 K. No negative impact on image quality has been detected so far.

Figure 101: SLSTR-A blackbody temperature and baseplate gradient trends for Jan 2022 to Jan 2023. The vertical dashed lines indicate the start and end of each cycle. Discontinuities are caused by the decontamination in August 2022, and the black-body crossover test at the end of September.
6.1.4.2 SLSTR-B Black-Bodies

Figure 102 shows the blackbody temperatures and baseplate gradients for SLSTR-B. The difference of the 5 PRTs located on the blackbody baseplate with the average base temperature are also plotted in Figure 102. The spread in temperature of the baseplate PRTs is largest when the blackbody is heated. In particular when the +YBB is hot, PRT1 is warmer than the average by approximately 70 mK whereas the other PRTs all cluster closely together. This difference was expected before launch and is consistent with measurements made during the ground testing.

As for SLSTR-A, the commanded blackbody heater power was reduced by one bit on 1st October 2021. No significant impact can be seen on the hot Black Body temperature in 2022.
6.1.4.3 Blackbody Cross-Over Tests

Blackbody cross-over tests are carried out at yearly intervals to compare the radiometric signals in the thermal channels when the two blackbodies are at identical temperatures. The test is performed to determine the effects of any drifts in the blackbody thermometer calibration or change in target emissivity caused by a deterioration of the black surface finish.

The method is based on that for AATSR on ENVISAT and has been performed for SLSTR during pre-launch calibration, then in-flight during commissioning and at yearly intervals to determine any changes in the blackbody performance.

It is important to note that this is not an absolute test of the blackbody performance since we do not have an independent method to evaluate the absolute radiances from the blackbodies on-orbit to sufficient accuracy (SLSTR is intended to have a radiometric uncertainty <0.1 K which is at the limit for most space-borne instruments). However, we are able to deduce any relative calibration errors between channels or trends in the blackbody thermometer calibration or change in target emissivity caused by a deterioration of the black surface finish. The method does not distinguish which effect is dominant because the two are highly correlated. However, the results do provide a means to verify the uncertainty in the BB radiances.

The test was performed by switching the heated blackbody from the +YBB to the –YBB (and vice versa) and allowing the temperatures to cross over and stabilise. The most recent tests for this reporting period were performed between 27th September (SLSTR-A) and 16th October (SLSTR-B), with crossover temperatures of 290.161/291.650 K for SLSTR-A and 289.607/291.210 K for SLSTR-B.

The analysis is performed by comparing the radiometric signals close to the cross-over times as a function of the baseplate temperatures as measured by the PRTs. Here, we can estimate the effective temperature difference between the two BBs from the slope \( \frac{dT}{dN} \), which is obtained by a simple linear fit to the data. So,

\[ \Delta T = \frac{dT}{dN} \Delta DN \]

Figure 103 shows \( \Delta T \) versus time for all of the BB cross-over tests performed to date, including the pre-launch measurements (6 tests for SLSTR-A and 4 tests for SLSTR-B).

For SLSTR-A, the results show that there has been some steady drift with time, and there is a possible correlation with the baseplate gradients for the second cross-over. For SLSTR-B, the results are largely consistent with the pre-launch measurements on-orbit. What is not expected is the most recent test results show a drop in the S8 and S9 channel differences for the SLSTR-A nadir view and the SLSTR-B oblique view. At the time of writing the cause of this is under investigation. The S8 and S9 channel differences for SLSTR-A oblique view recovered their long-term level after a drop in 2021. For SLSTR-B oblique view on the other hand the value remained lower in 2022 as in 2021.

Note that analysis reported in previous reports suggested a large discrepancy with S8 and S9 in the nadir view. This has since been traced to an incorrect ordering of the detector counts for odd/even integrators.
in the processed data and has since been rectified. Although the detector is the same, the pixel integrators have a slightly different gain resulting in different counts for the same scene temperature. The L1 data processing in the IPF does handle this correctly.

![Figure 103: BT differences vs time for all of the blackbody cross-over tests performed to date (including pre-launch measurements) for SLSTR-A (left) and SLSTR-B (right). The part 1 crossover is shown in the top plots, and part 2 in the lower plots. Different symbols indicate different channels (S7: square, S8: triangle, S9: diamond) and different colours indicate nadir (red) and oblique (blue) views. Error bars are derived from the blackbody temperature gradients and standard deviations of the BB signals during the cross-over.](image)

### 6.1.5 IR Channels

#### 6.1.5.1 Dynamic Range and Digitisation

The TIR channels (S7-S9, F1 and F2) are all functioning with no reported loss of data or digital resolution. The IR gains show an increase as detector temperatures warm-up between decontamination cycles (Figure 104 and Figure 105, left). Comparisons between nadir and oblique views show that the radiometric gains are consistent (Figure 106), within 1-2%.

The IR offsets show small variations due to detector and optics temperature variations (Figure 104 and Figure 105, right). These offset variations determine the minimum BTs detectable for channels S8 and S9, which also change with time. Note that each detector and odd/even pixels have different offset values.
Figure 104: SLSTR-A gain (left) and offset (right) trends for the TIR channels in nadir view. The different colour symbols show the response for each of the detector elements and integrators in the channels. The discontinuity is due to the decontamination (July 2022) and the blackbody crossover test (end Sept 2022). The offset plot for S8 and F2 also shows a step at the end of January 2021 when the commanded detector offset was updated for one detector.

Figure 105: SLSTR-B gain (left) and offset (right) trends for the TIR channels in nadir view. The different colour symbols show the response for each of the detector elements and integrators in the channels. The discontinuity in June is due to the decontamination while the spike in Sept 2022 is due to the blackbody crossover test.
6.1.5.2 Radiometric Noise

The thermal channel NEDT values derived from the on-board blackbody sources are consistent with previous operations and within the requirements – see Figure 108, Table 13 and Table 14. Noise levels haven’t changed significantly following the decontaminations. The NEDT levels are roughly consistent between SLSTR-A and SLSTR-B. The higher variability of the F1 channel NEDT for SLSTR-B observed in the
previous year is no longer visible in 2022. For SLSTR-B there appears to be a small but gradual increase of S8 and S9 NEDT over time.

![Figure 108: NEDT trend for the thermal channels for SLSTR-A (left) and SLSTR-B (right). Blue points were calculated from the cold blackbody signal and red points from the hot blackbody.](image)

**Table 13: NEDT for SLSTR-A in cycles 067-078 averaged over all detectors for both Earth views towards the hot +YBB (top) and the cold -YBB (bottom).**

<table>
<thead>
<tr>
<th>SLSTR-A</th>
<th>JAN22</th>
<th>FEB22</th>
<th>MAR22</th>
<th>APR22</th>
<th>MAY22</th>
<th>JUN22</th>
<th>JUL22</th>
<th>AUG22</th>
<th>SEP22</th>
<th>OCT22</th>
<th>NOV22</th>
<th>DEC22</th>
</tr>
</thead>
<tbody>
<tr>
<td>+YBB temp (K)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>17.0</td>
<td>17.3</td>
<td>17.3</td>
<td>17.5</td>
<td>17.7</td>
<td>17.7</td>
<td>17.6</td>
<td>83.8</td>
<td>17.4</td>
<td>17.3</td>
<td>17.1</td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>11.8</td>
<td>11.8</td>
<td>11.9</td>
<td>12.0</td>
<td>12.1</td>
<td>12.1</td>
<td>12.2</td>
<td>11.9</td>
<td>16.0</td>
<td>12.0</td>
<td>12.0</td>
<td>12.0</td>
</tr>
<tr>
<td>S9</td>
<td>18.5</td>
<td>18.6</td>
<td>18.6</td>
<td>18.6</td>
<td>18.7</td>
<td>18.8</td>
<td>18.9</td>
<td>18.3</td>
<td>18.5</td>
<td>18.4</td>
<td>18.5</td>
<td>18.4</td>
</tr>
<tr>
<td>F1</td>
<td>275</td>
<td>279</td>
<td>281</td>
<td>284</td>
<td>286</td>
<td>288</td>
<td>289</td>
<td>283</td>
<td>311</td>
<td>282</td>
<td>279</td>
<td>278</td>
</tr>
<tr>
<td>F2</td>
<td>35.4</td>
<td>35.2</td>
<td>34.8</td>
<td>34.8</td>
<td>34.9</td>
<td>35.4</td>
<td>34.7</td>
<td>34.6</td>
<td>53.0</td>
<td>34.9</td>
<td>35.2</td>
<td>35.4</td>
</tr>
</tbody>
</table>
### Table 14: NEDT for SLSTR-B in cycles 048-059 averaged over all detectors for both Earth views towards the hot +YBB (top) and the cold -YBB (bottom)

<table>
<thead>
<tr>
<th>SLSTR-B</th>
<th>JAN22</th>
<th>FEB22</th>
<th>MAR22</th>
<th>APR22</th>
<th>MAY22</th>
<th>JUN22</th>
<th>JUL22</th>
<th>AUG22</th>
<th>SEP22</th>
<th>OCT22</th>
<th>NOV22</th>
<th>DEC22</th>
</tr>
</thead>
<tbody>
<tr>
<td>+YBB temp (K)</td>
<td>304.027</td>
<td>303.642</td>
<td>303.260</td>
<td>303.043</td>
<td>302.754</td>
<td>302.642</td>
<td>302.565</td>
<td>302.706</td>
<td>303.138</td>
<td>303.596</td>
<td>303.894</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>16.1</td>
<td>16.0</td>
<td>16.1</td>
<td>16.2</td>
<td>16.5</td>
<td>16.5</td>
<td>16.5</td>
<td>16.5</td>
<td>16.8</td>
<td>16.2</td>
<td>16.1</td>
<td>16.1</td>
</tr>
<tr>
<td>S8</td>
<td>14.2</td>
<td>14.3</td>
<td>14.3</td>
<td>14.3</td>
<td>14.3</td>
<td>14.0</td>
<td>14.0</td>
<td>14.0</td>
<td>14.1</td>
<td>14.0</td>
<td>13.9</td>
<td>14.1</td>
</tr>
<tr>
<td>S9</td>
<td>16.5</td>
<td>16.5</td>
<td>16.5</td>
<td>16.5</td>
<td>16.4</td>
<td>15.8</td>
<td>15.8</td>
<td>15.9</td>
<td>16.0</td>
<td>15.9</td>
<td>15.9</td>
<td>16.4</td>
</tr>
<tr>
<td>F1</td>
<td>373</td>
<td>377</td>
<td>377</td>
<td>385</td>
<td>395</td>
<td>368</td>
<td>367</td>
<td>362</td>
<td>375</td>
<td>350</td>
<td>345</td>
<td>361</td>
</tr>
<tr>
<td>F2</td>
<td>30.5</td>
<td>30.3</td>
<td>30.3</td>
<td>30.3</td>
<td>30.3</td>
<td>30.2</td>
<td>30.1</td>
<td>30.1</td>
<td>30.4</td>
<td>30.4</td>
<td>30.6</td>
<td>30.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SLSTR-B</th>
<th>JAN22</th>
<th>FEB22</th>
<th>MAR22</th>
<th>APR22</th>
<th>MAY22</th>
<th>JUN22</th>
<th>JUL22</th>
<th>AUG22</th>
<th>SEP22</th>
<th>OCT22</th>
<th>NOV22</th>
<th>DEC22</th>
</tr>
</thead>
<tbody>
<tr>
<td>-YBB temp (K)</td>
<td>266.579</td>
<td>266.052</td>
<td>266.540</td>
<td>265.531</td>
<td>265.342</td>
<td>265.060</td>
<td>264.868</td>
<td>264.982</td>
<td>265.525</td>
<td>266.220</td>
<td>266.554</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>42.6</td>
<td>43.3</td>
<td>44.2</td>
<td>44.3</td>
<td>44.3</td>
<td>44.3</td>
<td>44.7</td>
<td>45.6</td>
<td>45.2</td>
<td>43.3</td>
<td>43.2</td>
<td>42.9</td>
</tr>
<tr>
<td>S8</td>
<td>18.1</td>
<td>18.1</td>
<td>18.2</td>
<td>18.2</td>
<td>18.3</td>
<td>18.1</td>
<td>17.9</td>
<td>17.9</td>
<td>17.9</td>
<td>17.8</td>
<td>17.9</td>
<td>17.9</td>
</tr>
<tr>
<td>S9</td>
<td>21.0</td>
<td>21.1</td>
<td>21.3</td>
<td>21.4</td>
<td>21.4</td>
<td>21.1</td>
<td>20.3</td>
<td>20.3</td>
<td>20.4</td>
<td>20.2</td>
<td>20.4</td>
<td>20.4</td>
</tr>
<tr>
<td>F1</td>
<td>1483</td>
<td>1514</td>
<td>1568</td>
<td>1544</td>
<td>1585</td>
<td>1600</td>
<td>1526</td>
<td>1549</td>
<td>1521</td>
<td>1441</td>
<td>1435</td>
<td>1427</td>
</tr>
<tr>
<td>F2</td>
<td>33.6</td>
<td>33.7</td>
<td>33.9</td>
<td>33.9</td>
<td>33.9</td>
<td>33.7</td>
<td>33.1</td>
<td>33.3</td>
<td>33.2</td>
<td>33.1</td>
<td>33.1</td>
<td>33.1</td>
</tr>
</tbody>
</table>
6.1.6 VIS/SWIR Channels

6.1.6.1 Radiometric gain variation SLSTR-A

Overall the S1-S6 channels are functioning well with no reported loss of data or digital resolution.

The main issue affecting the S1-S3 channels are oscillations in the radiometric response due to the build-up of ice on the optical path within the FPA. This is illustrated in Figure 109, which shows the variation of the radiometric gain derived from the VISCAL signals. These oscillations were observed for the corresponding channels on ATSR-2 and AATSR. Periodic decontamination of the IR FPA is necessary to remove the water ice contamination.

The trends of the radiometric gain variation clearly show where the decontamination took place, and that the signal was reset afterwards. During the decontamination, only the VIS channels are operating and the SWIR channels are switched off, causing a gap in the trends due to the loss of data.

The radiometric responses of S4-S6 appear to be more stable and not affected by the build-up of water ice contamination, Figure 110. There is a seasonal cycle of the response of ±1% that could be caused by variations in the solar zenith angle on the diffuser or partial vignetting of the Sun’s disc by the VISCAL baffle.
Figure 109: Gain trend for VIS channels (nadir view) for SLSTR-A. The data have been adjusted to allow for the variation of the solar intensity. The oscillations in the signal are due to the build-up of a thin condensation layer causing a thin film interference effect. The different colour symbols show the response for each of the 4 detector elements in the VIS channels.
Figure 110: Gain trend for SWIR channels (nadir view) for SLSTR-A. Outliers in the plots are due to gaps in L0 data or decontamination cycles. The different colour symbols show the response for each of the 8 detector elements of the A and B stripes of the SWIR channels.
6.1.6.2 Radiometric gain variation SLSTR-B

As in SLSTR-A, one of the main issues affecting the S1-S3 channels are oscillations in the radiometric response due to the build-up of ice on the optical path within the FPA. However, there is also a problem with S1 and S2 in particular, which show noisy behaviour and numerous drops in signal as shown in Figure 111. This gives 2-3% errors in the radiometric calibration of these channels. The effect has been the subject of a major NCR led by ESA-ESTEC. A number of candidate root causes have been identified, with the most likely due to motional chopping of the VIS detectors by an internal aperture in the VIS optical bench. Because the effect appears to be random it is most likely affecting all the data for S1 and S2.

The radiometric responses of S4-S6 appear to be more stable and not affected by the build-up of water ice contamination, Figure 112.

![Radiometric gain variation SLSTR-B](image)

*Figure 111: Gain trend for VIS channels (nadir view) for SLSTR-B. The data have been adjusted to allow for the variation of the solar intensity. The oscillations in the signal are due to the build-up of a thin condensation layer*
causing a thin film interference effect. The different colour symbols show the response for each of the 4 detector elements in the VIS channels.

Figure 112: Gain trend for SWIR channels (nadir view) for SLSTR-B. Outliers in the plots are due to gaps in L0 data or decontamination cycles. The different colour symbols show the response for each of the 8 detector elements of the A and B stripes of the SWIR channels.
6.1.6.3 Dark signal variation SLSTR-A

The dark signal variation derived from the nadir blackbody signals for the VIS and SWIR channels is stable – see Figure 113 and Figure 114.

Figure 113: Dark signal trend for VIS channels (nadir view) for SLSTR-A. The different colour symbols show the signal for each of the 4 detector elements in the VIS channels.
6.1.6.4 Dark signal variation SLSTR-B

The dark signal variation derived from the nadir blackbody signals for the VIS and SWIR channels is stable for SLSTR-B (Figure 115 and Figure 116).

Figure 114: Dark signal trend for SWIR channels (nadir view) for SLSTR-A. The different colour symbols show the signal for each of the 8 detector elements of the A and B stripes of the SWIR channels. The gap/jump in April is due to the decontamination.
Figure 115: Dark signal trend for VIS channels (nadir view) for SLSTR-B. The different colour symbols show the signal for each of the 4 detector elements in the VIS channels.
Figure 116: Dark signal trend for SWIR channels (nadir view) for SLSTR-B. The different colour symbols show the signal for each of the 8 detector elements in the SWIR channels.

6.1.6.5 Radiometric noise for SLSTR-A

The VIS/SWIR channel signal-to-noise ratio is derived from the VISCAL signal at full solar illumination. The measurements show that the SNR is stable and consistent over the year and largely unaffected by anomalies and decontamination.
Figure 117: SLSTR-A VIS and SWIR channel signal-to-noise. Different colours indicate different detectors.
6.1.6.6 Radiometric noise for SLSTR-B
6.1.6.7 Contamination

The monitoring of the VISCAL signal shows that the performance of the VIS and the SWIR channels has been affected by the build-up of a condensation layer on the FPA. The build-up of condensation on the optics was expected since similar patterns were observed previously in AATSR and ATSR-2.

The periodic pattern observed in the VISCAL signals depends on the rate of build-up of the condensation layer and the wavelength of the channel. So, an estimation of the layer thickness can be obtained from the oscillations in the visible channels signal that occurred at $x = l/2, l, 3l/2$, etc.
The growth of the ice layer is slow and decontamination activities are performed only once or twice per year. The rate of growth of the ice layer has reduced significantly with respect to that observed after the first cool down. It is expected that the rate of build-up will decrease with time resulting in longer periods between decontamination cycles.

### 6.2 L1 products performances

#### 6.2.1 TIR Channel Calibration

##### 6.2.1.1 SLSTR comparisons with IASI

The absolute radiometric calibration of the IR channels is being validated by EUMETSAT using comparisons against IASI-A and B (Tomazic et al 2018). Comparisons were performed during the commissioning phases for SLSTR-A and SLSTR-B. Currently there are no updates since 2018. The mission requirement is that the absolute radiometric calibration should be accurate to 0.2 K traceable to ITS-90, and that at a minimum this should be met in the temperature range between the two blackbodies.

The latest results have not changed since the previous annual reports – i.e. from Tomazic et al (2018). These results confirmed very good performance with almost no bias (<0.1 K) for channels S8 and S9 in the nadir view over the temperature range 220 – 280 K.

#### 6.2.2 VIS/SWIR Channel Calibration

Vicarious calibration methods are used to verify the radiometric calibration of the SLSTR visible (VIS) and shortwave infrared (SWIR) channels, and currently two methods are used.

1. Inter-comparisons of SLSTR with similar sensors such as OLCI, AATSR and MODIS using stable desert targets.

2. Compare SLSTR observed radiances over scenes containing sun-glint against the predicted top-of-atmosphere radiances computed radiative transfer models.

Both approaches provide consistent results. Table 15 and Table 17 show the relative differences obtained with the different calibration methods.

For analysis over desert sites we have used the extractions provided by the S3ETRAC tool, which contain the sensor reflectance values, cloud fraction, geometric and meteorological information needed for the analysis. For analysis over sun-glint regions we have used L1 products directly rather than the S3ETRAC analysis as the latter only contains a single value, and the analysis requires the full image context to model the sunglint.

##### 6.2.2.1 Inter-comparisons of SLSTR over desert sites

The analysis performed follows the methodology used for the comparisons of AATSR with MERIS and MODIS-A (see Smith and Cox 2013). The analysis needs to take into consideration a number of effects:

- Temporal differences: in particular, direct comparisons of SLSTR with AATSR or MERIS are not possible because the latter are no longer operating. Also, sensors such as MODIS-A do not observe the site at the same time. So, to perform the comparisons we compare for the same view/solar geometry assuming that the site is stable over long timescales.
Spectral differences: although SLSTR and OLCI have common spectral bands, the spectral responses are not exactly the same, which can give rise to differences in spectral reflectance of a few percent if not accounted for. Hence, we need to account for differences in atmospheric transmission and surface spectral reflectance.

Geometric differences: although the method attempts to perform the comparisons with the same view/solar geometry, an exact match is not always possible. To account for this, we construct a basic geometric model from the reference sensor to interpolate to the correct geometry.

The data are extracted by S3ETRAC tool over a series of pre-defined sites. These sites have been selected for their appropriate optical properties to validate the radiometry of optical sensors. Table 15 shows the desert sites and their geographical limits used for the assessment and monitoring of the VIS and SWIR radiometric calibration.

**Table 15: The list of these sites and their geographical limits**

<table>
<thead>
<tr>
<th>Site</th>
<th>North Latitude</th>
<th>South Latitude</th>
<th>East Longitude</th>
<th>West Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEOS_ALGERIA-3</td>
<td>30.82</td>
<td>29.82</td>
<td>8.16</td>
<td>7.16</td>
</tr>
<tr>
<td>CEOS_ALGERIA-5</td>
<td>31.52</td>
<td>30.52</td>
<td>2.73</td>
<td>1.73</td>
</tr>
<tr>
<td>CEOS_LIBYA-1</td>
<td>24.92</td>
<td>23.92</td>
<td>13.85</td>
<td>12.85</td>
</tr>
<tr>
<td>CEOS_LIBYA-4</td>
<td>29.05</td>
<td>28.05</td>
<td>23.89</td>
<td>22.89</td>
</tr>
<tr>
<td>CEOS_MAURITANIA-1</td>
<td>19.9</td>
<td>18.9</td>
<td>-8.8</td>
<td>-9.8</td>
</tr>
<tr>
<td>CEOS_MAURITANIA-2</td>
<td>21.35</td>
<td>20.35</td>
<td>-8.28</td>
<td>-9.28</td>
</tr>
<tr>
<td>RAL_Algeria-1</td>
<td>24.3</td>
<td>23.3</td>
<td>0.1</td>
<td>-0.9</td>
</tr>
<tr>
<td>RAL_Algeria-2</td>
<td>26.59</td>
<td>25.59</td>
<td>-0.88</td>
<td>-1.88</td>
</tr>
<tr>
<td>RAL_Algeria-4</td>
<td>30.54</td>
<td>29.54</td>
<td>6.09</td>
<td>5.09</td>
</tr>
<tr>
<td>RAL_Arabia-1</td>
<td>19.38</td>
<td>18.38</td>
<td>47.26</td>
<td>46.26</td>
</tr>
<tr>
<td>RAL_Arabia-2</td>
<td>20.63</td>
<td>19.63</td>
<td>51.46</td>
<td>50.46</td>
</tr>
<tr>
<td>RAL_Arabia-3</td>
<td>29.42</td>
<td>28.42</td>
<td>44.23</td>
<td>43.23</td>
</tr>
<tr>
<td>RAL_Sundan-1</td>
<td>22.24</td>
<td>21.24</td>
<td>28.72</td>
<td>27.72</td>
</tr>
<tr>
<td>RAL_Niger-1</td>
<td>20.17</td>
<td>19.17</td>
<td>10.31</td>
<td>9.31</td>
</tr>
</tbody>
</table>
Figure 119 shows the combined results for all the desert sites when SLSTR-A and B are compared with AATSR in nadir view, for the VIS and S5 channels.

Figure 120 shows comparisons for the oblique view.

Overall the calibration of SLSTR-A and B is very stable over the mission lifetime. However, there does appear to be a small drift of >1% for SLSTR-A channel S3 (nadir and oblique views) and S1 nadir view (both satellites).
Figure 119: Comparisons of SLSTR-A (red) and SLSTR-B (blue) S1-S3 and S5a and S5b channels vs. the corresponding channels for AATSR over desert sites.
Figure 120: Inter-comparisons between SLSTR-A and SLSTR-B S1-S3 and S5, oblique view and AATSR.
6.2.2.2 Radiometric validation with DIMITRI

Radiometric Validation is performed at the Optical MPC using three indirect methods, comparing simulated TOA radiances to that measured by the SLSTR instrument.

❖ The “Rayleigh” method: measurement of the Rayleigh atmospheric backscattering over open ocean sites in clear sky off-glint conditions with low aerosol load to provide absolute calibration in the blue-to-red spectral domain.

❖ The “Glint” method: using the specular reflection of the sun (i.e. sun glint) on the open ocean surface and its known spectral dependency to assess inter-band calibration in the red-to-SIWR spectral range. Its results are then scaled to an absolute result at a reference band, 659 nm, taken from the PICS method for DIMITRI.

❖ The PICS method: measurement over well characterized, temporally stable desert areas (Pseudo-Invariant Calibration Sites or PICS) to provide absolute calibration over the whole spectral domain. This method also allows cross-mission intercomparison with other sensors providing comparable spectral channels (e.g. Aqua/MODIS, S3/OLCI and AATSR/3REP).

6.2.2.2.1 Results of Validation over desert-PICS sites

The ingestion of the available L1-RBT-NT products from SLSTR-A and SLSTR-B over the 6 CEOS-desert CalVal-sites (Algeria3 & 5, Libya 1 & 4 and Mauritania 1 & 2) has been performed until the 30th December 2022. Automated cloud screening is performed using Globcarbon-algorithm implemented in DIMITRI. Then the Desert-PICS method applied over VNIR bands S01, S02 and S03).

The results are consistent over all the six used PICS sites (Figure 121 and Figure 122). Both sensors show a good stability over the analysed period over VNIR bands for the NADIR view.

The temporal average over the period January-December 2022 of the elementary ratios (observed reflectance to the simulated one) for SLSTR-A and SLSTR-B are reported below.

Both sensors show gain values between 3-5% over the VNIR bands S1-S3 (Figure 125).
Figure 121: Time-series of the elementary ratios (observed/simulated) signal from SLSTR-A for (top to bottom) bands S01, S02 and S03 respectively over January - December 2022 from the six PICS Cal/Val sites. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the desert methodology uncertainty.
Figure 122: Time-series of the elementary ratios (observed/simulated) signal from SLSTR-B for (top to bottom) bands S01, S02 and S03 respectively over January - December 2022 from the six PICS Cal/Val sites. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the desert methodology uncertainty.
6.2.2.2.2 Results of the validation over Rayleigh methodology

This activity is dedicated to the assessment of the radiometry measurements of SLSTR on board of Sentinel-3A and Sentinel-3B over the visible wavelength spectrum S01-S02 using Rayleigh methodology (supported by DIMITRI). The Rayleigh method is applicable over oligotrophic ocean with low cloud cover. The mini-file products over six regions are collected by the OPT-MPC operators and provided to the ESLs for analysis.

The dataset ingested to DIMITRI then automated cloud-screening is performed using the Globcarbon algorithm (available in DIMITRI).

Then the TOA radiance/reflectance, the sun and viewing angles, cloud-mask and auxiliary variables are stored for each pixel, and quick-looks are generated.

Figure 5 and Figure 6 display the results of Rayleigh methodology averaged over the ocean Cal/Val test sites. The results are consistent between the different test sites. The results seem sufficiently reliable to provide an estimate of the absolute vicarious calibration coefficients for both SLSTR-A and SLSTR-B (see STD values in Table 16).

6.2.2.2.3 Results of the validation over glint methodology

This activity is dedicated to the assessment the inter-band calibration of the radiometry measurements of SLSTR on board of Sentinel-3A and Sentinel-3B over the NIR-SWIR wavelength spectrum S03, S05 and S06 using S02 as reference spectral band (supported by DIMITRI). The Glint method is applicable over oligotrophic ocean with low cloud cover and high specular reflection of the sun on the ocean surface.

The mini-file products over six ocean regions are collected by the OPT-MPC operators and provided to the ESLs for analysis.

The dataset ingested to DIMITRI then automated cloud-screening is performed using the Globcarbon algorithm (available in DIMITRI).

Figure 125 displays the results of glint methodology averaged over the ocean Cal/Val test sites. The results are consistent between the different test sites.

Glint methodology allows the temporal monitoring of the SWIR band for both SLSTR-A and SLSTR-B (see Figure 123 and Figure 124). The relative difference of the simulated reflectance from DIMITRI for SLSTR-A is -8% and -15% for S5 and S6 in nadir view and for SLSTR-B is -7% and -18% for S5 and S6 in nadir view. The results are in good agreement with the previous results observed by RAL.
Figure 123: SLSTR-A measured TOA reflectance with respect to the simulated TOA reflectance over sun-glints using DIMITRI for the Nadir view.
Figure 124: SLSTR-B measured TOA reflectance with respect to the simulated TOA reflectance over sun-glints using DIMITRI for the Nadir view.

The synthesis of the results shows a good consistency over Rayleigh, Glint and PICS methods (Figure 125) over the VNIR spectral range from SLSTR-A and SLSTR-B over the period January - December 2022.
Figure 125: The estimated gain values for (top) S3A/SLSTR and (bottom) S3B/SLSTR from Glint, Rayleigh and PICS methods as a function of wavelength. We use the gain value of S2 from PICS method as reference gain for Sunglint method. Dashed-green and orange lines indicate the 2% and 5% respectively. Error bars indicate the methods uncertainties.
Table 16: Summary of Vicarious Radiometric Validation Results performed by ARGANS using DIMITRI. Results presented here are the ratios \( \frac{R_{\text{meas}}}{R_{\text{ref}}} \) averaged over the period January-December 2022.

Nadir View

<table>
<thead>
<tr>
<th>Band /Method</th>
<th>$S1$</th>
<th></th>
<th>$S2$</th>
<th></th>
<th>$S3$</th>
<th></th>
<th>$S5$</th>
<th></th>
<th>$S6$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rmeas/Ref</td>
<td>Uncert</td>
<td>Rmeas/Ref</td>
<td>Uncert</td>
<td>Rmeas/Ref</td>
<td>Uncert</td>
<td>Rmeas/Ref</td>
<td>Uncert</td>
<td>Rmeas/Ref</td>
<td>Uncert</td>
</tr>
<tr>
<td>SLSTR-A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rayleigh</td>
<td>1.08</td>
<td>0.02</td>
<td>1.05</td>
<td>0.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Desert-PICS</td>
<td>1.05</td>
<td>0.02</td>
<td>1.04</td>
<td>0.02</td>
<td>1.03</td>
<td>0.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glint</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.05</td>
<td>0.01</td>
<td>0.92</td>
<td>0.01</td>
<td>0.85</td>
<td>0.01</td>
</tr>
<tr>
<td>SLSTR-B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rayleigh</td>
<td>1.09</td>
<td>0.02</td>
<td>1.06</td>
<td>0.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Desert-PICS</td>
<td>1.04</td>
<td>0.02</td>
<td>1.04</td>
<td>0.02</td>
<td>1.03</td>
<td>0.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glint</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.05</td>
<td>0.01</td>
<td>0.93</td>
<td>0.01</td>
<td>0.82</td>
<td>0.01</td>
</tr>
</tbody>
</table>

6.2.2.3 Combined Results

The goal of the vicarious calibration analyses is to determine the offsets of SLSTR to a common reference that can be traced to a primary standard, and to implement these in the IPF.

In addition to the analysis performed by the MPC, independent studies by different groups have also been conducted to assess the post launch calibration of these channels. We have compared the results the analyses performed by RAL Space using comparisons with AATSR and MODIS-A over desert sites, CNES using the SADE/MUSCLE vicarious calibration system over desert sites, Rayference using a Radiative Transfer Model of the Libya-4 site, and the University of Arizona’s comparisons against in-situ field measurements of the Railroad Valley Playa RadCalNet site.

The comparisons performed by RAL and CNES have been made against other satellite sensors where there are known differences that need to be accounted for. For example, previous analyses of AATSR found systematic offsets compared to MERIS of approximately 1.03 for channels S1-S3. So, for instance, where AATSR is used as the reference for SLSTR channels S1-S3, the results are adjusted to MERIS by applying the corresponding difference reported in the literature. The analysis performed by Rayference and University of Arizona are independent of any satellite measurements and so no adjustment is needed.

For the reported uncertainties we attempt to combine the information provided using the Guide to expression of Uncertainties in Measurement (GUM). Uncertainties in the calibration factors are based on those reported by the different teams and are the best estimate at the time of writing.

Results presented in Table 17 and Figure 126 show good agreement within the reported uncertainties. We do not attempt to state which method is closest to the true value since all methods are relative to a different reference.
Table 17: Summary of Vicarious Radiometric Calibration Results performed by all groups. Comparisons are performed by comparing the measured reflectance vs. reference reflectance. Results presented here are the ratios $R_{\text{meas}}/R_{\text{ref}}$.

### Nadir View

<table>
<thead>
<tr>
<th>Method</th>
<th>S1 Rmeas/Ref</th>
<th>Uncert</th>
<th>S2 Rmeas/Ref</th>
<th>Uncert</th>
<th>S3 Rmeas/Ref</th>
<th>Uncert</th>
<th>S5 Rmeas/Ref</th>
<th>Uncert</th>
<th>S6 Rmeas/Ref</th>
<th>Uncert</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPC (RAL)</td>
<td>-</td>
<td>-</td>
<td>1.02</td>
<td>0.04</td>
<td>1.02</td>
<td>0.04</td>
<td>0.89</td>
<td>0.04</td>
<td>0.88</td>
<td>0.03</td>
</tr>
<tr>
<td>CNES</td>
<td>1.02</td>
<td>0.05</td>
<td>1.02</td>
<td>0.05</td>
<td>1.01</td>
<td>0.04</td>
<td>0.89</td>
<td>0.03</td>
<td>0.89</td>
<td>0.04</td>
</tr>
<tr>
<td>RTM (Rayference)</td>
<td>1.05</td>
<td>0.03</td>
<td>1.03</td>
<td>0.03</td>
<td>1.02</td>
<td>0.03</td>
<td>0.90</td>
<td>0.03</td>
<td>0.90</td>
<td>0.03</td>
</tr>
<tr>
<td>Railroad Valley</td>
<td>1.02</td>
<td>0.04</td>
<td>1.02</td>
<td>0.04</td>
<td>1.02</td>
<td>0.04</td>
<td>0.92</td>
<td>0.04</td>
<td>0.88</td>
<td>0.04</td>
</tr>
<tr>
<td>Median</td>
<td>1.02</td>
<td>-</td>
<td>1.02</td>
<td>-</td>
<td>1.02</td>
<td>0.90</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.03</td>
<td>0.03</td>
<td>1.02</td>
<td>0.02</td>
<td>1.02</td>
<td>0.02</td>
<td>0.90</td>
<td>0.02</td>
<td>0.89</td>
<td>0.02</td>
</tr>
<tr>
<td>Weighted Average</td>
<td>1.03</td>
<td>0.03</td>
<td>1.02</td>
<td>0.02</td>
<td>1.02</td>
<td>0.02</td>
<td>0.90</td>
<td>0.02</td>
<td>0.89</td>
<td>0.02</td>
</tr>
</tbody>
</table>

### Oblique View

<table>
<thead>
<tr>
<th>Method</th>
<th>S1 Rmeas/Ref</th>
<th>Uncert</th>
<th>S2 Rmeas/Ref</th>
<th>Uncert</th>
<th>S3 Rmeas/Ref</th>
<th>Uncert</th>
<th>S5 Rmeas/Ref</th>
<th>Uncert</th>
<th>S6 Rmeas/Ref</th>
<th>Uncert</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPC (RAL)</td>
<td>-</td>
<td>-</td>
<td>1.04</td>
<td>0.04</td>
<td>1.06</td>
<td>0.04</td>
<td>0.95</td>
<td>0.04</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CNES</td>
<td>1.03</td>
<td>0.06</td>
<td>1.04</td>
<td>0.07</td>
<td>1.04</td>
<td>0.05</td>
<td>0.95</td>
<td>0.06</td>
<td>0.89</td>
<td>0.08</td>
</tr>
<tr>
<td>RTM (Rayference)</td>
<td>1.09</td>
<td>0.03</td>
<td>1.07</td>
<td>0.03</td>
<td>1.07</td>
<td>0.03</td>
<td>0.99</td>
<td>0.03</td>
<td>0.96</td>
<td>0.03</td>
</tr>
<tr>
<td>Railroad Valley</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Median</td>
<td>1.09</td>
<td>-</td>
<td>1.04</td>
<td>-</td>
<td>1.06</td>
<td>0.95</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.06</td>
<td>0.06</td>
<td>1.05</td>
<td>0.04</td>
<td>1.06</td>
<td>0.03</td>
<td>0.96</td>
<td>0.03</td>
<td>0.92</td>
<td>0.07</td>
</tr>
<tr>
<td>Weighted Average</td>
<td>1.07</td>
<td>0.05</td>
<td>1.05</td>
<td>0.03</td>
<td>1.06</td>
<td>0.03</td>
<td>0.97</td>
<td>0.03</td>
<td>0.94</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note: Uncertainty estimates are based on the reported uncertainties at $k=1$ and do not necessarily account for all effects.
Figure 126: Summary of comparisons of SLSTR VIS/SWIR channel reflectances vs. Reference methods used to provide vicarious correction factors.

Using the combined weighted averages, we are able to provide vicarious adjustment factors to align SLSTR reflectances to MERIS and MODIS Aqua L1 calibrations, Table 18. This is on the basis that MERIS and MODIS calibrations have been assessed over many years and are considered as reference sensors in the VIS/SWIR and relative differences with other sensors are reported. Alignment to a different reference sensor, e.g. Sentinel-2 would be possible provided that relative differences and uncertainty estimates are provided. The correction factor is the inverse of the vicarious calibration results – i.e. $1/(R_{\text{meas}}/R_{\text{ref}})$. 
Table 18: Proposed VIS-SWIR Calibration Adjustments Based on Vicarious Calibration analysis. Note S4 is not included because the vicarious calibration techniques do not extend to this band.

Nadir View

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S5</th>
<th>S6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.97</td>
<td>0.98</td>
<td>0.98</td>
<td>1.11</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Input Analysis</td>
<td>UoAz Rayference</td>
<td>UoAz MPC (RAL) Rayference</td>
<td>UoAz MPC (RAL) Rayference</td>
<td>UoAz MPC (RAL) Rayference</td>
<td>UoAz MPC (RAL) Rayference</td>
</tr>
<tr>
<td></td>
<td>CNES</td>
<td>CNES</td>
<td>CNES</td>
<td>CNES</td>
<td>CNES</td>
</tr>
</tbody>
</table>

Oblique View

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S5</th>
<th>S6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.94</td>
<td>0.95</td>
<td>0.95</td>
<td>1.04</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>Input Analysis</td>
<td>Rayference</td>
<td>MPC (RAL) Rayference</td>
<td>MPC (RAL) Rayference</td>
<td>MPC (RAL) Rayference</td>
<td>Rayference</td>
</tr>
<tr>
<td></td>
<td>CNES</td>
<td>CNES</td>
<td>CNES</td>
<td>CNES</td>
<td>CNES</td>
</tr>
</tbody>
</table>

Note: Uncertainty estimates are at k=1.

The following tables report the estimated drift rate since the start of the mission for SLSTR A and B. The trend is generally lower or close to 1%, with slightly higher values observed for S1 (nadir view) and S3 (SLSTR-A).

Table 19: SLSTR-A temporal drift rates for each year since the start of the mission (in %)

<table>
<thead>
<tr>
<th>Year</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S5a</th>
<th>S5b</th>
<th>S6a</th>
<th>S6b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Na</td>
<td>Ob</td>
<td>Na</td>
<td>Ob</td>
<td>Na</td>
<td>Ob</td>
<td>Na</td>
</tr>
<tr>
<td>2017</td>
<td>0.3</td>
<td>0.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>2018</td>
<td>0.7</td>
<td>1.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.8</td>
<td>0.9</td>
<td>0.0</td>
</tr>
<tr>
<td>2019</td>
<td>0.9</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>1.1</td>
<td>1.2</td>
<td>0.0</td>
</tr>
<tr>
<td>2020</td>
<td>1.1</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>1.4</td>
<td>1.6</td>
<td>0.0</td>
</tr>
<tr>
<td>2021</td>
<td>1.4</td>
<td>0.9</td>
<td>0.6</td>
<td>0.6</td>
<td>1.6</td>
<td>1.8</td>
<td>0.0</td>
</tr>
<tr>
<td>2022</td>
<td>2.2</td>
<td>1.3</td>
<td>1.1</td>
<td>0.7</td>
<td>2.2</td>
<td>2.3</td>
<td>0.2</td>
</tr>
</tbody>
</table>
6.2.3 Geometric Calibration

The verification of the geolocation accuracy of the SLSTR Level-1 products is performed using the GEOCAL tool developed by ACS under ESTEC contract and running within the MPC. GEOCAL monitors the geolocation performance in Level-1 images by correlation of images with ground control points (GCP). GEOCAL takes into account each GCP’s pixel position, the predicted and the found direction cosines in the satellite control frame, and using the thermo-elastic quaternions, provides an estimation of the SLSTR orientation with respect to the satellite control frame in the form of boresight distortions angles, error estimates in the form of covariance matrices, and the optimal direction of each GCP.

Each Level-1 granule typically contains several hundred GCPs. Only GCPs with signal-to-noise ratio larger than 10 are taken into account to obtain a daily average of positional offsets in the across and along track directions.

Figure 127 and Figure 128 present the geolocation performance of SLSTR-A and SLSTR-B showing the average positional offsets in pixels (0.5 km) for Nadir and Oblique views since the beginning of the mission.

---

**Table 20: SLSTR-B temporal drift rates for each year since the start of the mission (in %)**

<table>
<thead>
<tr>
<th>Year</th>
<th>S1 Na</th>
<th>S1 Ob</th>
<th>S2 Na</th>
<th>S2 Ob</th>
<th>S3 Na</th>
<th>S3 Ob</th>
<th>S5a Na</th>
<th>S5a Ob</th>
<th>S5b Na</th>
<th>S5b Ob</th>
<th>S6a Na</th>
<th>S6a Ob</th>
<th>S6b Na</th>
<th>S6b Ob</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>0.3</td>
<td>0.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
<td>0.5</td>
<td>0.1</td>
<td>0.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2018</td>
<td>0.7</td>
<td>1.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.8</td>
<td>0.9</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.0</td>
<td>-0.1</td>
<td>-0.6</td>
<td>-</td>
<td>-0.7</td>
<td>-</td>
</tr>
<tr>
<td>2019</td>
<td>0.9</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>1.1</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.2</td>
<td>-</td>
<td>-0.3</td>
<td>-</td>
</tr>
<tr>
<td>2020</td>
<td>1.1</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>1.4</td>
<td>1.6</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.0</td>
<td>-0.1</td>
<td>-0.2</td>
<td>-</td>
<td>-0.3</td>
<td>-</td>
</tr>
<tr>
<td>2021</td>
<td>1.4</td>
<td>0.9</td>
<td>0.6</td>
<td>0.6</td>
<td>1.6</td>
<td>1.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.2</td>
<td>-</td>
<td>-0.3</td>
<td>-</td>
</tr>
<tr>
<td>2022</td>
<td>2.2</td>
<td>1.3</td>
<td>1.1</td>
<td>0.7</td>
<td>2.2</td>
<td>2.3</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
<td>0.0</td>
<td>0.7</td>
<td>-</td>
<td>0.6</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 127: Daily offset results from the GEOCAL Tool analysis for Nadir view along and across track (top two plots) and Oblique view along and across track (bottom two plots) for SLSTR-A.
Figure 128: Daily offset results from the GEOCAL Tool analysis for Nadir view along and across track (top two plots) and Oblique view along and across track (bottom two plots) for SLSTR-B.

The positional offset in nadir view meets the mission requirements and remained constant throughout 2022 for both SLSTR-A and SLSTR-B. The average geometric offset for SLSTR-A and SLSTR-B is within 0.1 pixel in nadir view along- and across-track and in oblique view across-track. In oblique view, the offset varies seasonally. This offset variation is well correlated with a variation in the number of ground control points observed during the year and is still within the requirements.

6.2.4 Cloud Screening

The Level 1 cloud screening monitoring continues in the fifth year of SLSTR-A operations, and fourth year of SLSTR-B. The cloud screening available within the Level 1 product consists of the basic cloud mask that uses a set of 14 different tests that combine to form the Basic ‘summary_cloud’ flag, the Bayesian, operating over ocean, and the Probabilistic, operating over land.

There is algorithm development in progress following the proposed updates and investigations in to several of the existing cloud tests, and proposals for new cloud tests. This included suggested updates to the sun-glint identification and cloud tests in the presence of glint, a new visible channel cloud test, a new 2.25 channel cloud test, and updates to the fog/low stratus test over ocean and land.
6.2.4.1 Summary of basic cloud tests

Currently, all tests but one (infrared histogram test) are included in the summary cloud flag. A short test-by-test summary is provided below.

❖ Visible (NDVI) cloud test
  ❖ The visible cloud test is a per-pixel test operating over land only. Two Normalised Differential Indices that are sensitive to vegetated and desert surfaces are calculated using the visible channels. An empirical-based look-up composed of a number of cloudy zones is used to determine if a pixel might contain cloud. There is, however, now an issue with this test missing patches of light cloud over vegetated areas. This could be fixed in the future by using land biome map in the algorithm.

❖ Fog/low stratus test
  ❖ The fog/low stratus test is a per-pixel threshold test that only operates on both land and ocean at night. It uses brightness temperature differences between the 11 µm and 3.7 µm channels to determine if there is cloud present. However, cloud can still be missed at night time. This could be improved with further parametrisation of the look-up table.

❖ Gross cloud test
  ❖ The gross cloud test identifies the coldest clouds, based on a threshold value on the 12 µm brightness temperatures. There is variation in the thresholds with latitude and season (month).

❖ Thin cirrus test
  ❖ This test analyses the BT11-BT12 vs Threshold(BT11, across-track band). It operates on each view separately. This is a reliable test. There is some dependence on atmospheric path and therefore further tuning of the LUTs to reflect this may bring small improvements.

❖ Medium high cloud test
  ❖ This test analyses BT3.7-BT12 vs Threshold(BT12). It operates on each view separately, only at night. The value of BT3.7 is always higher than BT12 due to partially cloud filled pixels and thin cirrus being present. There may be some discrepancies around twilight regions.

❖ 1.375 threshold test
  ❖ This test analyses R1.375 vs Threshold(across-track band). It is based on the high absorption from water vapour in this band, meaning any signal in this channel is likely to be from cloud.

❖ 1.6/2.25 large and small scale histogram tests
  ❖ The large-scale part of this test works on the basis that the signal received from clear-sky pixels will have a low value that has little variation, whereas any cloudy pixels will have a higher-varying bright signal. The pixels from a small area are formed into a histogram and the ‘shape’ of the low dark clear pixels is automatically identified from the brighter, wider peak of the cloudy pixels. The small-scale part of this test looks at the variability of the signal. It is intended to be used in sun-glinted regions when the large-scale test cannot be operated. These tests are not optimized for sun-glinted regions and significant cloud is still missed.
when the sun-glint flag is raised. It is recommended that an update to the algorithms be developed to counter this.

❖ Spatial coherence test
❖ This test assesses the standard deviation of the measured BTs over a small area of ocean. It is assumed that over clear sky, the signal variation will be small against the background of a homogeneous ocean. This test has a tendency to over-mask cloud and is one of the priorities for algorithm development.

❖ Infrared histogram test
❖ This test uses the 11 µm brightness temperature to identify cloud that all other tests may have missed. This is not a reliable test and when used in AATSR, was often seen to falsely classify clear-sky as cloud. It is rarely set. This test is not yet included in the summary cloud.

6.2.4.2 Summary of Bayesian test

The Bayesian mask is carried through to Level-2 Marine Products, and is currently also provided in the Level 1 product. The Bayesian cloud screening method makes use of measurements in the S2, S3, S5, S8 and S9 channels during the day, and S7, S8 and S9 channels at night. These are compared to radiative transfer modelling and pre-calculated look-up tables to infer the probability of a pixel being cloudy given the observations and background meteorological state. The method has previously been applied successfully in the context of the ESA SST CCI to the AVHRR and other ATSR instruments.

6.2.4.3 Summary of Probabilistic tests

The Probabilistic Cloud Mask is implemented in the IPF at Level-1 and carried through to Level-2.

6.2.4.4 Monitoring cloud masking performance

No monitoring results are available for the reporting period.

6.2.5 References


6.3 L2 product performances

6.3.1 Land Surface Temperature (LST)

6.3.1.1 Validation approach

The formal missions’ requirement for LST specifies that:

❖ S3-MR-420: Sentinel-3 shall be able to measure Land Surface Temperature (LST) to an accuracy of < 1K with a resolution of 1 km at nadir. This capability shall not reduce the quality of the SST retrievals.

A four-phase approach is detailed in the Cal/Val plan, which follows both the ESA LST validation protocol (Schneider et al., 2012) and the CEOS LPV Best Practices guide for LST (Guillevic et al., 2017):

❖ Comparison of satellite-retrieved LST with in situ measurements collected from radiometers sited at a number of stations spread across the Earth, for which the highest-quality validation can be achieved;

❖ Radiometric-based validation, which offers an alternative to validation with in situ LST measurements as it does not require measurements of LST on the ground, and can provide a viable alternative for long-term, semi-operational LST product evaluation at the global scale;

❖ Inter-comparisons with similar LST products from other sources such as AATSR, AVHRR, MODIS, SEVIRI, and VIIRS, which give important quality information with respect to spatial patterns in LST deviations;

❖ Time series analysis to quantify trends and to identify potential instrument drift or persistent cloud contamination.

We have focussed here only on the first approach. This responds directly to the formal mission requirements on accuracy for LST.

The SLSTR-A SL_2_LST product from SLSTR went operational in the Sentinel 3 PDGS on 5th July 2017 with PB 2.16. No additional updates to the retrieval algorithm have been implemented in the IPF since. However, Processing Baseline 2.29 released on 4th April 2018 included the new Probabilistic Cloud Mask implemented in the IPF at Level-1 and carried through to Level-2. Furthermore, from 26th February 2019 an updated ADF of retrieval coefficients has been implemented in PB 2.47, IPF 06.14. We show results on a monthly basis from 1st January 2022 to 31st December 2022. In all cases the Probabilistic Cloud Mask is applied. An improvement to the cloud coefficients ADF was made on 23rd October 2020 in PB 2.73.

The SLSTR-B SL_2_LST product from SLSTR went operational in the Sentinel 3 PDGS on 26th February 2019 with PB 1.19 IPF 06.14. We show results on a monthly basis from 1st January 2022 to 31st December 2022. In all cases the Probabilistic Cloud Mask is applied. An improvement to the cloud coefficients ADF was made on 23rd October 2020 in PB 1.50.

For both SLSTR-A and SLSTR-B all matchups have been performed for non-time critical (NTC) only since this is deemed to be the data of highest quality.
For the in situ validation ten “Gold Standard” stations were used in the matchups process, seven from the SURFRAD network; two from the ARM network; and one from the UOL network: i) Bondville, Illinois; ii) Desert Rock, Nevada; iii) Fort Peck, Montana; iv) Goodwin Creek, Mississippi; v) Penn State University, Pennsylvania; vi) Sioux Falls, South Dakota; vii) Table Mountain, Colorado; viii) Southern Great Plains, Oklahoma; ix) Barrow, North Slopes Alaska; x) Kanpur, India. Overall the matchups show good agreement between the satellite LST and the in situ LST across a broad range of LST values. This is the case for each of the “Gold Standard” stations (Figure 129 – SLSTR-A; Figure 130 – SLSTR-B).
Figure 129: In situ validation of S3A SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022. 1st row: Bondville (left), Desert Rock (right); 2nd row: Fort Peck (left), Goodwin Creek (right); 3rd row: Penn State (left), Sioux Falls (right); 4th row: Table Mountain (left); Southern Great Plains (right); 5th row: North Slopes Alaska (left), Kanpur (right).
Figure 130: In situ validation of S3B SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022. 1st row: Bondville (left), Desert Rock (right); 2nd row: Fort Peck (left), Goodwin Creek (right); 3rd row: Penn State (left), Sioux Falls (right); 4th row: Table Mountain (left); Southern Great Plains (right); 5th row: North Slopes Alaska (left), Kanpur (right).

The statistics are shown in Table 21 (SLSTR-A) and Table 22 (SLSTR-B). The accuracy can be directly compared with mission requirement S3-MR-420. For SLSTR-A, overall the absolute daytime accuracy is 0.87 K and the absolute night-time accuracy is 0.47 K. Both of which are within the mission requirements for LST. For SLSTR-B, overall the absolute daytime accuracy is 1.02 K and the absolute night-time accuracy is 0.46 K. Both of which are within or very close to the mission requirements for LST.
Table 21: Statistics of In situ validation for S3A SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022

<table>
<thead>
<tr>
<th>Network</th>
<th>Site</th>
<th>Day</th>
<th></th>
<th>Night</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SURFRAD</td>
<td>Bondville</td>
<td>1.80</td>
<td>3.44</td>
<td>0.33</td>
<td>1.05</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Table Mountain</td>
<td>-0.03</td>
<td>3.38</td>
<td>-0.50</td>
<td>1.05</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Desert Rock</td>
<td>0.08</td>
<td>1.57</td>
<td>-0.30</td>
<td>1.54</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Fort Peck</td>
<td>-0.25</td>
<td>2.45</td>
<td>0.09</td>
<td>1.45</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Goodwin Creek</td>
<td>0.10</td>
<td>2.03</td>
<td>0.02</td>
<td>1.39</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Penn State University</td>
<td>0.59</td>
<td>1.91</td>
<td>1.53</td>
<td>0.97</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Sioux Falls</td>
<td>0.40</td>
<td>2.54</td>
<td>0.10</td>
<td>1.55</td>
</tr>
<tr>
<td>ARM</td>
<td>Southern Great Plains</td>
<td>-2.13</td>
<td>2.43</td>
<td>-0.39</td>
<td>1.07</td>
</tr>
<tr>
<td>ARM</td>
<td>North Slopes Alaska</td>
<td>-2.59</td>
<td>5.31</td>
<td>-0.87</td>
<td>3.65</td>
</tr>
<tr>
<td>UOL</td>
<td>Kanpur, India</td>
<td>0.76</td>
<td>3.83</td>
<td>0.61</td>
<td>1.46</td>
</tr>
</tbody>
</table>

Table 22: Statistics of In situ validation for SL_2_LST product at ten “Gold Standard” stations for the period 1st January 2022 to 31st December 2022

<table>
<thead>
<tr>
<th>Network</th>
<th>Site</th>
<th>Day</th>
<th></th>
<th>Night</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SURFRAD</td>
<td>Bondville</td>
<td>2.21</td>
<td>4.30</td>
<td>-0.09</td>
<td>0.90</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Table Mountain</td>
<td>-0.28</td>
<td>3.34</td>
<td>-0.40</td>
<td>1.32</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Desert Rock</td>
<td>0.51</td>
<td>1.82</td>
<td>0.43</td>
<td>2.42</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Fort Peck</td>
<td>0.29</td>
<td>2.91</td>
<td>0.37</td>
<td>1.38</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Goodwin Creek</td>
<td>0.04</td>
<td>1.30</td>
<td>0.41</td>
<td>1.22</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Penn State University</td>
<td>0.69</td>
<td>1.94</td>
<td>1.39</td>
<td>0.90</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Sioux Falls</td>
<td>0.77</td>
<td>2.39</td>
<td>0.28</td>
<td>1.39</td>
</tr>
<tr>
<td>ARM</td>
<td>Southern Great Plains</td>
<td>-1.83</td>
<td>2.00</td>
<td>0.00</td>
<td>0.85</td>
</tr>
<tr>
<td>ARM</td>
<td>North Slopes Alaska</td>
<td>-2.96</td>
<td>5.36</td>
<td>-0.86</td>
<td>3.66</td>
</tr>
<tr>
<td>UOL</td>
<td>Kanpur, India</td>
<td>0.64</td>
<td>4.14</td>
<td>0.37</td>
<td>1.76</td>
</tr>
</tbody>
</table>
6.3.1.2 Validation metrics and visualisation

The following metrics are calculated for day and night-time overpasses for each station:

❖ **Accuracy**: the degree of conformity of the measurement of a quantity to the accepted value or the “true” value, defined here as the median bias between the Sentinel-3 and in-situ LST for all matchup pairs [K].

❖ **Precision**: the closeness of agreement between independent measurements of a quantity under the same conditions, defined here as the robust standard deviation of the bias between the Sentinel-3 and in-situ LST for all matchup pairs [K].

The validation results for each site are visualised as scatter plots of Sentinel-3 and in-situ LST, using a 1:1 line to illustrate deviation from the “true” value. Separate plots are created for Sentinel-3A and 3B, with day and night matchup pairs shown together on the same plots.

In addition to this, this work also validates the uncertainty associated with the satellite LST. The standard deviation of the satellite – in-situ LST bias is compared with the total uncertainty of the matchup pairs, and a goodness-of-fit statistic is calculated based on this comparison. The total satellite product uncertainty for each associated matchup pair ($\sigma_{total}$) is determined using the following relation:

$$\sigma_{total} = \sqrt{\sigma_{sat}^2 + \sigma_{in-situ}^2 + \sigma_{space}^2 + \sigma_{time}^2}$$

*Equation 1: Estimation of the total satellite product uncertainty for each Sentinel-3 – in-situ LST matchup pair. Equation 17 from [RD – 13]*

$\sigma_{total}$ is calculated from four components:

❖ **$\sigma_{sat}$**: the total LST uncertainty for the Sentinel-3 pixel, as reported by the “LST_uncertainty” variable in the “satellite_data” group in the matchup file.

❖ **$\sigma_{in-situ}$**: the uncertainty associated with the in-situ LST (0.5 K, as reported by the “LST_uncertainty” variable in the “insitu_data” group in the matchup file.

❖ **$\sigma_{space}$**: the uncertainty associated with matching a satellite and in-situ observation in a spatial context. In practice, this is calculated as the standard deviation of the LSTs reported by a 5 x 5 ground pixel grid centred on the in-situ station/alternative location (see above). It should be noted that $\sigma_{space}$ (and so $\sigma_{total}$) is only calculated for matchup pairs where at least 80% of the ground pixels in the 5 x 5 grid were not cloudy, cosmetically filled, or otherwise unusable.

❖ **$\sigma_{time}$**: the uncertainty associated with matching a satellite and in-situ observation in time. Because the available in-situ LST measurements were always within 1 minute of the satellite overpass, it was assumed that $\sigma_{time}$ was negligible.

The estimated $\sigma_{total}$ for all matchup pairs for a single satellite platform was binned to 0.1 K bins, and was compared against the standard deviation of the corresponding satellite – in-situ LST bias. The reduced chi-squared statistic was calculated to determine whether $\sigma_{total}$ was a reliable estimate of the observed LST uncertainty. A chi-squared value < 1.0 would indicate that $\sigma_{total}$ was an overestimate, while a chi-squared value > 1.0 would suggest that $\sigma_{total}$ was an underestimate.

This work also determines whether a long-term trend exists in the observed satellite – in-situ LST bias. For each satellite platform, the biases of all daytime and night-time matchup pairs was plotted as a time...
series, and a linear function was fitted using weighted least squares, where the uncertainty of the bias was used as observational error. The statistical significance of the fitted trend was established using Student’s t-test.

### 6.3.1.3 Results

This section details the results of the validation of Sentinel-3A and 3B for each in-situ site. The metrics calculated for each site are summarised in Table 23. Figure 131 - Figure 145 show the validation results for each site. The validation results for each site are discussed herein.

**Table 23: The results of the validation of Sentinel-3A and 3B LST using in-situ measurements.**

<table>
<thead>
<tr>
<th>Site name</th>
<th>Sentinel-3A</th>
<th></th>
<th>Sentinel-3B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day</td>
<td>Night</td>
<td>Day</td>
<td>Night</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Acc</td>
<td>Prec</td>
<td>N</td>
</tr>
<tr>
<td>Svarberget</td>
<td>79</td>
<td>-0.508</td>
<td>0.988</td>
<td>82</td>
</tr>
<tr>
<td>Hyytiälä</td>
<td>53</td>
<td>-0.636</td>
<td>0.522</td>
<td>65</td>
</tr>
<tr>
<td>KIT forest site</td>
<td>123</td>
<td>0.184</td>
<td>0.686</td>
<td>171</td>
</tr>
<tr>
<td>Robson Creek</td>
<td>19</td>
<td>0.286</td>
<td>1.332</td>
<td>61</td>
</tr>
<tr>
<td>Puéchabon</td>
<td>71</td>
<td>0.742</td>
<td>0.978</td>
<td>81</td>
</tr>
</tbody>
</table>

#### 6.3.1.3.1 Svarberget

Analysis of the matchup pairs for this site was hampered by the limited amount of data available; at the time of this work, only data from October 2021 – October 2022 was ingested into the MDB. That being said, Figure 131 shows that there are roughly equivalent night-time and daytime overpasses for both Sentinel-3A and 3B, suggesting that there may be no coverage bias from cloud masking for both satellite platforms over this site.

Both Sentinel-3A and 3B show good agreement with the in-situ data, with accuracy below the 1 K threshold set by [RD – 1]. However, the precision for Sentinel-3A slightly exceed the threshold set by [RD – 12] for night-time matchup pairs, while Sentinel-3B daytime matchup pairs also slightly exceed this threshold. It is possible that ground pixel coverage of heterogeneous scenes may introduce some additional variability in the matchup pairs, which may result in a larger than expected spread. However, this would need to be confirmed with more in situ observations around the Svarberget site.

It should also be noted that from mid-June – 7th July 2022 there was an unexplained positive trend in the brightness temperature observed by the skyward pointing radiometer at this site. This trend was thought to be caused by water contamination of the radiometer lens. It is currently thought that this effect adds an additional ~0.2 K uncertainty to the observed LST. Consequently, extra caution is required when interpreting the results from this site, and observations from this time period and subsequent months may need to be flagged and removed from the analysis in future iterations of this work.
The Hampel filtering removed ~28% of the filtered ground pixels, suggesting the cloud detection algorithm, and corresponding cloud coefficients are also performing well. These results overall suggest that the SL_2_LST retrieval algorithm is performing well over this site.

Figure 131: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Svartberget between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively.

Figure 132 shows the results of the LST uncertainty validation. The reduced chi-squared value reported for both Sentinel-3A and 3B exceeds 1.0, indicating that the estimated total uncertainty is on average an underestimate of the observed satellite – in-situ bias. However, it should be noted that the \( \sigma_{\text{total}} \) bands which underestimated the observed bias mostly had < 10 observations binned, which suggests that a lack of observations led to a biased estimate for the standard deviation. Indeed, Figure 132 shows that \( \sigma_{\text{total}} \) bands which had > 10 observations were in agreement with the observed satellite – in-situ bias. It is possible that more observations in this analysis would lead to a better agreement between the observed bias and \( \sigma_{\text{total}} \).

Figure 132: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Svartberget between October 2021 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B.
Figure 133 shows the time series in the observed satellite – in-situ bias for both Sentinel-3A and 3B. No statistically significant linear trend (i.e. p < 0.05) exists for either Sentinel-3A or 3B matchup pairs, indicating that the biases are stable over time. However, no night-time matchup pairs exist between May – August 2022, which may indicate a coverage bias for summer months. Further observations in subsequent years would be required to confirm if this is the case.

![Figure 133: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Svartberget between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup pair added together in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot.](image)

6.3.1.3.2 Hyytiälä

As with Svartberget, analysis of Hyytiälä matchup pairs is limited by the MDB ingesting only data between October 2021 – October 2022. Figure 134 shows that both Sentinel-3A and 3B matchup pairs contain ~20% more night-time than daytime observations, which may be evidence of a coverage bias for this site. Analysis of both Sentinel-3A and 3B matchup pairs shows that the retrieved LSTs are either below or very close to the accuracy and precision requirements set by [RD – 1] and [RD – 12], though the night-time Sentinel-3A observations very slightly exceed the < 1 K threshold for accuracy (see Figure 134). Day-time observations have better accuracy and precision than night-time observations, with both time periods exhibiting slightly negative biases compared to the in-situ data.

The Hampel filtering removed ~36% of all observations, suggesting the cloud detection algorithm, and corresponding cloud coefficients are also performing well. Overall these results suggest that the SL_2_LST algorithm is performing well for this site.
As shown in Figure 135, the LST uncertainty analysis showed generally good agreement between the estimated $\sigma_{\text{total}}$ values and the reported LST uncertainty, for most observations. The reduced chi-squared values are > 1.0 for both Sentinel-3A and 3B, indicating that the estimated $\sigma_{\text{total}}$ values are underestimates of the observed satellite–in-situ uncertainty. However, Figure 20 shows that most of the bands with > 5 observations binned fit neatly within the bounding cone, which suggests that poor sampling of the data is the main cause of binned $\sigma_{\text{total}}$ values not matching the observed uncertainty.

Indeed, the vast majority of matchup pairs shown in Figure 134 were not included in this analysis, as insufficient surrounding cloud-free or not cosmetically filled ground pixels were available to calculate $\sigma_{\text{space}}$ for many matchup pairs. Either more observations, or a looser criteria for calculating a reliable $\sigma_{\text{space}}$ value will be required to confirm these results.

---

**Figure 134**: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Hyytiälä between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively.

**Figure 135**: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Hyytiälä between October 2021 – October 2022. The standard deviation of the satellite–in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B.
Figure 136 shows the time series in the observed satellite – in-situ bias for both Sentinel-3A and 3B. No statistically significant linear trend exists for either Sentinel-3A or 3B matchup pairs, indicating that the biases are stable over time.

![Figure 136: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Hyytiälä between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot.]

6.3.1.3.3 KIT forest site

Unlike the other sites, matchup data from a much longer time period was available from the KIT forest site for both Sentinel-3A and 3B (August 2020 – October 2022). Figure 137 shows ~39% more night-time than daytime observations for both satellite platforms, potentially indicating a coverage bias over this site.

Both Sentinel-3A and 3B appear to agree very well with the in-situ data, with both day and night-time precision and accuracy being well within the requirements set in [RD–1] and [RD–12]. The precision is below the 1 K threshold set in [RD–12], indicating no impact of site variability or cloud misclassification. Overall, the site-specific coefficients of the SL_2_LST algorithm appear to be working well.

The Hampel filtering removed ~27% of all observations suggesting the cloud detection algorithm, and corresponding cloud coefficients are also performing well. Overall these results remain encouraging, and since the matchup pairs cover more than one year they provide good evidence of the seasonal performance of the LST retrieval algorithm for this site.
Figure 137: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from the KIT forest site between August 2020 – October 2022. Red and blue points indicate day and night-time overpasses, respectively.

Figure 138 shows the results of the LST uncertainty analyses. A reduced chi-squared value of < 1.0 was reported for both Sentinel-3A and 3B, indicating that $\sigma_{\text{total}}$ was overall an overestimate of the true satellite – in-situ LST uncertainty. Despite this, both plots show that the vast majority of matchup pairs reported $\sigma_{\text{total}} < 1.5$ K, which agreed well with the standard deviation of the satellite – in-situ bias. Because of the large number of matchup pairs available, it is likely that these findings are robust.

However, the $\sigma_{\text{total}}$ distributions for both the Sentinel-3A and 3B matchups have long tails, indicating that a small minority of matchup pairs have $\sigma_{\text{total}}$ values > 1.5 K which greatly overestimate the true satellite – in-situ bias.

Figure 138: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from the KIT forest site between August 2020 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B.
One possible reason for this distribution is that the forest measured by the site is surrounded by urban areas. The $5 \times 5$ grid may therefore contain heterogeneous land cover from other biomes which would in turn would bias the comparison with the in-situ data. In order to investigate this, the uncertainty analysis was repeated but with $\sigma_{\text{space}}$ this time being calculated using a $3 \times 3$ instead of a $5 \times 5$ grid. Figure 139 shows that while this approach removes some of the extreme outliers, the long tail in the $\sigma_{\text{total}}$ distributions persists. Consequently, the disagreement between $\sigma_{\text{total}}$ and the observed biases may be caused by other issues, such as incomplete cloud masking or the residual influence of the surrounding heterogeneous land cover.

![Figure 139: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from the KIT forest site as in Figure 138, but using a $3 \times 3$ grid to calculate $\sigma_{\text{space}}$ in Equation 1.](image)

Figure 140 shows the time series in the observed satellite – in-situ bias for both Sentinel-3A and 3B. The long multi-year continuity of measurements offered by this site provides robust trend estimation. Consequently, a statistically significant trend of $-8.07 \times 10^{-4} \text{ K day}^{-1}$ (or -0.024 and -0.016 K month$^{-1}$) was present in the daytime matchup pairs for Sentinel-3A and 3B, respectively. For night-time matchup pairs, a statistically significant trend of $4.70 \times 10^{-4} \text{ K day}^{-1}$ (or 0.014 K month$^{-1}$) was reported for Sentinel-3A, while no statistically significant trend was reported for Sentinel-3B. Despite their statistical significance, the trends reported by these results suggest that the long-term trend in both Sentinel-3A and 3B is very small, and so the performance of the SL_2_LST product remains encouraging.

![Figure 140: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over the KIT forest site between August 2020 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot.](image)
6.3.1.3.4 Robson Creek

As with the other sites excluding KIT, only data from October 2021 – October 2022 was available to be ingested into the MDB. Many of the matchup pairs for both Sentinel-3A and 3B exhibited very large satellite – in-situ LST biases, which is anomalous for a spatially homogeneous site. Also, Figure 141 shows that approximately 3 times as many night-time as daytime matchup pairs exist for both Sentinel-3A and 3B after Hampel filtering.

The reason for this coverage bias is twofold; between May – October 2022 almost no cloud-free daytime overpasses were observed from either Sentinel-3A or 3B. As well as this, the night-time observations during this time period decreased the overall standard deviation of the satellite – in-situ bias for this site. As a result, the Hampel filter here for removing potentially contaminated matchup pairs was stricter than in prior months. Therefore, not only was no new daytime matchup pairs added to the MDB during this time period, but previously allowed daytime matchup pairs were now being removed from the analyses, therefore causing a reduction in available data compared to previous months.

The filtered matchup pairs for Sentinel-3A and 3B both show biases which are below the mission requirement for accuracy from [RD – 1]. Similarly, the robust standard deviations for Sentinel-3A and 3B meet the < 1 K precision criteria set in [RD – 12], as the strict Hampel filter resulted in many of the matchup pairs with significantly large satellite – in-situ biases to be removed, which has reduced the overall spread of the data. It may be possible that future months may introduce more cloud-free daytime observations over this site, which would once again introduce the anomalously high precision values seen in previous months.

There are several possible explanations for the large biases observed for this site: i) the biome coefficients are not well tuned for this biome; ii) failures in the cloud masking; or iii) non-optimal setting of the across-track parameters in the algorithm for this biome which have higher impact in high water vapour regions.

The results from the matchup pairs could suggest; i) is unlikely since coefficient representativeness is based around the emissivity and would more generally lead to systematic bias, which is not evident. Both the probabilistic and basic cloud masks were tested and both performed similarly suggesting it is not an issue with any one particular cloud test as postulated in ii), but the region is very cloudy meaning both cloud tests could have some misses. Point iii) is a possibility and the in-situ data for this biome allows for a data-driven investigation of the parameterisation for this biome.

As with Svartberget, the lens of the skyward-facing radiometer at this site was also found to be contaminated with water, which had a similar effect on the observed atmospheric brightness temperatures and retrieved LST between early January – May 12th 2022. Therefore, observations from this and subsequent months may need to be flagged and removed from the analysis.
The comparisons between $\sigma_{\text{total}}$ and the standard deviation of the satellite – in-situ bias are shown in Figure 142. The analyses report a reduced chi-squared value of 0.82 for Sentinel-3A and 1.34 for 3B, suggesting that $\sigma_{\text{total}}$ is slightly overestimating the true satellite – in-situ bias for Sentinel-3A, while being an underestimate for Sentinel-3B. However, as before it should be noted that only a very small proportion of the matchup pairs used in Figure 141 are used here, because not enough surrounding cloud-free satellite ground pixels were available to reliably estimate $\sigma_{\text{space}}$. As well as this, the aforementioned stricter Hampel filter has removed many of the more problematic observations from this analysis, compared with previous months.

For both Sentinel-3A and 3B, $\sigma_{\text{total}}$ bands with more than 3 observations binned ($\sigma_{\text{total}} < 1.2$ K) agreed very well with the observed bias. This may indicate that the inclusion of more observations with a stricter Hampel filter will result in more robust standard deviations and so a better agreement between $\sigma_{\text{total}}$ and the observed bias.

![Figure 141: Validation of the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Robson Creek between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively.](image)

![Figure 142: Validation of the satellite LST uncertainty estimated by the SL_2_LST product for Sentinel-3A (left) and Sentinel-3B (right) against LST observations from Robson Creek between October 2021 – October 2022. The standard deviation of the satellite – in-situ LST bias is plotted against bands of the total uncertainty calculated by Equation 1. The bounding cone is distorted for near-zero total uncertainties because of the inherent calibration uncertainty of the in-situ instruments (0.5 K). The reduced chi-squared goodness-of-fit statistic is given for both Sentinel-3A and 3B.](image)
Figure 143 shows the time series in the observed satellite – in-situ bias for both Sentinel-3A and 3B. At first glance it appears that a positive linear trend exists for daytime matchup pairs for both satellite platforms. However, no statistically significant trend was found for either satellite platform, as there are too few matchup pairs to infer a robust trend. Additionally, a statistically significant trend of \(-2.40 \times 10^{-3}\) K day\(^{-1}\) (or \(-0.072\) K month\(^{-1}\)) was observed for Sentinel-3B night-time matchup pairs, though no statistically significant trend was observed for Sentinel-3A. Given the aforementioned issues with this site, more observations are required to ascertain whether a trend in bias genuinely exists. In spite of this, the results are still very encouraging overall.

![Figure 143: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Robson Creek between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot.](image)

6.3.1.3.5 Puéchabon

Only matchup pairs of data between October 2021 – October 2022 were available at the time of this work for Puéchabon. Figure 144 shows slightly more night-time than daytime observations exist for both Sentinel-3A and 3B, which may indicate a night-time coverage bias.

For both day-time and night-time observations, both Sentinel-3A and 3B appear to largely meet the [RD – 1] accuracy and [RD – 12] precision criteria. Sentinel-3B also shows a slight negative bias for night-time observations, while the day-time bias for both satellite platforms is positive. Also, Sentinel-3A observations show a larger spread than those for 3B for daytime observations, while the reverse is true for night-time observations. Overall, these results continue to show good performance of the SL_2_LST product over this site.

The Hampel filtering removed \(~26\%\) of all observations suggesting the cloud detection algorithm, and corresponding cloud coefficients are also performing well. Overall, these results indicate good performance is being achieved over this site, and remain encouraging.
As shown in Figure 145, the uncertainty analysis shows that $\sigma_{\text{total}}$ is overestimating the observed satellite – in-situ uncertainty, particularly for Sentinel-3B (reduced chi-squared $< 1.0$). However, Sentinel-3A matchup pairs where $\sigma_{\text{total}} < 1.5$ K appear to fit neatly within the bounding cone, suggesting that only a minority of matchup pairs are overestimates of the observed satellite – in-situ bias.

Again, very few matchup pairs are binned for most $\sigma_{\text{total}}$ bands due to inadequate surrounding cloud-free ground-pixels being available to calculate $\sigma_{\text{space}}$ for many matchup pairs, which may lead to inaccurate standard deviations of the satellite – in-situ bias being reported here.

Figure 146 shows the time series in the observed satellite – in-situ bias for both Sentinel-3A and 3B. For Sentinel-3A, a statistically significant trend of $4.10 \times 10^{-3}$ K day$^{-1}$ (or 0.12 K month$^{-1}$) for both daytime and
night-time matchup pairs. While Sentinel-3B daytime matchup pairs reported no statistically significant trend, night-time matchup pairs reported a statistically significant trend of $3.67 \times 10^{-3}$ K day$^{-1}$ (or 0.11 K month$^{-1}$). The trends reported here are the largest observed over all 5 sites, and suggest that both satellite platforms are subject to a positive long-term trend over this site which requires rectifying.

![Figure 146: Time series for the satellite – in-situ LST bias for Sentinel-3A (left) and Sentinel-3B (right) observed over Puéchabon between October 2021 – October 2022. Red and blue points indicate day and night-time overpasses, respectively. The error bars are the in-situ and satellite LST uncertainties for each matchup added in quadrature. The linear trend fitted to the biases is represented in all the plots by a dotted line, and the gradient and associated p-value is quoted in the legend of each plot.](image)

### 6.3.1.4 Conclusions

The mean absolute accuracy for Sentinel-3 is as follows:

- **Sentinel-3A**:
  - **Day**: Absolute Accuracy = 0.471 K
  - **Night**: Absolute Accuracy = 0.440 K

- **Sentinel-3B**:
  - **Day**: Absolute Accuracy = 0.557 K
  - **Night**: Absolute Accuracy = 0.585 K

Overall, both Sentinel-3A and Sentinel-3B appear to be within or very close to the mission accuracy requirement set in [RD – 1] for all of the LAW sites for both day and night-time observations. These results are consistent with the performance of the SL_2_LST product across multiple sites as reported in the Data Quality Reports.

Svartberget and Puéchabon report roughly the same number of day and night-time observations, suggesting that no significant differences in day or night coverage exist for these sites. However, KIT forest, Robson Creek, and Hyytiälä report greater night-time than daytime matchup pairs for both Sentinel-3A and 3B observations, which suggest that a coverage bias may be present for these sites. In particular, Robson Creek reported virtually no cloud-free overpasses after May 2022. This coverage bias could be due to stricter daytime cloud masking in the SL_2_LST algorithm, which may need addressing in future algorithm development.
While overall performance is good, both Sentinel-3A and 3B have mixed performance when considering individual sites. For instance, the largest biases observed were for daytime Sentinel-3B overpasses over Robson Creek, which caused the accuracy to exceed the [RD – 1] threshold. Such biases may be addressed, if they persist over the course of following months, by updating the SL_2_LST retrieval algorithm site-specific coefficients. Overall though, the results for these sites are good and provide further confidence on the quality of the SL_2_LST product for Sentinel-3A and 3B.

The results from Robson Creek were anomalous, as the matchup pairs greatly exceeded the precision threshold prior to Hampel filtering. This is in contrast to the expectations given the homogeneity of the land cover surrounding the site. Further analysis will be performed to determine whether these biases are the result of cloud contamination affecting retrieval accuracy, data processing issues with the in situ or matchup data, or parameterisation of the water vapour for sites with very high water vapour.

As well as this, the recently discovered contamination of the skyward-facing radiometer at both Svarterberget and Robson Creek may also contribute to these anomalous results. Therefore, extra caution is required when interpreting the results from these sites. Identification and flagging of problematic LST observations due to this effect will be carried out in future iterations of this work, and may affect the final conclusions drawn from these analyses.

However, between May – October 2022 only cloud-free night-time overpasses were observed over Robson Creek. These observations had similar small satellite – in-situ biases, which caused a reduction in the Hampel filter threshold. As a result, the number of passable daytime overpasses for both Sentinel-3A and 3B was greatly reduced, resulting in a significant night-time coverage bias to be observed compared to previous reports. However, the stricter Hampel filter also caused the overall accuracy and precision for this site to greatly improve for both satellite platforms, to the point where Robson Creek no longer has anomalous results compared to other sites. It is possible that future cloud-free observations over this site may reintroduce the atypical behaviour observed in prior months, for which further analysis would still be required.

Time series analysis of the satellite – in-situ biases found no statistically significant trend for Svarterberget and Hyytiälä. KIT forest reported statistically significant yet minor trends for both Sentinel-3A and 3B, while Robson Creek reported a statistically significant trend of a similarly small magnitude for Sentinel-3B night-time matchup pairs only. However, a statistically significant trend of 0.11 K month$^{-1}$ was observed over Puéchabon for both satellite platforms. Such trends may be caused by drifts in instrument calibration or changes in the cloud masking algorithm over time; future SL_2_LST retrieval algorithm development for this biome will need to investigate and rectify these causes.

It must be noted that with the exception of the KIT forest site, the matchup pairs for all sites used in this work came from only 1 year of data, between October 2021 – October 2022. Therefore, the seasonality of these biases could not be assessed. Additionally, these preliminary statistics may not be robust, and more data may be required to make confident recommendations on what, if any, improvements should be made to the Sentinel-3 LST product.

This work includes an attempt at validating the SL_2_LST uncertainty budget for all sites, introduced in [RD – 13]. While the analyses for Hyytiälä and the KIT forest site suggest that the uncertainty budget is largely accurate for those sites, the results from the other sites indicate that the uncertainty budget is either underestimating (i.e. Svarterberget) or overestimating (i.e. Puéchabon) the observed satellite – in-situ uncertainty.
However, a key issue with the uncertainty validation attempt was that in order to accurately estimate the spatial matching uncertainty ($\sigma_{\text{space}}$) it was necessary to only analyse matchup pairs where a suitable number of surrounding cloud-free or good-quality satellite ground pixels existed. This led to very few matchup pairs being used to compute the observed retrieval uncertainty and reduced chi-squared metrics, so these uncertainty validation results are also not robust. For KIT forest, long tails were also observed in the $\sigma_{\text{total}}$ distributions, which may be due to the influence of the surrounding heterogeneous land cover. Future iterations of this work using more data will hopefully provide more reliable analyses of the uncertainty budget accuracy.
6.3.2 Fire Radiative Power (FRP)

6.3.2.1 Validation approach

The SLSTR FRP validation uses inter-comparisons with similar FRP products from other sources such as other satellite sensors, which give important quality information with respect to active fire detection and fire clusters characterisation. Here we compare the SL_2_FRP product from both SLSTR-A and SLSTR-B with the operational MODIS MOD14 FRP product (from MODIS Terra) available from the LAADS DAAC. It is important to note that the employed products have slightly different overpass times, implying that the two sensors do not observe fires in the exact same configuration nor with the same atmospheric conditions. Thus, for these reasons, and for the nature of the procedure delineated below, this inter-comparison should not be interpreted as a full validation exercise but rather as a check of the consistency of the FRP products derived from SLSTR with the ones from MODIS.

The inter-comparison procedure is divided into two main parts:

The first part is related to omission and commission fire pixels, i.e., fire pixels detected by MODIS without any SLSTR fire pixel in a 7x7 window around it (omissions), and fire pixels detected by SLSTR without any MODIS fire pixel in a 7x7 window around it (commissions).

The second part is related to the characterisation of fire clusters, i.e., groups of one or more pixels spatially adjacent to each other and corresponding to a single fire.

An improved algorithm – called FRP V2 algorithm – has been developed in 2021 by KCL. This algorithm is now performing also Fire detection over daytime S7-saturated scene by creating a combined F1/S7 channel.

This improved algorithm has been deployed in operations on 28th February. This analysis has then been performed between 1st March 2022 and 31st December 2022. This analysis is performed over five areas of high fire activity, re-evaluated every 3 months, where fire pixels are aggregated and compared. All areas of interest defined in 2022 are: the Amazonia forest (Brazil, Bolivia, Paraguay, Uruguay), Central Africa (forests band), south-east Africa (Tanzania, Malawi, Zambia, Mozambique, Eswatini), the exterior of Australia and the south-east of Asia (India, Bangladesh, Myanmar, Thailand, Cambodia, Vietnam,...). From around four thousand SLSTR scenes encompassing these areas for each three-month periods, around 220 products which respected all the criteria delineated above were selected to perform this analysis.

6.3.2.2 Night-Time Validation

Night-time validation is performed considering three 27-days orbital cycle and around 120 SLSTR FRP products are compared to MODIS MOD14 fire products to identify fires and other thermal anomalies for each 3-months periods between March and December 2022. Then, results over each trimester have been aggregated into the table below.

The comparison might be done with caution as we compare different time and different validation sites (areas of high fire activity are re-evaluated at each period). However, this comparison highlights the stability of the results.

Table 24: Summary of the inter-comparison between night-time SLSTR FRP and MODIS FRP – October, November and December 2022. Results from previous comparison obtained between July and August 2022, April and May
2022 also obtained with FRP have been added respectively in the second and third column for information and comparison.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value (Oct. 2022 – Dec. 2022; FRP V2)</th>
<th>Value (July – Sept. 2022; FRP V2)</th>
<th>Value (April – May 2022; FRP V2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of commission AFP</td>
<td>13,093 (55.1% of Total SLSTR AFP)</td>
<td>20,092 (62.3% of Total SLSTR AFP)</td>
<td>5,531 (53.7% of Total SLSTR AFP)</td>
</tr>
<tr>
<td>Number of omission AFP</td>
<td>694 (2.9% of Total SLSTR AFP)</td>
<td>883 (2.7% of Total SLSTR AFP)</td>
<td>401 (3.9% of Total SLSTR AFP)</td>
</tr>
<tr>
<td>FRP of commission AFP (MW)</td>
<td>40,359</td>
<td>58,380</td>
<td>20,164</td>
</tr>
<tr>
<td>FRP of omission AFP (MW)</td>
<td>30,641</td>
<td>34,123</td>
<td>7,940</td>
</tr>
<tr>
<td>Number of SLSTR AFP detected by both sensors</td>
<td>10,665 (44.9% of Total SLSTR AFP)</td>
<td>12,147 (37.7% of Total SLSTR AFP)</td>
<td>4,775 (46.3% of Total SLSTR AFP)</td>
</tr>
<tr>
<td>Number of MOD14 AFP detected by both sensors</td>
<td>3,341 (82.8% of Total MOD14 AFP)</td>
<td>4,089 (82.1% of Total MOD14 AFP)</td>
<td>1,798 (81.8% of Total MOD14 AFP)</td>
</tr>
<tr>
<td>Total number of AFP detected by SLSTR</td>
<td>23,758</td>
<td>32,239</td>
<td>10,306</td>
</tr>
<tr>
<td>Total number of MOD14 AFP</td>
<td>4,035</td>
<td>4,972</td>
<td>2,199</td>
</tr>
<tr>
<td>Mean number of SLSTR AFP per cluster</td>
<td>6.4</td>
<td>5.9</td>
<td>5.8</td>
</tr>
<tr>
<td>Total SLSTR FRP within clusters (MW)</td>
<td>81,672</td>
<td>95,704</td>
<td>27,436</td>
</tr>
<tr>
<td>Mean SLSTR FRP per cluster (MW)</td>
<td>54.0</td>
<td>50.4</td>
<td>42.9</td>
</tr>
<tr>
<td>Median SLSTR FRP per cluster (MW)</td>
<td>21.0</td>
<td>21.7</td>
<td>20.3</td>
</tr>
<tr>
<td>Mean number of MOD14 AFP per cluster</td>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Total MOD14 FRP within clusters (MW)</td>
<td>70,530</td>
<td>85,814</td>
<td>23,367</td>
</tr>
<tr>
<td>Mean MOD14 FRP per cluster (MW)</td>
<td>46.7</td>
<td>45.2</td>
<td>36.5</td>
</tr>
<tr>
<td>Median MOD14 FRP per cluster (MW)</td>
<td>13.6</td>
<td>14.3</td>
<td>13.0</td>
</tr>
<tr>
<td>Mean bias of FRP per cluster (MW)</td>
<td>7.4</td>
<td>5.2</td>
<td>6.4</td>
</tr>
<tr>
<td>Median of FRP scatter per cluster (MW)</td>
<td>6.0</td>
<td>5.2</td>
<td>5.2</td>
</tr>
<tr>
<td>Root-mean-square deviation of FRP per cluster</td>
<td>59</td>
<td>71</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>25-50-75 percentiles of SLSTR clusters FRP</td>
<td>25-50-75 percentiles of MOD14 clusters FRP</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.1, 21.0, 40.2</td>
<td>12.1, 21.7, 42.6</td>
<td>11.2, 20.3, 43.7</td>
</tr>
<tr>
<td></td>
<td>7.8, 13.6, 29.5</td>
<td>8.1, 14.3, 31.2</td>
<td>7.8, 13.0, 28.7</td>
</tr>
</tbody>
</table>
Optical MPC
Sentinel-3 optical Annual Performance Report
Year 2022

Ref.: OMPC.ACR.APR.002
Issue: 1.1
Date: 17/03/2023
Page: 175

Figure 147: Areas of high fire activity selected for the inter-comparison. The basemap shows night-time fires detected by SLSTR (top) and MODIS (bottom) for the months of respectively (1) April, May – (2) July, August, September and (3) October, November and December 2022.

Overall, there is good agreement between SLSTR FRP and MODIS FRP, as can be seen in Figure 147. The comparison shows that SLSTR detects in general more fire pixels than MODIS (over the period from March 66,303 vs 11,206), albeit many of them with very low FRP. Furthermore, there is a large number of commission fire pixels and a low number of omission fire pixel. Such pixels indicate fires that were detected only by one sensor, however, these are not necessarily incorrect/missed detections.

It is important to highlight, in fact, that the two sensors observe the scenes at slightly different times (the MODIS product has to be within an interval of ± 6 minutes with respect to the SLSTR acquisition). This difference in time translates into different conditions of the observed fires and also of the cloud coverage. A fire could move, increase/decrease in extent and power, whereas clouds could cover different portions of the image. On the other hand, a fraction of these pixels may represent real fires that are undetected
by MODIS but are detected by the SLSTR product - for example because of the use of the smaller pixel footprint F1 channel – or vice versa. For these reasons, the reported values should not be interpreted as full validation, but rather as indication of the consistency between the two sensors.

A comparison of fire pixels detected distinctly by Sentinel 3A and Sentinel 3B is visualised in Figure 148 considering only last 3-months SLSTR FRP data because the results remain almost the same for all 2022 periods of time.

![Figure 148: The basemap shows night-time fires detected by Sentinel 3A (top) and Sentinel 3B (bottom) for the months of October, November and December 2022.](image)

The fire pixels dispersion is significantly similar between both satellites Sentinel A and B in all the areas of high fire activity for night-time SLSTR products. However, it could be noted that for night-time and day-time Sentinel 3B provide respectively 55% and 52% of the fire pixels detection for this period.

The dispersion of active fires is lightly not uniform. Clusters are generated based on active fire density zones (groups of one or more pixels spatially adjacent to each other and corresponding to a single fire -
hotspot location that radiates a heating signal within a pixel size of 1 km²). Figure 149 is providing an example of what is called a cluster:

The left panel is showing all the pixels confirmed as an active fire. The FRP product will provide the position and an equivalent FRP value for all these pixels;

However, some of these pixels are close enough to be interpreted as a single fire event, called fire cluster (see distribution of the pixels into three distinct fire clusters in the middle panel). Each cluster will then be analysed separately and each single FRP value will be computed taken into account the whole cluster;

The right panel is representing 3 fire clusters, each of them associated with an global average FRP and an specific geographical extent (i.e. the number of fire pixel included in this cluster). This cluster distribution will be the basis of the following validation activities.

![Cluster formation for a specific area.](image)

In the following, we are then comparing clusters detected by MODIS and by SLSTR. In particular, if a fire cluster is detected by both satellites, the number of individuals fire pixels inside this cluster can be different for SLSTR than for MODIS (for example on a specific cluster, MODIS could detect 3 fire pixels whereas SLSTR detect 6 fire pixels). These proportions for both sensors are visualised in Figure 150.
Figure 150: Scatterplot between the number of fires detected by both sensors (night-time). SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. (1) April-May, (2) July, August and September, (3) October, November and December. Colour of circles represents the SLSTR fire radiative power while diameter represents the number of occurrences.
The graph compares the number of fires detected by SLSTR with the one detected by MODIS over a single cluster. Each circle counts then for a fire cluster and its coordinates on the graph indicates the number of fires detected by the sensors SLSTR (X-axis) and MODIS (Y-axis).

The first information provided by this graph is then that all clusters detected by MODIS are also detected by SLSTR (no circle on the Y axis), where we can have small fire cluster detected by SLSTR and missed by MODIS (see the small point on X-axis).

In addition, the size of each circle specifies the number of cases with a similar MODIS/SLSTR fires detection ratio (i.e. the same couple of coordinates corresponding to the same proportion of fires detected by both sensors) in the manner of a 3D graph. The colour of each point indicates the total Fire Radiative Power of the cluster detected by satellites.

Figure 151 shows that:

**SLSTR detects more fire pixels than MODIS in a single cluster.** These clusters are associated with low FRPs and with larger circle size indicating a higher sensitivity of SLSTR to smaller fires.

Considering large fire clusters, they are always associated with large FRP and the ratio SLSTR/MODIS regarding the number of pixel fires included in these cluster shows again a higher sensitivity of SLSTR with more individual pixel detected as active fires.

The distribution of FRP for fire clusters detected by both sensors is quite similar between SLSTR and MODIS, even though MODIS exhibits a higher peak for low FRP values and SLSTR shows a higher curve for intermediate values. Figure 5 highlights the fact that SLSTR is detecting more fire pixels than MODIS for the same fire clusters, many of them with very low FRP, and the total FRP of all clusters is higher for SLSTR, although this number is heavily affected by a few outliers with high FRPs.

Contrary to the case of omission/commission fire pixels, the cluster analysis includes a step for checking the relevant flags associated with the fire detection, in particular those related to water or clouds in the background window around the fire cluster. Thus, results of the cluster analysis are more robust against differences in the atmospheric conditions or cloud masking algorithm. Nonetheless, the detections could still have been affected by the fact that the different sensors do not observe the fires exactly at the same time and are not perfectly equivalent. Hence, some fluctuations are expected.
Cyclic Validation Cluster Distribution

Figure 151: Distribution of the FRP for fire clusters detected by both sensors. The lines represent kernel density estimates (KDE) computed over the barplot.
Figure 152: Scatterplot between the FRP of fire clusters detected by both sensors. SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. The regression line (with a shaded 0.05 confidence interval) is obtained with a robust linear method (RLM), which is less affected by outliers. The scatter points are color-coded according to their cluster number, so that they can be traced back to the original datasets. (1) April, May $R=0.785$; (2) July, August and September $R=0.835$ and (3) October, November and December $R=0.925$. 
The ratio SLSTR/MODIS regarding the number of included pixel is also closer to 1 specifically for large FRP values, showing equivalent sensitivity in case of strong fire for both MODIS and SLSTR. It’s interesting to highlight that high FRP fire clusters (>400 MW) show a high correlation to the regression line, the dispersion of the points differs from the previous validation report, the R-squared value is closer to 1.

6.3.2.3 Day-Time Validation

As for Night-time validation, Day-time validation is performed considering three 27-days orbital cycle and around 120 SLSTR FRP products are compared to MODIS MOD14 fire products to identify fires and other thermal anomalies for each 3-months periods **between March and December 2022**.

A summary of the results is reported in Table 25.

**Table 25: Summary of the inter-comparison between SLSTR FRP and MODIS FRP – From October to December 2022 data.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of commission AFP</td>
<td>5,344 (47.9% of Total SLSTR AFP)</td>
<td>7,613 (48.6% of Total SLSTR AFP)</td>
<td>3,071 (42.7% of Total SLSTR AFP)</td>
</tr>
<tr>
<td>Number of omission AFP</td>
<td>3,693 (33.1% of Total SLSTR AFP)</td>
<td>5,720 (36.5% of Total SLSTR AFP)</td>
<td>2,484 (31.5% of Total SLSTR AFP)</td>
</tr>
<tr>
<td>FRP of commission AFP (MW)</td>
<td>67,342</td>
<td>96,572</td>
<td>40,883</td>
</tr>
<tr>
<td>FRP of omission AFP (MW)</td>
<td>146,446</td>
<td>191,942</td>
<td>52,621</td>
</tr>
<tr>
<td>Number of SLSTR AFP detected by both sensors</td>
<td>5,819 (52.1% of Total SLSTR AFP)</td>
<td>8,054 (51.4% of Total SLSTR AFP)</td>
<td>4,127 (57.3% of Total SLSTR AFP)</td>
</tr>
<tr>
<td>Number of MOD14 AFP detected by both sensors</td>
<td>4,421 (54.5% of Total MOD14 AFP)</td>
<td>5,845 (50.5% of Total MOD14 AFP)</td>
<td>2,947 (54.2% of Total MOD14 AFP)</td>
</tr>
<tr>
<td>Total number of AFP detected by SLSTR</td>
<td>11,163</td>
<td>15,667</td>
<td>7,198</td>
</tr>
<tr>
<td>Total number of MOD14 AFP</td>
<td>8,114</td>
<td>11,565</td>
<td>5,431</td>
</tr>
<tr>
<td>Mean number of SLSTR AFP per cluster</td>
<td>2.5</td>
<td>2.7</td>
<td>3.0</td>
</tr>
<tr>
<td>Total SLSTR FRP within clusters (MW)</td>
<td>73,916</td>
<td>156,059</td>
<td>56,580</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------------------------------</td>
<td>----------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Mean SLSTR FRP per cluster (MW)</td>
<td>48.0</td>
<td>59.0</td>
<td>64.1</td>
</tr>
<tr>
<td>Median SLSTR FRP per cluster (MW)</td>
<td>27.4</td>
<td>27.4</td>
<td>34.0</td>
</tr>
<tr>
<td>Mean number of MOD14 AFP per cluster</td>
<td>1.7</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>Total MOD14 FRP within clusters (MW)</td>
<td>68,320</td>
<td>143,963</td>
<td>51,904</td>
</tr>
<tr>
<td>Mean MOD14 FRP per cluster (MW)</td>
<td>44.3</td>
<td>54.4</td>
<td>58.9</td>
</tr>
<tr>
<td>Median MOD14 FRP per cluster (MW)</td>
<td>22.3</td>
<td>22.5</td>
<td>25.7</td>
</tr>
<tr>
<td>Mean bias of FRP per cluster (MW)</td>
<td>3.6</td>
<td>4.6</td>
<td>5.3</td>
</tr>
<tr>
<td>Median of FRP scatter per cluster (MW)</td>
<td>3.7</td>
<td>3.4</td>
<td>5.7</td>
</tr>
<tr>
<td>Root-mean-square deviation of FRP per cluster</td>
<td>58.9</td>
<td>64.0</td>
<td>66.0</td>
</tr>
<tr>
<td>25-50-75 percentiles of SLSTR clusters FRP</td>
<td>17.0, 27.4, 52.0</td>
<td>16.0, 27.4, 54.4</td>
<td>19.2, 33.9, 63.5</td>
</tr>
<tr>
<td>25-50-75 percentiles of MOD14 clusters FRP</td>
<td>13.0, 22.3, 42.9</td>
<td>12.2, 22.5, 48.6</td>
<td>13.5, 25.7, 50.4</td>
</tr>
</tbody>
</table>

The results remain relatively consistent with the Night-time one, with more fire detected by SLSTR than MODIS (over the period from March 34,028 vs 25,110) but a good agreement regarding fire position and regarding FRP values. The comparison shows that SLSTR detects in general more fire pixels than MODIS and many of them with very low FRP. There is a quite similar number of commission and omission fire pixels comparing with Night-time data. The results remain also consistent with the previous report.
Figure 153: Areas of high fire activity selected for the inter-comparison. The basemap shows day-time fires detected by SLSTR (top) and MODIS (bottom) for the months of respectively (1) April, May – (2) July, August, September and (3) October, November and December 2022.
Figure 154: Scatterplot between the number of fires clusters detected by both sensors (day-time). SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. (1) April-May, (2) July, August and September, (3) October, November and December.
To compare with night-time results, dots are less aggregated close to the SLSTR axis, reflecting the closer number of omissions and commissions. The high FRP clusters are more detected by MODIS while SLSTR detect more low FRP fires.
Figure 155: Scatterplot between the FRP of fire clusters detected by both sensors (with an FRP < 500MW). SLSTR values are reported on the x-axis, whereas MODIS ones are on the y-axis. (1) April, May $R=0.915$; (2) July, August and September $R=0.925$ and (3) October, November and December $R = 0.775$.

The last report shows a higher dispersion of the points than for night-time data.
This section has provided a detailed performance evaluation of SLSTR FRP product generated by Sentinel-3 A and Sentinel-3 B. On a “per-fire” basis there is a fair correlation (on average over the year $r^2=0.85$ for night-time data and $r^2=0.82$ for day-time data) between the FRP measures recorded by Sentinel 3 and by MODIS near simultaneously, and a slope of the linear best fit line close to the unity (on average all over the year 0.871 for day-time and 0.848 for night-time). Overall, the number of SLSTR AFP detected by both sensors reaches 45% of night-time products and almost 52% of day-time products.
6.3.2.4 Conclusion

Intercomparison with MODIS fire products has been used to assess the quality of the SLSTR products. A very good consistency is observed for the period from March to December 2022. Spatially, we find a good agreement between MODIS and SLSTR on different areas of interests corresponding to large fire activity zones. MODIS detects slightly fewer fires in the Northern hemisphere at night-time, whereas results are closer in the Southern hemisphere. For day-time fires, MODIS detects clearly more fires in the Northern hemisphere. More investigations are needed to better assess those differences. In future, comparisons should also be extended to other sensors like VIIRS to overcome the end of MODIS.
7 Summary of performances – SYN

7.1 L1 products performances

The quality assessment of the misregistration data between OLCI and SLSTR has been done before February 2017, in particular with the inclusion of updated intra-instrument misregistration Auxiliary Data files. As a consequence, we focused on operational issues to ensure the production of SYNERGY products all over the globe and at any time.

In particular, SYNERGY L1 processing was specified to be processed only on descending half-orbits. As consequences, the seasonal variation of OLCI orbit – implying a starting point of the Earth Observation acquisition in ascending mode – was not well handled by SYNERGY Level 1 module and was filtering out all SLSTR radiometric measurements from SYN L1 dataset. The resulting SYN L2 Aerosol optical thickness was then retrieved using only OLCI measurements and no SLSTR surface reflectance were included in affected products.

Following several discussions and investigations by the ground segment, the operational processing teams and the SYNERGY ESLs, this specificity is now considered and well-handled and SYN L2 product is not affected anymore by this seasonal variation.

The SYNERGY Level 1 processing is now performed without issue related to an incompatibility between SLSTR and OLCI products.

7.2 L2 product performances

7.2.1 SYN L2 SDR quality assessment: intercomparisons with other missions

7.2.1.1 Introduction

Intercomparison of Sentinel 3 derived SYNERGY (SYN) Surface Directional Reflectances (SDR) with comparable products from other missions is carried out as no in-situ reference data is available for direct validation.

The SYN SDR product is derived from the combined measurements of OLCI and SLSTR (Nadir and Oblique) onboard Sentinel 3 and is produced at 300 m spatial resolution. The SYN SDR measurements are compared against the normalized surface reflectance measurements observed from MODIS (MCD43A4.061, a combined Terra and Aqua Product) at 500 m spatial resolution and VIIRS (VNP43MA4.001) at 1 km resolution. This product Nadir BRDF-adjusted Reflectances (NBAR), where surface reflectances are corrected to a common nadir view geometry at the local solar noon zenith angle of the day of interest. The SYNERGY (SYN) derived SDR products thus requires to be adjusted to the same illumination and viewing conditions before any intercomparison exercise can be undertaken. The geometric correction is performed with RTLS-R (Ross-Thick-Li-Sparse Reciprocal) BRDF (Bidirectional Reflectance Distribution Function) models, taken from the MODIS product MCD43A1.061 (version 6.1), which is also a combined Terra and Aqua product. Also, in a similar way, for VIIRS, the BRDF products are observed from VNP43MA1.001 (collection 1). Considering the differences in spatial resolution for SYN (300 m), for MODIS (500 m) and VIIRS (1 km), the intercomparison exercise is always performed for the intersection grid of scenes and re-projected at 0.001-degree resolution (~ 1 km).
For the current study, the performance assessment of the intercomparison is represented in terms of statistical variables, namely, Accuracy (A), Precision (P) and Uncertainty (U). For brevity, the Accuracy is known to represent the mean bias of the estimates (or the mean bias difference), Precision represents the repeatability, and the uncertainty stands for the root mean squared error difference. The equations of the metrics (A, P, U) are shown below.

\[ A = \frac{1}{N} \cdot \sum_{i=1}^{N} \epsilon_i \]  
\[ P = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^{N} (\epsilon_i - A)^2} \]  
\[ U = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \epsilon_i^2} \]  

where \( \epsilon_i \) is the observed difference MODIS (or VIIRS) minus SYN SDR.

7.2.1.2 Results and Discussion:

After the geometric adjustment, the SYN SDR products are evaluated against those observed from MODIS nadir adjusted products. Only the closely matching spectral bands between MODIS and OLCI (difference of 10 nm or less) are selected. The same exercise has been also started with VIIRS in order to prepare for the decommissioning of MODIS; preliminary results are also presented below.

7.2.1.2.1 Regional inter-comparison

Several regions namely United States, Amazonia, Europe, France, Spain, Africa, Congo and Australia, are selected across the globe to perform regional inter-comparison of SYN SDR products. These regions are represented as area polygons in Figure 157.

![Figure 157: Selected regions across the globe for inter-comparison of SYN SDR products.](image-url)
The regional inter-comparisons are performed between the satellite reflectance products (SYN SDR vs MODIS, SYN SDR vs VIIRS) when there is overlap between the two image granules. The necessary steps, such as filtering of bad quality data, observed among satellite products, reprojection and BRDF correction are implemented before any inter-comparisons are made.

Figure 158 shows the inter-comparison exercise performed between SYN SDR (OLCI), Oa17 and MODIS, b02, for the selected regions across the globe and for a randomly selected day.
Figure 158: density scatter plots of inter-comparisons of SYN SDR, Oa17 with that of observed from MODIS, b02, for the regions (a) Congo, (b) Europe, (c) Africa, (d) Australia, (e) Central USA and (f) region of France and for randomly selected days in the year 2022.
It is clearly observed from Figure 158 that SYN SDR reasonably compares to that of MODIS after the geometric adjustment, where the correlation scores, $R^2$ reach up to 0.929 for Central US and 0.855 for Congo.

Similarly, Figure 159 shows the inter-comparison exercise performed between SYN SDR, OLCI bands Oa08, Oa17, Oa04, Oa06 and MODIS, b01, b02, b03, b04 for the region of Australia.

![Figure 159: density scatter plots of inter-comparisons between SYN SDR (OLCI), bands Oa08, Oa17, Oa04, Oa06 and MODIS, b01, b02, b03, b04 for the region of Australia and for a selected day during the year 2022.](image)

It is clearly observed from Figure 3 that SYN SDR reasonably compares to that of MODIS for the pair of bands (b01 vs Oa08), (b02 vs Oa17), (b03 vs Oa04) and (b04 vs Oa06) used in the inter-comparison exercise, with $R^2$ values higher than 0.8.

An attempt is also made in this study to inter-compare SYN SDR (OLCI) products with the surface reflectance products observed from VIIRS. Shown in the Figure 160 is the inter-comparison exercise...
performed between SYN SDR (OLCI), Oa17 and VIIRS, bands M7, for the selected regions across the globe and for random selected days during 2022, vis-à-vis to that of MODIS, b02.

![Image of density scatter plots](image)

**Figure 160:** density scatter plots of inter-comparisons of SYN SDR (OLCI), Oa17 with that of observed from VIIRS, bands M7, for the regions (b) Africa, (d) Australia. Shown in the same figure, the counter parts from MODIS, b02 for the same regions, (a) Africa, (c) Australia and for randomly selected days during the year 2022.

It is observed from the figure that the SYN SDR compares well with that of VIIRS and MODIS. This shows that VIIRS derived SR products are a good alternative to MODIS products after the decommissioning of MODIS surface reflectance products and upon their unavailability.

### 7.2.1.2.2 Time series intercomparison
For the inter-comparison exercise, the SYN SDR and MODIS surface directional reflectance (SDR) products are extracted over a 50 X 50 km box, and around several CEOS (Committee on Earth Observation Satellites), LPV (land product validation) sites, during the period between 2022-09 and 2022-11. The selected CEOS LPS sites are known to represent varying biome classes across the globe. The list of CEOS LPV sites is shown in Table 26 below.

**Table 26: List of CEOS LPV sites used for intercomparison of SYN SDR products.**

<table>
<thead>
<tr>
<th>No</th>
<th>Site acronym</th>
<th>Country</th>
<th>Network</th>
<th>Lat</th>
<th>Lon</th>
<th>Land cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AU-Cumberland</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-33.615</td>
<td>150.723</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>2</td>
<td>AU-Great-Western</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-30.192</td>
<td>120.654</td>
<td>Broadleaved, deciduous, open</td>
</tr>
<tr>
<td>3</td>
<td>AU-Litchfield</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-13.180</td>
<td>130.790</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>4</td>
<td>AU-Robson-Creek</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-17.117</td>
<td>145.630</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>5</td>
<td>SP-Ali</td>
<td>Spain</td>
<td>CORE</td>
<td>38.452</td>
<td>-1.065</td>
<td>Cropland</td>
</tr>
<tr>
<td>6</td>
<td>US-Moab-Site</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>38.248</td>
<td>-109.388</td>
<td>Shrub, closed-open, deciduous</td>
</tr>
<tr>
<td>7</td>
<td>US-Talladega</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>32.950</td>
<td>-87.393</td>
<td>Needle-leaved, evergreen</td>
</tr>
<tr>
<td>8</td>
<td>AU-Wombat</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-37.422</td>
<td>144.094</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>9</td>
<td>FR-Guayaflux</td>
<td>France</td>
<td>ICOS Associated</td>
<td>5.279</td>
<td>-52.925</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>10</td>
<td>FR-Hesse</td>
<td>France</td>
<td>ICOS</td>
<td>48.674</td>
<td>7.065</td>
<td>Broadleaved, deciduous, closed</td>
</tr>
<tr>
<td>11</td>
<td>US-Harvard</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>42.537</td>
<td>-72.173</td>
<td>Broadleaved, deciduous, closed</td>
</tr>
<tr>
<td>12</td>
<td>US-Mountain-Lake</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>37.378</td>
<td>-80.525</td>
<td>Broadleaved, deciduous, closed</td>
</tr>
<tr>
<td>13</td>
<td>AU-Calperum</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-34.003</td>
<td>140.588</td>
<td>Shrub, closed-open, deciduous</td>
</tr>
<tr>
<td>14</td>
<td>AU-Cape-Tribulation</td>
<td>Australia</td>
<td>TERN-SuperSites, OzFlux</td>
<td>-16.106</td>
<td>145.378</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>15</td>
<td>AU-Rushworth</td>
<td>Australia</td>
<td>TERN-AusCover</td>
<td>-36.753</td>
<td>144.966</td>
<td>Broadleaved, deciduous, open</td>
</tr>
<tr>
<td>16</td>
<td>AU-Tumbarumba</td>
<td>Australia</td>
<td>TERN-SuperSites, AusCover/OzFlux</td>
<td>-35.657</td>
<td>148.152</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>17</td>
<td>FR-Puechabon</td>
<td>France</td>
<td>ICOS</td>
<td>43.741</td>
<td>3.596</td>
<td>Needle-leaved, evergreen</td>
</tr>
<tr>
<td>18</td>
<td>IT-Cat</td>
<td>Italy</td>
<td>CORE</td>
<td>37.279</td>
<td>14.883</td>
<td>Cropland</td>
</tr>
<tr>
<td>No</td>
<td>Site acronym</td>
<td>Country</td>
<td>Network</td>
<td>Lat</td>
<td>Lon</td>
<td>Land cover</td>
</tr>
<tr>
<td>----</td>
<td>--------------</td>
<td>---------------</td>
<td>----------------</td>
<td>-------</td>
<td>-------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>19</td>
<td>IT-Lison</td>
<td>Italy</td>
<td>ICOS</td>
<td>45.740</td>
<td>12.750</td>
<td>Cropland</td>
</tr>
<tr>
<td>20</td>
<td>US-Oak-Rige</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>35.964</td>
<td>-84.283</td>
<td>Broadleaved, deciduous, closed</td>
</tr>
<tr>
<td>21</td>
<td>AU-Watts-Creek</td>
<td>Australia</td>
<td>TERN-AusCover</td>
<td>-37.689</td>
<td>145.685</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>22</td>
<td>FR-Montiers</td>
<td>France</td>
<td>ICOS</td>
<td>48.538</td>
<td>5.312</td>
<td>Broadleaved, deciduous, closed</td>
</tr>
<tr>
<td>23</td>
<td>US-Bartlett</td>
<td>United States</td>
<td>NEON, AERONET</td>
<td>44.064</td>
<td>-71.287</td>
<td>Broadleaved, deciduous, closed</td>
</tr>
<tr>
<td>24</td>
<td>BR-Mata-Seca</td>
<td>Brazil</td>
<td>ENVIRONET</td>
<td>-14.880</td>
<td>-43.973</td>
<td>Herbaceous, closed-open</td>
</tr>
<tr>
<td>25</td>
<td>IT-Collelongo</td>
<td>Italy</td>
<td>EFDC</td>
<td>41.849</td>
<td>13.588</td>
<td>Broadleaved, deciduous, closed</td>
</tr>
<tr>
<td>26</td>
<td>SE-Dahra</td>
<td>Senegal</td>
<td>KIT / UC</td>
<td>15.400</td>
<td>-15.430</td>
<td>Cultivated and managed areas</td>
</tr>
<tr>
<td>27</td>
<td>AU-Zigzag-Creek</td>
<td>Australia</td>
<td>TERN-AusCover</td>
<td>-37.474</td>
<td>148.339</td>
<td>Broadleaved, evergreen</td>
</tr>
<tr>
<td>28</td>
<td>FR-Estrees-Mons</td>
<td>France</td>
<td>ICOS Associated</td>
<td>49.872</td>
<td>3.021</td>
<td>Cultivated and managed areas</td>
</tr>
<tr>
<td>29</td>
<td>NE-Loobos</td>
<td>Netherlands</td>
<td>ICOS Associated</td>
<td>52.166</td>
<td>5.744</td>
<td>Needle-leaved, evergreen</td>
</tr>
<tr>
<td>30</td>
<td>FR-Aurade</td>
<td>France</td>
<td>ICOS</td>
<td>43.550</td>
<td>1.106</td>
<td>Cropland</td>
</tr>
</tbody>
</table>

Figure 161 to Figure 167 below show the time series of intercomparison between SYNERGY SDR (OLCI) and MODIS SDR products, for selected stations in Table 26: Talladega (needle-leaved evergreen), Robson-Creek (broadleaved evergreen), Estrees-Mons (Cultivated) and Lison (Crop), MOAB, United States (Shrub closed open deciduous), Dahra, Senegal (Cultivated) and Mata-Seca, Brazil (Herbaceous).
Figure 161: (top 4 plots) time series of surface reflectance (SR) of MODIS b01, b02, b03 and b04 and their Sentinel3-A SYN SDR (OLCI) counterpart Oa08, Oa17, Oa04 and Oa06; and (bottom 4 plots) time series of the comparison statistics in terms of A,P,U for the selected bands, for the station Talladega, United States (needle-leaved evergreen).
Figure 162: same as Figure 161 but for the station Robson-Creek, Australia (Broadleaved Evergreen).
Figure 163: same as Figure 161 but for the station Estrees-Mons, France (Cultivated).
Figure 164: same as Figure 161 but for the station Lison, Italy (Cropland).
Figure 165: same as Figure 161 but for the station MOAB, United States (Shrub closed open deciduous).
Figure 166: same as Figure 161 but for the station Dahra, Senegal (Cultivated).
It is observed from Figure 161 to Figure 167 that the influence of underlying surface biome is clearly seen on the results, with a good agreement observed for needle and broad evergreen biomes (Figure 161 and Figure 162), and a slight degradation observed for cultivated and crop biomes (Figure 163 and Figure 164).

The retrieved time series of surface reflectance data as observed from sentinel S3A SYNERGY OLCI and MODIS and for all the CEOS LPV sites are represented as scatter plot, for all the surface classification types and is shown in Figure 168 and for several spectral band pairs, b01 vs Oa08 (Figure 168a), b02 vs Oa17 (Figure 168b), b03 vs Oa04 (Figure 168c) and b04 vs Oa06 (Figure 168d). The scatter points are also indicated by their surface biome classification as indicated by colour code.
Figure 168: Scatter plot of intercomparisons between Sentinel3-A SYN SDR (OLCI) and MODIS for selected band pairs: b01 vs Oa08 (a), b02 vs Oa17 (b), b03 vs Oa04 (c) and b04 vs Oa06 (d).

Similarly, shown in Figure 169, is the overall scatter plot from the intercomparison of Sentinel3-B SYN SDR with MODIS for several spectral band pairs, b01 vs Oa08 (Figure 169a), b02 vs Oa17 (Figure 169b), b03 vs Oa04 (Figure 169c) and b04 vs Oa06 (Figure 169d)
An overall good comparison is observed between Sentinel S3A/S3B derived SYN SDR (OLCI) and MODIS surface reflectance products, and for all the selected spectral band pairs. Slightly better statistical scores are observed for Sentinel S3B SYN SDR (Figure 169) as compared to Sentinel S3A (Figure 168) as can be seen from the statistical indicators in the figures.

To understand the influence of biome classification, we have selected few biome classes and the intercomparison exercise is repeated separately for each of these biomes. Figure 170 shows the intercomparison between Sentinel S3B SYN SDR (OLCI) and MODIS derived surface reflectance for a single data pair (B01 vs. Oa08) but for each single biome: broad and needle evergreen, cultivated and shrub.

Figure 169: Scatter plot of intercomparisons between Sentinel3-B SYN SDR (OLCI) and MODIS for selected band pairs: b01 vs Oa08 (a), b02 vs Oa17 (b), b03 vs Oa04 (c) and b04 vs Oa06 (d).
Similarly, Figure 171 shows the inter-comparisons between Sentinel S3A SYN SDR (OLCI) and MODIS derived surface reflectance data pairs but shown only for the single biome classification types: broad and needle evergreen, cultivated and shrub.
It is clearly observed from Figure 170 and Figure 171 that there is an influence of surface biome on the inter-comparison as observed from both S3A (Figure 171) and S3B (Figure 170). The derived uncertainty (0.01) is better for Needle and Broad evergreen types and as when compared to the uncertainty value of 0.04 observed with shrub and cultivated types.

7.2.1.3 Conclusion

Under the OPT-MPC project and as a part of routine service validation, the Sentinel 3, derived SYNERGY SDR products are geometrically adjusted as a first step and are inter-compared against the MODIS nadir adjusted surface reflectance products. In general, a good agreement is observed between both Sentinel 3 A/B derived SYN SDR and MODIS (VIIRS) surface reflectance products. The observed statistical scores are slightly better with S3B than compared to S3A. The correction of discrepancies with respect to acquisition geometry greatly improved the bias associated with SYN SDR products and hence it is advisable to implement such correction while carrying the inter-comparison exercise. Also, it was observed that
there is a clear influence of surface biome on the inter-comparisons, where the uncertainty estimates are better for needle and broad evergreen classes than to cultivated and shrub classes. With MODIS coming to its end of commissioning phase, it is envisaged to use fully the VIIRS derived surface reflectance products instead of MODIS surface reflectance products and the inter-comparisons exercise will be adapted.

7.2.2 SYN VGT: consistency evaluation with SPOT/Vegetation and PROBA-V

In order to assess the possible extension of the SPOT/Vegetation (SPOT/VGT) – PROBA-V data archive with Sentinel-3 SYN VGT, the consistency between Sentinel-3 SYN VGT and SPOT/VGT and PROBA-V products is evaluated. Since the operational phase of PROBA-V has ended on June 30 2020, and also the PROBA-V experimental phase has ended, direct intercomparison is no longer possible between the operational Sentinel-3 Level 2 synergy products and PROBA-V products.

The evaluation of SYN VGT product performance focuses on:
Visual checks and checks on missing data.
Consistency evaluation for SYN VGT V10 products is done through indirect comparison of SYN VGT V10 products of 2022 with 5-years long term statistics (LTS) averages based on SPOT/Vegetation Collection 3 [1] and PROBA-V Collection 2 (soon to be released to the public, see https://proba-v.vgt.vito.be/en/proba-v-collection-2-reprocessing-campaign).

To complete this analysis, the previous results obtained on direct comparison of SYN VGP products with experimental PROBA-V data using the S3A PB 2.56, S3B PB 1.28 are still valid and provided below.

7.2.2.1 Data

For direct comparison on top-of-atmosphere (TOA) products the following data was used:
❖ PROBA-V L2A products of the experimental phase (July/2020 – October/2021) with limited data acquisitions over Europe and Africa only. The number of product match-ups with closest acquisition time is 608 for S3A, 610 for S3B. The number of pairwise valid observations is illustrated in Figure 172.

Figure 172: Number of pairwise TOA observations used in the intercomparison between SY_VGP and PROBA-V L2A from 01/07/2020 – 30/09/2020.

For indirect comparison on 10-daily top-of-canopy (TOC) composite products with the LTS, the following data was used:
Sentinel-3A and B SY_V10 products from January/2022 to December/2022, based on processing baselines:

For S3A:
- IPF/PB 06.09/2.77, in operations since 14/June/2021
- IPF/PB 06.10/3.03, in operations since 27/January/2022
- IPF/PB 06.11/3.11, in operations since 23/August/2022

For S3B:
- IPF/PB 06.09/2.55, in operations since 14/June/2021
- IPF/PB 06.10/3.03, in operations since 27/January/2022
- IPF/PB 06.11/3.11, in operations since 9/September/2022

S3 SYN VGT processing baseline updates in 2022 had minor impacts on the SYN VGT product content. PB 06.11 solved the issue with data gaps at high latitudes and high viewing zenith angles (see below) and improved the status map.

SPOT5/Vegetation2 Collection 3 LTS based on 10-daily TOC composites (S10-TOC) from 2009 to 2013.

PROBA-V Collection 2 LTS based on 10-daily TOC composites (S10-TOC) from 2014 to 2018.

### 7.2.2.2 Methods

Validation metrics for statistical consistency analysis are calculated over a large number of samples (pixels) [2–4]. For both the SYN VGT and PROBA-V products, the SM was interpreted in order to exclude pixels labelled as ‘cloud’, ‘shadow’ (in case of PROBA-V), ‘snow/ice’ or ‘water’, or with bad radiometric quality or bad coverage in one of the spectral bands. A systematic spatial subsample of 0.25% (every 20th pixel in both X and Y) is applied in order to reduce processing time.

To identify the relationship between SYN VGT and SPOT/VGT or PROBA-V, the geometric mean (GM) regression model is used. Such an orthogonal (model II) regression is appropriate, because – unlike when comparing to an absolute reference – both datasets are subject to noise. By applying an eigen decomposition to the X and Y covariance metrics, two eigenvectors are obtained that describe the principal axes of the point cloud [4], i.e., the regression line. The GM regression intercept and slope value are added as quantitative information related to the scatterplots. The coefficient of determination ($R^2$) indicates the agreement or covariation between two datasets with respect to the linear regression model, summarizing the total explained variance by this model.

The results focus on the Accuracy (A), Precision (P) and Uncertainty (U) of the bias. The average bias is considered as a measure of Accuracy. The standard deviation of the bias, or repeatability of the observations, is a measure of Precision. Finally, the Root Mean Squared Difference (RMSD) expresses the difference magnitude between two datasets from 0 and is an expression of the overall difference, or uncertainty (U).

$$A = \frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i) = \bar{X} - \bar{Y}$$
7.2.2.3 Results and discussion

Statistical consistency of TOA reflectances (SY_VGP vs. PROBA-V L2A)

Figure 173 shows the distribution of the APU statistics per match-up in the direct comparison between S3 SY_VGP products and PROBA-V L2A TOA products. The statistics show a large range, caused by angular effects (i.e., large ranges in illumination and observation angles for the different observations). For bands B0 (blue), B2 (red) and B3 (NIR), an Accuracy of around 3% is observed, with SY_VGP being brighter than PROBA-V L2A (Figure 173). This is related to differences in absolute calibration, but also to the fact that SPOT4-VGT1 spectral response functions are used in the spectral band mapping procedure to generate SYN VGT products.

Both the Accuracy and Uncertainty indicate large inconsistencies between SY_VGP and PROBA-V L2A for the MIR band, with an average bias of around -8%. This can be largely attributed to the calibration issues of SLSTR, as reported in [5].

![Figure 173: Distribution of Accuracy (A), Precision (P) and Uncertainty (U) between SY_VGP and PROBA-V L2A match-ups (July-September 2020) for S3A (blue) and S3B (orange).](image)

Visual inspection

A number of quality issues were identified during visual inspection of SY_VG1 data.

- Spatial artefacts are introduced in the reprojection to the fixed 1/112° grid. The artefacts are observed at the west edge of the swath (e.g. Figure 174) and are related to the large ground sampling distance at large viewing angles. This issue was solved in IPF/PB 06.11/3.11.
The SYN VGT products show very high AOT values. Related to these high AOT values spatial artefacts in the Blue TOC are introduced by the atmospheric correction (e.g. Figure 175).

AOT values above a certain threshold (1.012) are masked as NaN, although TOC reflectances bands show valid numbers (e.g. Figure 175).

Omission of clouds, and no cloud shadow detection (as previously reported).
Data quality issues with a deviating amount of missing data and empty tiles were observed for the following dates: 2022/04/07, 2022/04/15, 2022/04/19, 2022/04/21, 2022/04/23-28, 2022/06/20-21, 2022/06/24, 2022/08/01-02, 2022/08/23, 2022/09/10, 2022/09/12, 2022/09/21, 2022/09/30, 2022/10/01.

**Statistical consistency of TOC reflectances (SY V10 vs. SPOT/VGT S10-TOC LTS and PROBA-V S10-TOC LTS)**

The statistical consistency of SYN VGT V10 products with the SPOT/VGT archive is evaluated through indirect comparison of the SY V10 products of 2022 with the SPOT/VGT S10-TOC LTS (2009-2013). Figure 176 shows the scatter density plots, frequency histograms and bias frequency plots between respectively S3A and S3B SY V10 products of 2022 and the SPOT/VGT LTS. Large inconsistencies between SY_V10 and SPOT/VGT S10-TOC are observed for the SWIR band, with an average bias (or accuracy) of -9%. This can be largely attributed to the calibration issues of SLSTR, as reported in [5]. The bias for Blue is very small (<1%). For Red and NIR TOC, an average bias of around 3% is observed, with slightly larger bias for S3A. This is related to differences in absolute calibration, but also to the fact that SPOT4-VGT1 spectral response functions are used in the spectral band mapping procedure to generate SYN VGT products. Precision and uncertainty are smallest for Blue (<3%), larger for Red and NIR (4 to 7%), and largest for SWIR.

Similarly, the statistical consistency of SYN VGT V10 products with the PROBA-V archive is evaluated through indirect comparison of the SY V10 products of 2022 with the PROBA-V S10-TOC LTS (2014-2018). Scatter density plots, frequency histograms and bias frequency plots between respectively S3A and S3B SY V10 products of 2022 and the PROBA-V LTS are shown in Figure 177. The results are very similar to the intercomparison with the SPOT/VGT LTS: largest inconsistencies for SWIR, small bias for Blue, and slightly larger biases for Red and NIR.
Figure 176: GMR scatter density plots, frequency histograms and bias frequency plots between S3A (top) and S3B (bottom) SY_V10 (2022) and SPOT-VGT S10-TOC LTS (2009-2013), for Blue, Red, NIR and SWIR bands.
S3A SY_V10 vs PROBA-V LTS

S3B SY_V10 vs PROBA-V LTS

Figure 177: GMR scatter density plots, frequency histograms and bias frequency plots between S3A (top) and S3B (bottom) SY_V10 (2022) and PROBA-V S10-TOC LTS (2014-2018), for Blue, Red, NIR and SWIR bands.
Differences between S3A and S3B SY_V10 products consistency with SPOT/VGT S10-TOC LTS and PROBA-V S10-TOC LTS are small (Table 27). This indicates the impact of differences in radiometric calibration between S3A-OLCI and S3B-OLCI, and S3A-SLSTR and S3B-SLSTR, respectively, are limited. The results are also very comparable for the intercomparison with SPOT/VGT and PROBA-V: SYN VGT V10 products show a similar bias with SPOT/VGT and with PROBA-V.

| Table 27 APU statistics for S3A and S3B SY_V10 products vs. SPOT/VGT LTS and PROBA-V LTS |
|-----------------|--------------|--------------|--------------|--------------|
| N               | Blue         | Red          | NIR          | MIR          |
| S3A vs. SPOT/VGT LTS | 8.3 10^6     | A            | 0.008        | 0.035        | 0.036        | -0.089       |
|                 |              | P            | 0.025        | 0.040        | 0.057        | 0.051        |
|                 |              | U            | 0.026        | 0.053        | 0.067        | 0.103        |
| S3B vs. SPOT/VGT LTS | 8.8 10^6     | A            | 0.006        | 0.032        | 0.031        | -0.095       |
|                 |              | P            | 0.025        | 0.039        | 0.055        | 0.051        |
|                 |              | U            | 0.025        | 0.050        | 0.063        | 0.108        |
| S3A vs. PROBA-V LTS | 10.8 10^6    | A            | -0.006       | 0.026        | 0.020        | -0.086       |
|                 |              | P            | 0.027        | 0.040        | 0.058        | 0.055        |
|                 |              | U            | 0.028        | 0.047        | 0.061        | 0.102        |
| S3B vs. PROBA-V LTS | 11.3 10^6    | A            | -0.009       | 0.023        | 0.015        | -0.091       |
|                 |              | P            | 0.027        | 0.039        | 0.056        | 0.055        |
|                 |              | U            | 0.029        | 0.045        | 0.058        | 0.106        |

Statistical consistency of TOC NDVI (SY_V10 vs. SPOT/VGT S10-TOC LTS and PROBA-V S10-TOC LTS)

As illustrated in Figure 178, the statistical consistency analysis of TOC NDVI shows an average bias of -0.05 with the SPOT/VGT LTS and -0.04 with the PROBA-V LTS. Both datasets show strong agreements ($R^2 \approx 0.9$). The bias is constant over the NDVI range, as the GMR slope is very close to 1.0. As for the TOC reflectances,
the differences between S3A and S3B, and between the comparison with SPOT/VGT LTS and PROBA-V LTS, are small.

**SY_V10 vs SPOT/VGT LTS**

![GMR scatter density plot (left), frequency histogram (middle) and bias frequency plot (right) between S3A resp. S3B SY_V10 (2022) and SPOT/VGT S10 TOC NDVI LTS (2009-2013) (top) and PROBA-V S10-TOC NDVI LTS (2014-2018) (bottom).](image)

**SY_V10 vs PROBA-V LTS**

![GMR scatter density plot (left), frequency histogram (middle) and bias frequency plot (right) between S3A resp. S3B SY_V10 (2022) and SPOT/VGT S10 TOC NDVI LTS (2009-2013) (top) and PROBA-V S10-TOC NDVI LTS (2014-2018) (bottom).](image)
7.2.2.4 Conclusions

The S3 SYN VGT 10-daily composite products (V10) of 2022 show a similar bias with the SPOT/VGT and PROBA-V archives, as shown through the intercomparison with long term statistics (LTS). Similar results were also obtained for S3A and S3B products. Comparable overall biases are observed at the TOA level, as through direct intercomparison at VGP level with PROBA-V L2A products.

Consecutive processing baseline updates have resulted in a large improvement of the statistical consistency between S3 SYN VGT products and PROBA-V, especially for the NDVI. However, some quality issues remain open:

- In the spectral resampling procedure, SPOT4-VGT1 spectral response functions (SRFs) are used. Better consistency, especially for the Red band, would be reached when using PROBA-V SRFs.
- Large systematic differences for SWIR are related to the absolute calibration of SLSTR. We advise to include the proposed correction factors [5] in the SYN VGT processing line. This will also have an impact on the image based AOT retrieval.
- Image retrieved AOT is very high, causing spatial artefacts in the Blue TOC.
- No cloud shadow detection is done for the SYN VGT products, and cloud detection is not accurate (see previous APR). This should be solved by incorporation of the latest IdePix version in the SYN VGT processing line.
- Interpolation errors, artefacts in the AOT and artefacts in the land/sea mask cause visual artefacts (see previous APR).
- S3A and S3B are not combined in one composite product (VG1 and V10).

While a number of recently solved issues have improved the quality of the current SYN VGT products and further enhancements are planned, a reprocessing action on the SYN VGT archive is required to allow for consistency analysis with nominal PROBA-V data and for users to have access to a continuous consistent data series from 1998 onwards.

7.2.2.5 References

5. Smith, D. *Assessment of visible and short wavelength radiometric calibration using vicarious calibration methods*; 2020;

### 7.2.3 SYN L2 global Aerosol performance

Since January 2020, a new SYNERGY product has been included in the Sentinel 3 portfolio. This product is dedicated to the Aerosol characteristics and provides - over both land and sea pixels and on a 4.5 km resolution – Aerosol optical depth (AOD; in other sources – aerosol optical thickness, AOT) dataset and its associated pixel-level uncertainties (AODunc), fine mode AOD (FMAOD), dust AOD, surface directional reflectance datasets for 440nm, 550nm, 670nm, 865nm, 1610nm, 2250nm, Angstrom Exponent (AE) for 440-879nm and 500-870nm, cloud fraction for nadir and oblique views.

The validation approach for satellite retrieved aerosol products suggested for the European Space Agency (ESA) Climate Change Initiative (CCI) AOD product validation (Product Validation and Intercomparison Report, PVIR, [https://climate.esa.int/media/documents/Aerosol_cci_PVIR_v1.2_final.pdf](https://climate.esa.int/media/documents/Aerosol_cci_PVIR_v1.2_final.pdf)) and further developed by Sogacheva et al. (2022) in the frame of the ESA LAW project was followed.

Validation period covers three years from 14 January 2020 to 31 December 2022. AOD, AODuns, FMAOD (all at 550nm) and AE were validated against AErosol RObotic NETwork (AERONET). Matchups database (created by ACRI and extended by FMAOD) includes collocation with AERONET for quality controlled SY_2_AOD (syAOD) product.

Validation was performed for:

* The whole product
  * Groups of pixels retrieved with dual or single (applied to nadir or oblique) view approaches, depending on the SLSTR and OLCI coverage and L1B data availability in different viewing angles (North and Heckel, 2019). Dual-view processor (dual) has been applied when SLSTR measurements from both views, nadir and oblique, were available. If measurements were available from one view only, the single view processor was applied to either nadir (singleN, over either land or ocean) or oblique view (singleO, over ocean or inland waters only).

  * Different regions, specified in Figure 179

  * Different aerosol types. Aerosol types were defined with AERONET AOD (aAOD) and AERONET AE (aAE) thresholds. Although these thresholds are subjective, we consider “background” aerosol to be cases where $\text{aAOD}_{550} \leq 0.2$, “fine-dominated” with $\text{aAOD}_{550} > 0.2$ and $\text{aAE} > 1$, and “coarse-dominated” with $\text{aAOD}_{550} > 0.2$ and $\text{aAE} < 1$ (e.g. Eck et al., 1999).

  * Different seasons. Winter in the NH has been combined with winter in the SH, etc.

  * Separate years to reveal, based on validation results, if drift of the instrument exists
Figure 179: Land and ocean regions defined for this study (as in Sogacheva et al., 2020): Europe (Eur), Boreal (Bor), northern Asia (AsN), eastern Asia (AsE), western Asia (AsW), Australia (Aus), northern Africa (AfN), southern Africa (AfS), South America (SA), eastern North America (NAE), western North America (NAW), Indonesia (Ind), Atlantic Ocean dust outbreak (Aod), Atlantic Ocean biomass burning outbreak (AOb). In addition, Southeast China (ChinaSE), which is part of the AsE region, marked with a blue frame, is considered separately. Land, ocean and global AOD were also considered.

7.2.3.1 AOD

Scatter density plots for S3A and S3B AOD$_{550}$ matchups between AERONET (aAOD) and Sy_2 (syAOD) for different groups of products: all matchups, pixels retrieved with dual, single applied to nadir (singleN) and single applied to oblique (singleO) approaches are shown globally in Figure 180.

Over land, overall correlation (R, Pearson) with AERONET is 0.59/0.63 for S3A and S3B, respectively (from here onwards, results, if shown for both instruments, will be presented as S3A/S3B). The performance shows a high rms error, 0.25/0.23 when compared to AERONET. Reduced performance over land is expected, since the surface reflectance and angular distribution of scattering is higher and more difficult to predict than ocean reflectance. Correlation coefficients are considerably higher (0.89/0.89) for pixels retrieved over ocean with the singleO approach. Second-high correlation (0.66/0.68, but highest rms 0.33/0.29) is observed for singleN approach. Results for dual view show similar to the whole product R (0.59/0.61), but lower rms (0.19/0.18).
Figure 180: Globally, for S3A (left) and S3B (right), scatter density plots for $\text{AOD}_{550}$ matchups between Sy_2 and AERONET for different groups of products: all matchups, pixels retrieved with dual, single applied to nadir (singleN) and single applied to oblique (singleO) approaches (top down).

Fractions of matchups which fit to MODIS expected error (EE, ±0.05±0.2*AOD) envelope or which satisfy GCOS requirements (of 0.03 or 10%*AOD) is low. For the whole product, EE is 52%/58%, GCOS is 22%/28%.

Analysis of the binned (with 0.1 AOD step, based on aAOD) S3A AOD offset a (Figure 181, left panel; all groups in one plot) shows that positive AOD offset (of ~0.1 for dual and singleO, up to 0.2 for singleN) for
AOD<0.4 (>90% of matchups fit into this AOD range). For AOD>0.4, an offset between Sy_2 and AERONET AOD turns to negative for dual group and becomes more positive for singleN, strongly increasing towards high AOD in both groups (note that less than 1.4% of the matchups fit to the range of aAOD>1). For S3B, offsets for different groups are slightly lower (Figure 181, right panel, plots for S3A/S3B for each group separately).

Figure 181: Left: For S3A, binned (based on AERONET AOD) Sy_2 AOD offsets (magenta dots in Figure 180) for different groups of matchups: all, dual, singleN and singleO. Right: For S3A and S3B, for different groups of matchups, binned offset to AERONET (colours for S3A are as on the left panel; S3B is shown in grey).

It can be seen from Figure 180 and Figure 181 that the behaviour of the two instruments is very similar, even if statistics are slightly better for OLCI-B (Figure 180). This is why, to preserve conciseness, the results presented hereafter are not always provided for the two platforms; however, they can be considered as applicable to both.

Binned syAOD offsets for selected regions are shown in Figure 182:. For dual pixels (upper panel), negative (for AOD>0.4) AOD offset is increasing faster towards high AOD in Europe (Eur), where AOD is, on average, around 0.2, and dust-exposed regions over both land (North Africa, AfN) and ocean (middle Atlantic, AOd), where AOD is reaching high values during the dust-transport episodes. Offsets are smaller for Boreal (Bor) regions and South America (SA). For singleN group of pixels, dAODs are positive, with some exceptions in Indonesia (Ind) and Bor. Exceptions can be explained by local fire events (or fire emission transport) over AERONET stations, while satellite AOD is averaged over 3*3 neighboring syAOD pixels which could have been mostly fire-free.
Figure 182: Binned syAOD offset for selected regions (see legend) for dual (upper panel) and singleN (lower panel) groups of matchups.

For the aAOD binned on 0.1 intervals, the global difference (dAOD) between syAOD and aAOD represented with the median bias and dAOD standard deviation is shown in Figure 183 for all aerosol types (including background (aAOD ≤ 0.2) AOD), fine-dominated and coarse-dominated AOD. Globally, background AOD (77% from all matchups) is overestimated by 0.04-0.06. Overestimation of fine-dominated matchups is increasing from 0.07 to 0.15 in the AOD range of 0.2-1.2 (18% of matchups). Overestimation for coarse-dominated matchups is about 0.05 for aAOD<0.7; for aAOD of 0.7-0.9, an overestimation for coarse-dominated matchups is within the GCOS requirements of <0.03 dAOD. For aAOD>1.2, dAOD is varying in the sign and in amplitude; however, the number of matchups in this size range is low (<1 %) and results are thus unstable. Fractions of the fine-dominated matchups is 60-70% for aAOD in the range of 0.2-0.9 and more than 70% for aAOD>0.9. Thus, binned offsets for all matchups follow closely offsets for fine-dominated matchups.
Figure 183: Global difference ($d_{\text{AOD}_{550}}$) between syAOD and aAOD for aAOD binned in 0.1 intervals: median bias (circles) and bias standard deviation (error bars) for all and background (aAOD ≤ 0.2) AOD types (purple), aerosol fine-dominated AOD (blue) and coarse-dominated AOD (green). The fraction ($F$) of points in each bin from the total number of matchups is represented with orange bars. The fraction of fine-dominated matchups in each bin is shown as the blue dashed-line.

Global seasonal scatter density plots with corresponding validation statistics are shown in Figure 184. Note, that seasons are combined from the season-corresponding months in both hemispheres (e.g., winter months are December, January and February in the Northern hemisphere and June, July and August in the Southern Hemisphere). Though $R$ is highest in spring, the fraction of positive outliers (and rms) is highest too, which may be result of snow contaminated signal. A “cloud” of strongly underestimated syAOD (in syAOD range <0.3) is clearly recognized in winter, spring and summer. Fraction of matchups in EE and fractions of matchups which satisfy GCOS requirements are higher in winter and summer.
We used validation statistics calculated for 3 consequent years (2020, 2021, 2022, all years starting from 14th of January) to reveal if instrument drift exists. As it is seen from the annual statistics (Figure 185), annual statistics are worsening slightly from year to year for both S3A and S3B AOD products. For S3A, R is 0.61-0.58-0.58, GCOS fraction is 24.8%-24.2%-23.5% for years 2020-2021-2022, respectively.

Figure 184: For S3A, seasonal scatter density plots with corresponding validation statistic.
Figure 185: For S3A (left panel) and S3B (right panel), seasonal–annual scatter density plots with corresponding validation statistic for three consequent years, 2020, 2021 and 2022.

To better understand where the difference in validation statistics between years comes from, we plotted binned syAOD offsets for separate years (Figure 186). For both S3A and S3b, low AOD (<0.4) offsets are within ±0.02 between the years. Interannual differences are growing considerably for aAOD>1.2. Thus, differences in interannual statistics are caused by annual syAOD outliers and thus, most likely, do not point to instrument drift.
We looked deeper at the inter-annual differences in syAOD “extreme” outliers (Figure 187). SyAOD was considered as extreme outlier if absolute difference between syAOD and aAOD was above 1.0 AOD. For S3A, the number of outliers (Figure 187, upper panel) has increased from year 2020 (N=206) to year 2021 (N =247), but then lowered again (N = 216). Similar tendency was observed in the number of AERONET stations, where “extreme” outliers are observed. Fraction of the “extreme” outliers per AERONET station (from the total number of matchups per station) is shown in Figure 187, lower panel. Number of AERONET station with extreme syAOD outliers is gradually decreasing in the US from year 2020 to year 2022. Number of AERONET stations with high (up to 30-40%) fraction of the extreme outliers is increasing the AfN (Sahara) region.

Figure 186: For S3A (left) and S3B (right), annual binned syAOD offsets for years 2020, 2021 and 2022

Figure 187: For S3A, scatter density plot for syAOD$_{550}$ “extreme” (>1.0AOD) outliers July, years 2020, 2021 and 2022 (upper panel) and for AERONET stations fraction of outliers for corresponding periods (lower panel)
Visual inspection of the S3A AOD scenes with syAOD “extreme” outliers. More than 200 cases were inspected. Typical syAOD spatial distribution for positive syAOD outlier is shown in Figure 188. For the majority of the cases, AOD outliers were retrieved in possibly cloud contaminated area. Thus, cloud screening should be improved to recognize better cloud-contaminated and/or cloud edge areas.

**Figure 188**: Examples of syAOD spatial distribution for the cases when matchup with AERONET station (pixels within the magenta circles) is classified as an “extreme” syAOD outlier.

### 7.2.3.2 AODunc

Below are definitions for uncertainties considered in the current exercise of the evaluation of uncertainties:

- **Prognostic uncertainties (PU)** for Sy₂₂_AOD product are provided at 440, 550, 670, 865, 1600 and 2250 nm wave lengths.

- **Expected discrepancy (ED)** is an uncertainty variable which accounts for the accuracy of the ground-based data (AERONET, AU), as defined by Sayer et al. 2020

  \[ ED = \sqrt{PU^2 + AU^2} \] (4)

- **AOD error (AODerror)** is a difference between satellite product AOD (syAOD) and AERONET AOD (aAOD)

- **AOD absolute error (absAODerror)** is an absolute value for AODerror

- **Normalised error (NE)** is a ratio of the ED and AODerror. If the uncertainty is a good representation of the expected discrepancy, \( \Delta \) will be normally distributed so that a fraction 68.3% of values should fall within the range \([-1, +1]\), with zero mean and unit standard deviation of 1. A non-zero mean indicates the presence of residual systematic errors (an issue for algorithm development). A standard deviation less than one indicates uncertainties are overestimated (NE fraction > 68.3%), which could result from overestimation of individual sources of error, while a standard deviation greater than one indicates an underestimate (NE fraction < 68.3%), e.g. due to neglecting an important source of error (shared from Aerosol_cci+, PVIR).
The goodness of the PU was estimated with the \(\chi^2\) test, as in eq. (5)

\[
c^2 = \frac{1}{N-1} \sum_{i=1}^{N} \overline{d_i} \tag{5}
\]

where individual weighted deviation \(\overline{d_i}\) is described in eq.(6).

\[
\overline{d_i} = \frac{\left(\text{syAOD}_i - \text{aAOD}_i - \text{mean(\text{syAOD}_i - \text{aAOD}_i)}\right)^2}{\text{PU}_i^2 + \text{AU}^2} \tag{6}
\]

If \(\chi^2 \sim 1\), PU describes well the AODerror. If \(\chi^2 >> 1\), PU are strongly underestimated; if \(\chi^2 << 1\), PU are strongly overestimated. \(\chi^2\) was calculated for the whole dataset and for different AOD bins to reveal if the goodness of the PU uncertainties is AOD dependent.

For the whole dataset, \(\chi^2 = 2.8\), which means that PU are slightly underestimated. For the binned AOD, \(\chi^2\) is varying strongly (Figure 189). For aAOD<0.4, which is ca 90% of all values, \(\chi^2\) fits into the interval [1.8 3.2]. Thus, for most of the matchups, PU is only slightly underestimated. For AOD>0.4 PU underestimation is more pronounced.

Though the number of the matchups in the whole dataset is high (which provides the confidence to \(\chi^2\) test results), it was noticed that high \(\overline{d_i}\) (up to 155) exists, which may bias the evaluation of the PU with \(c^2\). To remove possible contribution of the outliers on the \(\chi^2\) test results, cases with \(\overline{d_i} > 10\) (which are less than 5% of the total number of matchups) were removed from the analysis.

For the whole dataset with the removed outliers, \(\chi^2 = 1.2\), which means that PU describe well the AODerror.

Influence of \(\overline{d_i}\) outliers is more pronounced for AOD bins, where the number of matchups per bin is lower and thus the contribution of the outliers to the results is more expected. If \(\overline{d_i}\) outliers are removed from the binned analysis, \(\chi^2\) fits to the range [1 1.45] for AOD<0.4 (Figure 189).

![Figure 189: \(\chi^2\) for binned aAOD for all available matchups (magenta line) and after the outliers of the individual weighted deviations (\(\overline{d_i} > 10\)) are removed (red line). Density scatter plot for PU and syAOD.](image)

To qualitatively illustrate an accuracy of PU, we show in Figure 190 a comparison between the prognostic uncertainties, AOD error (which is a difference between syAOD and aAOD) distribution and theoretical Gaussian distribution (with a mean of 0 and standard deviation of the AODerror) for the whole Sy_2 AOD product, as well as for the groups of pixels retrieved with different (dual- or single-view) approaches. AOD error distributions are Gauss-like with partly some asymmetry in positive AODerror direction. PU distribution shows a double peak, (first peak is at ca. 0.02-0.04 for all groups; the second peak in a range of 0.12-18, for different groups). For singleN, two peaks are located close to each other. Mean PU for dual
group is higher; standard deviation (std) is higher for singleN group. AOD error distributions are Gauss-like with partly some asymmetry in positive AODerror direction.

Figure 190: Comparison between prognostic uncertainties, AOD error distribution and theoretical Gaussian distribution for the whole product (left panel), dual- (middle panel) and singleN (right panel) groups of matchups.

Comparison between the normalized error (NE, which is a ratio between AODerror and expected discrepancy, ED) and theoretical Gaussian distribution (with a mean of 0 and standard deviation of NE) is shown in Figure 191. Positive NE means that syAOD was higher than aAOD. Absolute NE<1 means that AOD error was lower than ED (or PU, since the contribution of AERONET error to the ED is not significant). The high peak of small positive NE, considerable positive bias, especially for singleN group, and longer positive than negative tail, negative tail below expected Gaussian distribution (positive tail is longer than negative) say that the fraction of small (<1) NE, when actual error is smaller than expected discrepancy, is higher than expected and ED, which is defined mostly from the PU, is overestimated (for NE<<1). Clear difference exists in NE footprints for dual- and single-view groups of pixels. PU was considerably underestimated for the cases when syAOD was lower than aAOD.

Figure 191: Comparison between normalized error and theoretical Gaussian distribution for the whole product (left panel), dual- (middle panel) and singleN (right panel) groups of matchups.

Sayer, et al. (2020) suggested the analysis of the potential of the PU to discriminate between (“good” / “bad”) pixels with likely small / large errors. Instead of PU, we perform analysis of the ED, which, besides PU, includes uncertainties of the ground-based measurements. To estimate the potential of ED, we plot the absolute errors which are 38%, 68% and 95% of all pixels with this uncertainty as a function of binned ED (Figure 192). These percentages relate to 0.5 σ, 1 σ, and 2 σ (σ = standard width) for normal error distributions in each bin (along the vertical axis). Those theoretically expected values are shown as dashed lines in black, red and blue. The amount of ED hits per bin is shown as a grey dashed line.
The percentile plots show a reasonable agreement (within statistical noise) with the theoretical lines of 38% and 68% for majority of the validation points in the lower range of ED (up to 0.05–0.1) for all groups. 98% show the overestimation of ED. For higher uncertainties the error values for all and dual groups are clearly below the expected lines, which means that ED is too large. For singleN, ED>~ 0.2 is underestimated.

Figure 192: Percentile plots of absAODerrors at 38% (black), 68% (red) and 95% (blue) as function of binned expected discrepancy.

7.2.3.3 AE

The Ångström exponent, AE, is often used as a qualitative indicator of aerosol particle size.

Satellite AE is calculated in the spectral interval 550–865 nm and AERONET AE is provided for 500–870 nm. The difference between AE_{550-865} and AE_{500-870} depends on the aerosol type and may be as high as 5–10% of AE (personal estimations). This difference must be considered for the interpretation of the evaluation results.

Scatter plots between sAE_{550-870} and aAE_{500-870} for S3A and S3B, for different groups of products (top-down: all, dual, single, singleN and singleO) and corresponding validation statistics are shown in Figure 193 for both S3A and S3B. Validation results for S3A and S3B products are similar.

Two “clouds” of satellite/AERONET AE matchups can be recognized. The first cloud (cloud1) is in the aAE interval of [1.2 1.6]. In that interval, the cloud of pixels is located around 1:1 line, which means that the agreement between syAE and aAE is quite good. The second cloud (cloud2) is in the aAE interval [1.4 1.9]. In that interval, syAE is overestimated by 0.3–0.6. For the whole global product, correlation coefficients between sAE_{550-870} and aAE_{500-870} are quite low, 0.33/0.32, rms is high, 0.61/0.62 for S3A/S3B, respectively. Validation statistics are slightly better for dual product. SingleO product shows best correlation (0.84/0.834), but worse rms and std.
Figure 193: For S3A (left panel) and S3B (right panel), global scatter plots between syAE$_{550-865}$ and aAE$_{500-870}$ for different groups of products (top-down: all, dual, singleN and singleO).

Regional differences in syAE evaluation results were revealed (Figure 194). Data the western part of the North America (NAW) mostly form the cloud1. Similar “cloud” is observed in Europe, where syAE overestimation is also rather high. Similar high is observed in Asia and eastern part of the North America. “Dust” regions, AfN and AOd, contribute mostly to cloud2.
Figure 194: For S3A, regional scatter plots between $sy\text{AE}_{550-865}$ and $a\text{AE}_{500-870}$.
7.2.3.4 FMAOD, FMF

Fine mode AOD in the SY_2 product (syFMAOD) is provided at 550nm, while AERONET Fine mode AOD (aFMAOD) is provided at 500 nm. As for aAOD, AOD spectral dependence (https://aeronet.gsfc.nasa.gov/new_web/man_data.html, O'Neill et al., 2003) and AERONET AE exponent were considered to convert aFMAOD$_{500}$ into aFMAOD$_{550}$.

Sy_2 fine mode AOD (syFMAOD) and Sy_2 Fine Mode Fraction (syFMF), which is a fraction of syFMAOD from the total syAOD, was validated against AERONET Fine Mode Fraction (aFMF). Density scatter plots for the relation between syFMAOD and aFMAOD, syFMF and aFMF, are shown in Figure 195 for S3A.

As syFMAOD is overestimated for low (<0.7) aFMAOD, we expect overestimation of syFMF, which is clearly seen. syFMF is overestimated in the aFMF range of 0-0.7; positive offset of 0.3-0.5 at low (<0.25) aFMF is gradually decreasing. At aFMF>0.9, syFMF is slightly underestimated.

![Figure 195: Scatter density plots for S3A syFMAOD and corresponding aFMAOD for matchups available globally.](image)

7.2.3.5 Conclusions

Extension of the Sy_2 _AOD validation period 2020-2021 (Sogacheva et al., 2022) with year 2022 have not change validation statistic considerably. The main results are:

- Sy_2 AOD product quality (over land) is below GCOS requirements.
- Validation statistics for AOD and AOD uncertainties are different for pixels retrieved with different approaches (dual, single).
- Validation statistics are slightly better for S3B.
- Small inter-annual differences in validation statistics are determined by extreme syAOD outliers.
- Cloud screening should be considerably revised.

7.2.3.6 References


8 Problems encountered in the reporting period

8.1 Product Notices Reports

Product notices are issued when a new processing baseline is deployed. Hereafter, for each instrument, the list of product notices issued are displayed.

8.1.1 OLCI

Table 28: List of OLCI Product Notices issued in 2022

<table>
<thead>
<tr>
<th>Level</th>
<th>Thematic</th>
<th>Reference</th>
<th>Date of issue</th>
<th>Version</th>
<th>Last update</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Global</td>
<td>S3.PN.OLCI-L1.09</td>
<td>20/01/2022</td>
<td>1.0</td>
<td></td>
<td>PB 3.02 (S3A/S3B)</td>
</tr>
<tr>
<td>L1</td>
<td>Global</td>
<td>S3.PN.OLCI-L1.10</td>
<td>29/08/2022</td>
<td>1.0</td>
<td></td>
<td>PB 3.10 (S3A/S3B)</td>
</tr>
<tr>
<td>L1</td>
<td>Global</td>
<td>S3.PN.OLCI-L1.11</td>
<td>In preparation</td>
<td></td>
<td></td>
<td>PB OL__L1_.003.01.00 (LAND) and_L1_.002.24.00 (MARINE)</td>
</tr>
</tbody>
</table>

8.1.2 SLSTR

Table 29: List of SLSTR Product Notices issued in 2022

<table>
<thead>
<tr>
<th>Level</th>
<th>Thematic</th>
<th>Reference</th>
<th>Date of issue</th>
<th>Version</th>
<th>Last update</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Global</td>
<td>S3.PN.SLSTR-L1.09</td>
<td>18/01/2022</td>
<td>1.0</td>
<td>20/01/22</td>
<td>PB 3.00 (S3A/S3B) - roll-out in ESA centre on 20/01/22</td>
</tr>
<tr>
<td>L2</td>
<td>FRP</td>
<td>S3.PN.SLSTR-FRP.03</td>
<td>28/02/2022</td>
<td>1.0</td>
<td></td>
<td>FRP_NTC.004.07.00 (PB 3.03)</td>
</tr>
<tr>
<td>L2</td>
<td>Land</td>
<td>S3.PN.SLSTR-L2L.05</td>
<td>16/09/2022</td>
<td>1.0</td>
<td></td>
<td>PB SL__LST.004.07.00</td>
</tr>
<tr>
<td>L1</td>
<td>Global</td>
<td>S3.PN-SLSTR-L1.10</td>
<td>In preparation</td>
<td></td>
<td></td>
<td>PB SL__L1_.004.05.00</td>
</tr>
</tbody>
</table>
8.1.3 SYN

Table 30: List of SYN Product Notices issued in 2022

<table>
<thead>
<tr>
<th>Level</th>
<th>Thematic</th>
<th>Reference</th>
<th>Date of issue</th>
<th>Version</th>
<th>Last update</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>Land</td>
<td>S3.PN.SYN-L2.08</td>
<td>16/09/2022</td>
<td>1.0</td>
<td>28/10/22</td>
<td>PB SYN_L2_.002.16.00 (updates required on known limitations)</td>
</tr>
<tr>
<td>L2</td>
<td>Land</td>
<td>S3.PN.SYN-L2.09</td>
<td>In preparation</td>
<td></td>
<td></td>
<td>PB SYN_L2_.002.17.00</td>
</tr>
<tr>
<td>L2</td>
<td>AOD</td>
<td>S3.PN.SYN-AOD.02</td>
<td>In preparation</td>
<td></td>
<td></td>
<td>PB AOD_NTC.002.07.00</td>
</tr>
</tbody>
</table>

8.2 Instrument anomalies

8.2.1 OLCI

The OLCI anomalies or events recorded by the S3MPC operators in 2021 are displayed in 2 forms:

❖ A calendar view, in Figure 196
❖ A table providing more details is available at

https://sentinel.esa.int/documents/247904/3810702/OLCI-Anomaly-Events.pdf
8.2.2 SLSTR

The SLSTR anomalies or events recorded by the S3MPC operators in 2021 are displayed in 2 forms:

❖ A calendar view, in Figure 197
❖ A table providing more details is available at

https://sentinel.esa.int/documents/247904/3810820/SLSTR-Anomaly-Events.pdf
8.2.3 SYN

The SYN anomalies or events recorded by the S3MPC operators in 2021 are displayed in 2 forms:

- A calendar view, in Figure 198
- A table providing more details is available at

https://sentinel.esa.int/documents/247904/3810953/SYNergy-Anomaly-Events.pdf
Figure 198: SYN anomalies/events in 2021

End of document