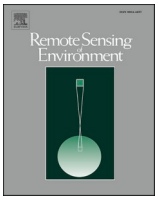
Remote Sensing of Environment 299 (2023) 113844

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse



# Radiometric assessment of OLCI, VIIRS, and MODIS using fiducial reference measurements along the Atlantic Meridional Transect

Silvia Pardo [a,*], Gavin H. Tilstone [a], Robert J.W. Brewin [a,b], Giorgio Dall'Olmo [a,c], Junfang Lin [a], Francesco Nencioli [d], Hayley Evers-King [e], Tânia G.D. Casal [f], Craig J. Donlon [f]

[a] Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth PL1 3DH, UK
[b] Centre for Geography and Environmental Science, Department of Earth and Environmental Science, Faculty of Environment, Science and Economy, University of Exeter, Penryn Campus, Penryn TR10 9FE, UK
[c] Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (OGS), Borgo Grotta Gigante, Trieste, 34010, Italy
[d] Collecte Localisation Satellites (CLS), No.11 Rue Hermès, Parc Technologique Du Canal, 31520 Ramonville Saint-Agne, France
[e] European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), 64295 Darmstadt, Germany
[f] European Space Agency (ESA) - European Space Research and Technology Centre (ESTEC), Keplerlaan 1, 2201 AZ Noordwijk, the Netherlands

ARTICLE INFO

Edited by Menghua Wang

*Keywords:*
Sentinel-3
OLCI
Validation
Atmospheric correction
Atlantic Ocean
MODIS
VIIRS
Sentinel-3A and 3B tandem phase
Water leaving radiance
Ocean colour

ABSTRACT

High quality independent ground measurements that are traceable to metrology standards, with a full uncertainty budget, are required for validation over the lifetime of ocean-colour satellite missions. In this paper, we used radiometric Fiducial Reference Measurements (FRM) collected during four Atlantic Meridional Transect (AMT) field campaigns from 2016 to 2019 to assess the performance of radiometric products from the Ocean and Land Colour Instrument (OLCI) aboard Sentinel-3A (S-3A) and 3B (S-3B), the Moderate Resolution Imaging Spectroradiometer instrument aboard Aqua (MODIS-Aqua), and the Visible Infrared Imaging Radiometer Suite instrument aboard Suomi NPP and NOAA-20 (Suomi-VIIRS and NOAA-20 VIIRS). The AMT provides one of the few sampling platforms that make high-quality *in situ* radiometric measurements in oligotrophic, low chlorophyll-a oceanic waters for ocean colour satellite validation. *In situ* data were acquired and processed following established FRM protocols, calibrated to metrology standards, referenced to inter-comparison exercises and with a full uncertainty budget. From these we selected an uncertainty threshold, which we used as part of a matchup procedure that takes into account the temporal and spatial variability of both the *in situ* and satellite data. Three atmospheric correction models were compared for S-3A and S-3B OLCI radiometric products; the standard OLCI IPF-OL-2, POLYMER and NASA SeaDAS l2gen. Based on the round-robin comparison, POLYMER provided the best performance in the retrieval of water-leaving radiances. The analysis showed that Suomi-VIIRS and MODIS-Aqua performed better than NOAA-20 VIIRS, and comparably with S-3B OLCI standard products. The S-3A OLCI standard product outperformed the NASA products. The S-3A OLCI and S-3B OLCI instruments were also compared during their tandem phase, which showed that S-3B OLCI radiances were systematically higher than S-3A OLCI across the spectrum.

## 1. Introduction

The study of the green pigment chlorophyll-a (Chl-a), a proxy for phytoplankton biomass, is fundamental to our understanding of the role of the oceans in the marine carbon cycle, the flow of carbon from the base of the trophic web to fish and marine mammals, and the response of marine algae to a changing climate (Werdell et al., 2018). Since 1978, when the first experimental ocean-colour satellite Coastal Zone Colour Scanner (CZCS) was launched as a proof of concept, Chl-a has been observed from space. From 1997, improvements in the design of ocean-colour satellites and algorithms applied to the data, resulted in quantitative estimates of Chl-a through a succession of missions planned and operated by the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA) and the National Oceanic and Atmospheric Administration (NOAA) among other aerospace agencies. The NASA ocean-colour missions carrying the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and Moderate Resolution Imaging Spectroradiometer (MODIS) and ESA mission carrying the Medium Resolution

* Corresponding author.
*E-mail address:* spa@pml.ac.uk (S. Pardo).

https://doi.org/10.1016/j.rse.2023.113844
Received 22 December 2022; Received in revised form 25 September 2023; Accepted 2 October 2023
Available online 31 October 2023

Imaging Spectrometer (MERIS) delivered synoptic coverage of changes in phytoplankton biomass on a global basis, at unprecedented temporal and spatial scales that had not previously been achievable using *in situ* or airborne sampling platforms. The combined data sets cover more than twenty-five years and have provided critical insight into global ocean chlorophyll trends (Martinez et al., 2009; Gregg et al., 2017; Siemer et al., 2021) and the effects of climate change on marine phytoplankton (Siegel and Franz, 2010; Dutkiewicz et al., 2019; Kulk et al., 2020). The fundamental measurements behind this ocean-colour record are the remote-sensing reflectance, $R_{rs}$, and the water-leaving radiance, $L_{WN}$. In the open ocean, empirical relationships are applied to specific $R_{rs}$ band ratios in the visible spectrum to estimate Chl-a. Using this approach, ocean-colour satellite Chl-a has a nominal uncertainty of ±35%, as long as the target uncertainty in $R_{rs}$ at 442 and 560 nm of ±5% is met (Moore et al., 2009). To verify whether the satellite $R_{rs}$ uncertainty requirement is met, it is necessary to evaluate the accuracy of radiometric products using independent, high-quality *in situ* datasets over the life time of the satellite mission, which for the ESA Sentinel-3 Ocean and Land Colour Instrument (OLCI), is expected to be ∼20 years.

### 1.1. The need for fiducial reference measurements

Comprehensive comparisons of satellite data with field measurements are a fundamental way of assessing the accuracy of satellite missions (McClain et al., 2004). Due to the diversity in the percentage uncertainty and the principal factors that contribute to it, validation of satellite $L_{WN}$ and $R_{rs}$ requires accurate *in situ* datasets covering a wide dynamic range of water properties, which take a significant amount of time and resources to collect (Barnes et al., 2019). For the NASA missions, this has been achieved through the use of SeaBASS (SeaWiFS Bio-optical Archive and Storage System) and NOMAD (NASA bio-Optical Marine Algorithm Data set) databases (Werdell and Bailey, 2005). NOMAD, like many *in situ* databases, does not represent well the global distribution of Chl-a, because there are more data collected in coastal regions than in the less accessible open ocean. The large errors in satellite data associated with high Chl-a values inflate the overall error calculated from the NOMAD data (Moore et al., 2009). Additionally, uncertainties in the *in situ* data can be highly variable depending on the regions, conditions, and platforms from which they are taken. In some cases these uncertainties can be larger than those in the satellite data being validated (Bailey and Werdell, 2006; Hooker and Maritorena, 2000), and disentangling the uncertainties arising from *in situ* and satellite sources can be problematic (Barnes et al., 2019). The spatial distribution of the validation stations within these databases can also be very uneven, due to sparse coverage in open-ocean regions (Gregg and Casey, 2004).

At the launch of S-3A, the European Space Agency recognised the need for field radiometric measurements of the highest accuracy. Based on the Guide for Expression of Uncertainty in Measurement (GUM, JCGM (2008)) and Zibordi and Voss (2014), ESA defined such highly-accurate measurements as Fiducial Reference Measurements (FRM): measurements that conform to accepted satellite protocols, that are traceable to metrology standards, referenced to intercomparison exercises, with a full uncertainty budget to provide independent, high-quality, satellite validation measurements for the duration of a satellite mission. The uncertainty budget should account not only for the performance of the radiometer, but also for the calibration source and its transfer, and should incorporate the effect of environmental variability and deployment conditions (Zibordi and Voss, 2014). There is also a requirement for radiometer intercomparison exercises, which guarantee data consistency among investigators and minimise uncertainties that can, in turn, affect our assessments of the accuracy of satellite missions (Zibordi et al., 2012; Tilstone et al., 2020). The FRM methodology is widely established for fixed platforms that are mainly located at coastal locations (Zibordi et al., 2009a,b) and also recently on moving ships that access the open ocean (Lin et al., 2022).

### 1.2. Radiometric assessment of Sentinel-3 OLCI data

S-3A was launched in February 2016, and was followed by S-3B in April 2018. Most papers published to date on the accuracy of Sentinel-3 ocean-colour data are focused on the assessment of derived products (Liu et al., 2018; Kyryliuk and Kratzer, 2019; Harshada et al., 2021; Tilstone et al., 2021; Kratzer and Plowey, 2021; Masoud, 2022; Salama et al., 2022). Of the papers focusing on radiometric validation, Zibordi et al. (2018) and Tilstone et al. (2022) found a systematic underestimation of OLCI Collection 2 $L_{WN}$ and $R_{rs}$, particularly in the blue and red bands. Zibordi et al. (2018) attributed this negative bias to errors in the estimation of OLCI aerosol optical depths at 865 nm resulting from an overestimate in OLCI-derived Ångström exponents.

These issues were addressed in the development of OLCI Collection 3 (OLCI OL L2M 003), deployed in February 2021, which introduced major updates to the system vicarious calibration gains, bright pixel correction, ocean-colour algorithms, masking, and flag recommendations (EUMETSAT, 2021a). Few papers have been published assessing OLCI Collection 3. Tilstone et al. (2021), Tilstone et al. (2022) and Zibordi et al. (2022) report an improvement in the performance of $R_{rs}$ and $L_{WN}$ retrievals at blue and green wavelengths for oligotrophic and mesotrophic waters, and systematic negative biases in $L_{WN}$ for CDOM-dominated and other optically complex waters. Zibordi et al. (2022) also report a strong dependence of $L_{WN}$ on the viewing angle for S-3A OLCI.

Other OLCI performance assessments have focused on the intercomparison of atmospheric correction schemes over optically complex waters. Several atmospheric corrections algorithms are currently available for the OLCI instrument: the baseline atmospheric correction (BAC) and alternative atmospheric correction (AAC) neural network (NNv2) as deployed in the operational level 2 OLCI processor (IPF-OL-2), the case 2 regional coast colour atmospheric correction algorithms C2R-CC and C2R-CCAltNets (Brockmann et al., 2016), the polynomial-based algorithm developed for MERIS (POLYMER, Steinmetz et al., 2011), the Dark Spectrum Fitting (DSF) atmospheric correction (Vanhellemont and Ruddick, 2018), and the standard NASA SeaDAS l2gen atmospheric correction. The AC intercomparison study conducted by Mograne et al. (2019) in French coastal waters found that C2R-CC and POLYMER performed best in independent and coincident matchup analysis, respectively. Similarly, Giannini et al. (2021) found that C2R-CC presented the best radiometric performance for Northeast Pacific coastal waters. Renosh et al. (2020) results for highly turbid estuarine waters indicate that most AC algorithms tend to underestimate $L_{WN}$ in the green, red and NIR parts of the spectrum. Windle et al. (2022) demonstrated that C2R-CC performed best in Chesapeake Bay, but POLYMER maximised the number of clear retrievals without substantially compromising performance.

A few weeks after its launch, from 6 June to 15 October 2018, the S-3B satellite was positioned into a tandem arrangement in orbit alongside its operational twin, S-3A. During this tandem phase, both satellites were flown in a synchronized manner on the same orbit ground track, only separated by a ∼30 second interval. The S-3A/B tandem phase presented an exceptional chance to enhance understanding of differences in payloads, standardize datasets and minimise data uncertainties as described by Lamquin et al. (2020a,b) and Clerc et al. (2020). From 16 October to 22 November 2018, S-3B underwent a short drift phase where the satellite was gradually shifted to its final orbit phase, 140° separated from S-3A.

The objective of this paper is, firstly, to assess the performance of the S-3A and -3B standard processing baseline $L_{WN}$ products (OLCI OL_L2M_003) in the open ocean; secondly, to intercompare the AC models available for OLCI, thirdly, to intercompare S-3A and 3B during their tandem phase, and finally, to compare the performance of S-3A and 3B OLCI $L_{WN}$ with MODIS-Aqua, Suomi-VIIRS and NOAA-20 VIIRS.

## 2. Data and methodology

### 2.1. In situ data

During the AMT campaigns, radiometric data were collected using a number of different radiometer systems. The primary instrument was a SeaBird HyperSAS system (Lin et al. (2022), Fig. 1) which collected above-water measurements both underway and whilst the ship was on station. The HyperSAS package is composed of three instruments which measure: radiance from the water surface ($L_t(\lambda)$), radiance from the sky ($L_i(\lambda)$), and downwelling irradiance ($E_s(\lambda)$). Further instrumentation was deployed on different campaigns (both above and in-water and from different manufacturers) to cross-compare performance as part of the FRM characterisation (Tilstone et al., 2020). The auxiliary data wind speed, and position were measured as part of the ship's navigation package.

Data were gathered through daylight hours (8 am to 7 pm) at intervals of ~1 to ~5 seconds, in the spectral range from ~305 nm to ~1142 nm, with a spectral resolution of 10 nm and a spectral step of 3.3 nm. The processing chain was built on that developed by Brewin et al. (2016) and is described in detail in Lin et al. (2022). Data were extracted using the relevant calibration files and any necessary correction for discrepancy between pre- and post cruise calibrations was applied. Once extracted, the individual data files were merged to produce daily files, and measurements were interpolated to common spectral and temporal resolutions. Measurements where tilt $> 5^{\circ}$ were filtered out. After aggregating the data in 2-min bins, a glint filter was applied such that the near-infrared signal was zero (Hooker et al., 2002). Further filtering was conducted for high solar zenith ($\theta > 80^{\circ}$) and relative azimuth ($\phi \notin [100^{o}, 170^{o}]$) angles. Finally, remote sensing reflectances were calculated according to eq. 1:

$$R_{rs}(\lambda) = \frac{L_t(\lambda) - \rho L_i(\lambda) - \Delta L}{E_s(\lambda)}, \quad (1)$$

where $L_t$, $L_i$ and $E_s$ were derived from the three respective instrument packages, $\rho$ is an estimate of the sea surface reflectance, and $\Delta L$ is a spectrally-flat residual term representing radiative contributions due to glint, foam, sea spray and whitecaps (Lin et al., 2022). Remote sensing reflectances $R_{rs}$ were converted to spectral normalised water-leaving radiances $L_{WN}$ according to Eq. 2:

$$L_{WN}(\lambda) = R_{rs}(\lambda)E_0(\lambda)BRDF(\theta_{insitu}, \theta_0, \Delta\phi_{insitu}, \lambda, chl) \quad (2)$$

where $BRDF(\theta_{insitu}, \theta_0, \Delta\phi_{insitu}, \lambda, chl)$ is a term correcting for bidirectional effects (Morel et al., 2002), and $E_0(\lambda)$ is the mean extraterrestrial solar irradiance (Thuillier et al., 2003). A simple linear interpolation was applied to match the *in situ* data to the centre wavelength of the corresponding satellite bands.

The uncertainty for the HyperSAS measurements was computed following the Guide for Expression of Uncertainty in Measurement (JCGM, 2008) as described in Lin et al. (2022). Uncertainty estimates were produced for each *in situ* $L_{WN}$ measurement for the whole spectrum. At 412 nm, 442 nm and 590 nm, $L_{WN}$ percentage uncertainties were typically below 100%, with 65% of the measurements being below 10% uncertainty, and 96% below 25% uncertainty. The mean percentage uncertainty for these bands was around 24%. At 510 nm and 560 nm, the overall percentage uncertainty increased, with 17% and 34% of the measurements for each band having uncertainties above 25%, respectively. At 620 nm and 665 nm, 50% of the measurements showed uncertainties above 100%. The spectral dependence of the *in situ* data uncertainties is assessed in detail by Lin et al. (2022).

Total uncertainty, defined as the 412–600 nm uncertainty average, was used to define uncertainty thresholds in this work. Fig. 1 shows the spatial distribution of this average uncertainty in *in situ* $L_{WN}$ along the AMT26 (2016-09-20 — 2016-11-04), AMT27 (2017-09-23 — 2017-11-05), AMT28 (2018-09-23 — 2018-10-29) and AMT29 (2019-10-13 — 2019-11-25) cruise tracks. Low uncertainty measurements were found along the entire transects, independent of location: while environmental factors caused uncertainty variability during the day, low-uncertainty measurements were obtained on a daily basis at specific times. In particular, a total of 5600 measurements were found to have an uncertainty average between 0% and 10%, and more than half of this total, 3918 out of 5600, were acquired while the ship was moving. A total of 10,735 measurements were in the [10%–25%] range, 4423 in the [25%–50%] range, 2177 in the [50%–100%] range, and 1557 were above 100% total uncertainty.

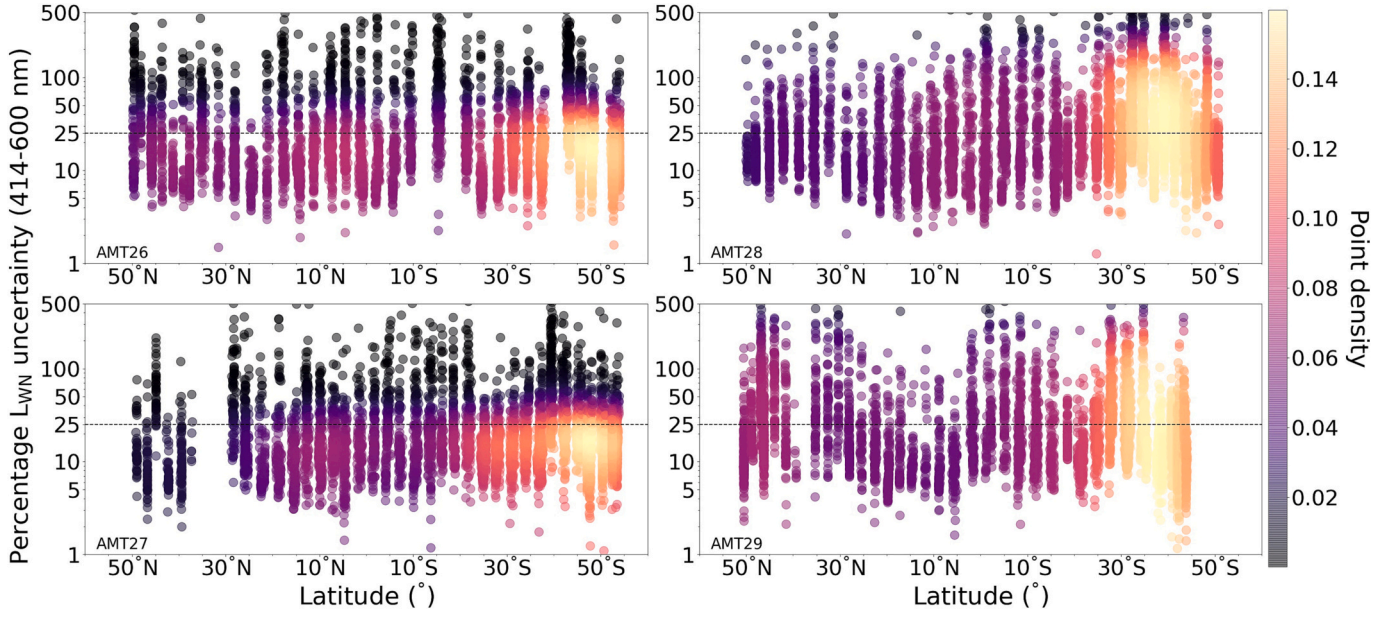



**Fig. 1.** Percentage uncertainty distribution of *in situ* $L_{WN}$ along the AMT26, AMT27, AMT28 and AMT29 campaign tracks. The colours indicate the probability density estimated by a Kernel density function (high density in yellow, low density in black). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 2.2. Satellite data

S-3A and -3B passes acquired within a ± 12-h time window along the AMT26, AMT27, AMT28 and AMT29 cruise tracks provided a total of 143 granules for S-3A and 79 granules for S-3B. The corresponding S-3 OLCI water full resolution (WFRM) data were acquired at levels 1 and 2. Level 1 and 2 products were downloaded from the EUMETSAT Data Centre and Copernicus Online Data Access (CODA) portals, respectively.

Data were also acquired for two non-standard processors: NASA's SeaDAS l2gen and POLYMER. SeaDAS l2gen is the default multi-sensor level 1 to level 2 processor used in the SeaDAS software package (https://seadas.gsfc.nasa.gov). The accuracy of the SeaDAS l2gen atmospheric correction has been extensively tested, *e.g.* Qin et al. (2017), Vanhellemont and Ruddick (2021) and Mélin (2022). SeaDAS l2gen OLCI level 2 data were acquired from the NASA Ocean Biology Distributed Active Archive Center through the Ocean Color Web portal (https://oceancolor.gsfc.nasa.gov). The SeaDAS l2gen OLCI dataset assessed here corresponds to the latest NASA ocean-colour data reprocessing, R2022.0 (NASA OBPG, 2022c, 2022d).

POLYMER (Steinmetz et al., 2011) is an atmospheric correction initially developed for the MEdium Resolution Imaging Spectrometer (MERIS) instrument for the retrieval of ocean-colour parameters in the presence of sun glint. It employs an iterative coupled ocean-atmosphere algorithm where the atmosphere contribution to the top-of-atmosphere reflectances is modelled as a polynomial. POLYMER has been shown to increase coverage, minimise noise and promote temporal stability in satellite time series at both global and regional scales (Müller et al., 2015). For this work we used the latest available version of the processor which, at the time of writing, was POLYMER v4.14. Bespoke processors for case-2 waters, such as C2R-CC (Brockmann et al., 2016) and DSF (Vanhellemont and Ruddick, 2021), were not taken into consideration for this study.

The OLCI standard radiometric product is a directional water-leaving reflectance $\rho_w(\lambda)$, *i.e.*, not corrected for bidirectional effects (EUMETSAT, 2021b). The normalised water-leaving radiances $L_{WN}$ assessed in this work were derived from the standard $\rho_w(\lambda)$ product according to Eq. 3:

$$L_{WN}(\lambda) = \rho_w^{OLCI}(\lambda)\frac{E_0(\lambda)}{\pi}BRDF(\theta_{sat}, \theta_0, \Delta\phi_{sat}, \lambda, chl), \quad (3)$$

where $BRDF(\theta_{sat}, \theta_0, \Delta\phi_{sat}, \lambda, chl)$ is a term correcting for bidirectional effects (Morel et al., 2002), and $E_0(\lambda)$ is the mean extraterrestrial solar irradiance (Thuillier et al., 2003).

As POLYMER and SeaDAS l2gen $R_{rs}$ products are already corrected for bidirectional effects, $L_{WN}$ for these processors were derived according to Eq. 4:

$$L_{WN}(\lambda) = R_{rs}(\lambda)E_0(\lambda), \quad (4)$$

The three processors approach the system vicarious calibration (SVC) of OLCI in different ways. The S-3 OLCI standard processor IPF-OL-2 applies the standard Ocean Colour SVC methodology (Franz et al., 2007) using a combination of South Pacific Gyre and Marine Optical BuoY (MOBY) hyperspectral measurements to compute the SVC gains (EUMETSAT, 2021a; Mazeran and Ruescas, 2021). POLYMER uses a dedicated, *in situ*-based SVC designed for ocean-atmosphere coupled algorithms (Steinmetz and Ramon, 2018). While NASA generally uses *in situ* forcing against MOBY data for vicarious calibration, this method currently results in limited OLCI matchups and is hence unable to remove intermission biases. The ocean reflectance model-based approach (Werdell et al., 2007) is used instead to derive the SeaDAS l2gen OLCI gains (NASA OBPG, 2022g).

Fig. S1 shows the values obtained by the three atmospheric correction models at 442 nm and 560 nm for the 18th of October 2017 S-3A overpass. While spatial patterns were similar, the values obtained by the l2gen processor (bottom row) were distinctly higher than for the other two processors. POLYMER (centre row) was able to coherently retrieve data for most of the area affected by sun glint (in orange).

**Table 1**
Satellite data quality flags used in this study.

| Processor | Flags |
|---|---|
| IPF-OL-2 | CLOUD, CLOUD AMBIGUOUS, CLOUD MARGIN, INVALID, COSMETIC, SATURATED, SUSPECT, HISOLZEN, HIGHGLINT, SNOW ICE, AC FAIL, WHITECAPS, ADJAC, RWNEG O2, RWNEG O3, RWNEG O4, RWNEG O5, RWNEG O6, RWNEG O7, RWNEG O8 |
| POLYMER | IDEPIX INVALID, IDEPIX CLOUD, IDEPIX CLOUD AMBIGUOUS, IDEPIX CLOUD SURE, IDEPIX CLOUD BUFFER, IDEPIX WHITE, IDEPIX CLOUD SHADOW, IDEPIX SNOW ICE, IDEPIX BRIGHT, bright, straylight risk, invalid, cosmetic, sun glint risk, dubious, land |
| SeaDAS l2gen | ATMFAIL, LAND, HIGLINT, HILT, HISATZEN, STRAYLIGHT, CLDICE, COCCOLITH, HISOLZEN, LOWLW, CHLFAIL, NAVWARN, MAXAERITER, ATMWARN, NAVFAIL |

Passes acquired over the areas of interest within a time window of ± 12 h were also identified for MODIS-Aqua, Suomi-VIIRS and NOAA-20 VIIRS, resulting in a total of 165 passes for MODIS, 208 passes for Suomi-VIIRS and 100 passes for NOAA-20 VIIRS. The corresponding level 2 granules were acquired at 1 km (MODIS) and 750 m (VIIRS) resolution from the NASA Distributed Active Archive Center through the Ocean Color Web portal. The datasets assessed here correspond to the latest NASA ocean-colour data reprocessing, R2022.0 (NASA OBPG, 2022a, 2022b, 2022e).

In order to exclude unreliable satellite measurements from the individual products, a set of quality flags recommended for validation were applied as a mask to each pass. Table 1 lists the processing flags applied for each of the processors. A common requirement in validation protocols is that matchups are performed under clear-sky conditions only. Consequently, we were particularly strict in filtering pixels affected by any kind of cloud: for the OLCI IPF-OL-2 processor that meant masking cloud, cloud border and cloud suspect pixels. When flags were available, other sources of contamination such as haze, whitecaps, and sun glint were also taken into consideration.

### 2.3. Matchup procedure

The matchup procedure follows the method outlined in Bailey and Werdell (2006), further refined for use on AMT underway data by Brewin et al. (2016), and extended here to include *in situ* uncertainty values. The following steps were implemented:

- In order to minimise atmospheric and oceanic variability, matchup pixels were extracted from passes acquired within a ± 1 h window over the location of the *in situ* measurement. The matchup exercise was repeated for temporal windows ranging from ± 0.5 h to ± 12 h.
- From the corresponding level 2 ocean-colour products, a 3 × 3 pixel box centred on the coordinates of the *in situ* observation was extracted. To ensure that the matchups were derived from independent satellite pixels, the validation statistics were computed using the central pixel. The remaining data within the 3 × 3 pixel box were used for quality control only.
- To account for sub-pixel variability, the average value of the *in situ* observations (and their associated uncertainties) was computed when a single pixel was matched to multiple *in situ* observations.
- To test the effect of the *in situ* uncertainty on the validation statistics, the validation exercise was repeated for different uncertainty thresholds including no threshold (*i.e.* all available matchups), then for matchup subsets with a maximum uncertainty of 75%, 50%, 25%, 10%, and 5%, while keeping the matchup temporal window (± 1 h) and coefficient of variation threshold (0.15) constant.
- Matchups were excluded if the median coefficient of variation (CV, defined as the standard deviation divided by the mean) of the

**Table 2**
Metrics definitions.

| Metric | Description | Formula |
|---|---|---|
| S | Slope of the Type-II regression[a] | $\frac{\sigma_{EE} - \sigma_{MM} + \sqrt{(\sigma_{MM} - \sigma_{EE})^2 + 4\sigma_{ME}^2}}{2\sigma_{ME}}$ |
| I | Intercept of the Type-II regression | $\bar{X^E} - S\bar{X^M}$ |
| r | Pearson correlation coefficient | $\sum_{i=1}^{n}(X_i^M - \frac{\bar{X^M})(X_i^E - \bar{X^E})}{\sqrt{\sum_{i=1}^{n}(X_i^M - \bar{X^M})^2 \sum_{i=1}^{n}(X_i^E - \bar{X^E})^2}}$ |
| Ψ | Root mean square difference | $\sqrt{\frac{1}{N}\sum_{i=1}^{n}(X_i^E - X_i^M)^2}$ |
| $\delta$ | Bias | $\frac{1}{N}\sum_{i=1}^{n}(X_i^E - X_i^M)$ |
| Δ | Bias-corrected root mean square difference | $\sqrt{\Psi^2 - \delta^2}$ |
| RPD | Relative percentage difference | $100\frac{1}{N}\sum_{i=1}^{n}\frac{\lvert X_i^E - X_i^M\rvert}{X_i^M}$ |

[a] $\sigma_{EE} = \frac{1}{N-1}\sum_{i=1}^{n}(X_i^E - \bar{X^E})^2, \sigma_{MM} = \frac{1}{N-1}\sum_{i=1}^{n}(X_i^M - \bar{X^M})^2$

$\sigma_{ME} = \frac{1}{N-1}\sum_{i=1}^{n}(X_i^E - \bar{X^E})(X_i^M - \bar{X^M})$

400–560 nm $L_{WN}$ over the 3 × 3 box was higher than 0.15. As with the uncertainty threshold, the CV threshold was varied to test the effect of spatial homogeneity on the performance metrics. Error bars for the *in situ* and satellite $L_{WN}$ were computed, respectively, from the *in situ* uncertainty budget per band and from the standard deviation over the 3 × 3 pixel extraction. Matchups were also excluded if less than 50% of the pixels were available within the 3 × 3 box.

- To assess the performance of the satellite retrievals, a set of widely-used statistical metrics (Brewin et al., 2015; Müller et al., 2015) were employed: the type-II slope (S) and intercept (I), the Pearson correlation coefficient (r), the root-mean-square difference (Ψ), the bias ($\delta$), the bias-corrected root-mean-square difference (Δ), and the relative percentage difference (RPD), as defined in Table 2. The notation follows the ocean-colour literature, with $X^M$ being the measured (*in situ*) variable and $X^E$ the estimated (satellite) variable. All statistical tests were performed in linear space. Note the use of the type-II slope and intercept, which were selected because both the *in situ* and satellite measurements are affected by uncertainties (Glover et al., 2011). The relative percentage difference, while conveying similar information to the bias, is included here as one of the metrics of choice in the ESA Mission Requirements Document (Donlon, 2011).
- The OLCI processors were compared in coincident and independent matchup analysis. In the coincident analysis, all processors had to provide a valid retrieval (according to the criteria above) for a matchup to be considered. In the independent analysis, all the valid matchups from each processor were considered to compute the statistics.
- Finally, we used the round-robin classification developed by Brewin et al. (2015) to assess and rank the quantitative performance of the different sensors and their corresponding ACs based on the independent analysis statistics. For each sensor AC processor, each band was tested independently and assigned a score. The final score for a given processor was calculated as the mean of its band scores, normalised by the average score of all processors being tested. A score greater than one indicates that the processors performed better than the average, while a score less than one indicates that the processors performed worse than the average. If a processor has a score of one, its performance was considered average when compared to all other processors. The round-robin exercise was split in two periods, AMT26–27 and AMT28–29, so the shortest data streams (NOAA-20 VIIRS and S-3B OLCI) and the longest data streams (MODIS-Aqua, Suomi-VIIRS and S-A OLCI) could be evaluated on equal terms regarding the number of retrievals.

## 3. Results

### 3.1. Impact of the uncertainty and homogeneity thresholds

Uncertainties in the radiometric measurements are routinely estimated at fixed platforms used for satellite validation, and an uncertainty threshold of 5% to 20% is often used to select the highest quality *in situ* data at fixed platforms (Gergely and Zibordi, 2014; Brown et al., 2007; Antoine et al., 2008; Białek et al., 2020). This threshold has not been assessed for moving platforms such as ships. To evaluate the impact of *in situ* uncertainty on our analysis, we tested different uncertainty thresholds, 100%, 75%, 50%, 25%, 10%, and 5%, while keeping the matchup temporal window (± 1 h) and coefficient of variation threshold (0.15) constant.

The metrics of Table 2 were calculated for each of these uncertainty thresholds for all the OLCI bands under assessment for the IPF-OL-2 product. Fig. 2a shows how the root-mean-square difference (Ψ) varies with the uncertainty threshold for the different bands. The minimum uncertainty threshold considered for this analysis was 5% due the number of matches being negligible for lower thresholds. Although Ψ decreased with the wavelength, the shape of the curve as a function of uncertainty was the same for all bands, with the minimum statistically-significant value of Ψ obtained at 25% uncertainty threshold for all bands. Results for band 442 nm summarised in Table S1 show that the rest of the metrics exhibited a similar behaviour. The number of matchups, the root mean square difference and the relative percentage difference increased, and correlation coefficient values decreased with the threshold. The bias-corrected root mean square difference remained stable across the different uncertainty ranges. Variation in all statistics between 5% to 25% uncertainty was small, except for an 85% increase in the number of matchups. We therefore selected 25% as the optimal uncertainty threshold for our analysis. It is critical to note that, while the higher uncertainties in the yellow and red parts of the spectrum drive the average total uncertainty, the uncertainty values in the 412–510 nm range are much lower (see Fig. 1 *versus* Lin et al. (2022) Fig. 8). In particular, for matchups with average uncertainty below the 25% threshold, the median uncertainty of water leaving radiances is ∼6% for the 412 nm, 442 nm and 490 nm bands, ∼8% at 510 nm, ∼13% at 550 nm, ∼61% at 620 nm, and ∼67% at 665 nm.

To test the effect of spatial and temporal variability on our analysis, we assessed how an increase or decrease in the time window and coefficient of variation threshold changed the statistics for the OLCI IPF-OL-2 product. Table S2 shows the results at 442 nm for each time window, while Table S3 lists the resulting metrics for the same band at different CV thresholds. The number of matchups decreased with a stricter time window, with numbers dropping by 52% from ±12 to ±3 h and 96% from ±12 to ±0.25 h. The bias-corrected root mean square difference and regression statistics remained relatively stable, while the relative differences improved slightly for the longer time windows. Based on metrics improvement *versus* number of matchups, we selected ± 1 h as the optimal time window for our matchup analysis. The spatial homogeneity test showed very small changes in all the metrics for the different CV values, with matchup numbers only dropping by 4% when we decreased the CV threshold from 1 to 0.1. Similar to the uncertainty threshold, a CV of 0.15 ensures that the RPD is low, whilst maintaining a high number of matchups. Following these results and for consistency with the existing literature, we selected a CV threshold of 0.15.

### 3.2. Comparison of OLCI processors

The satellite water-leaving radiances retrieved by the IPF-OL-2, POLYMER, and l2gen OLCI processors were compared with the Hyper-

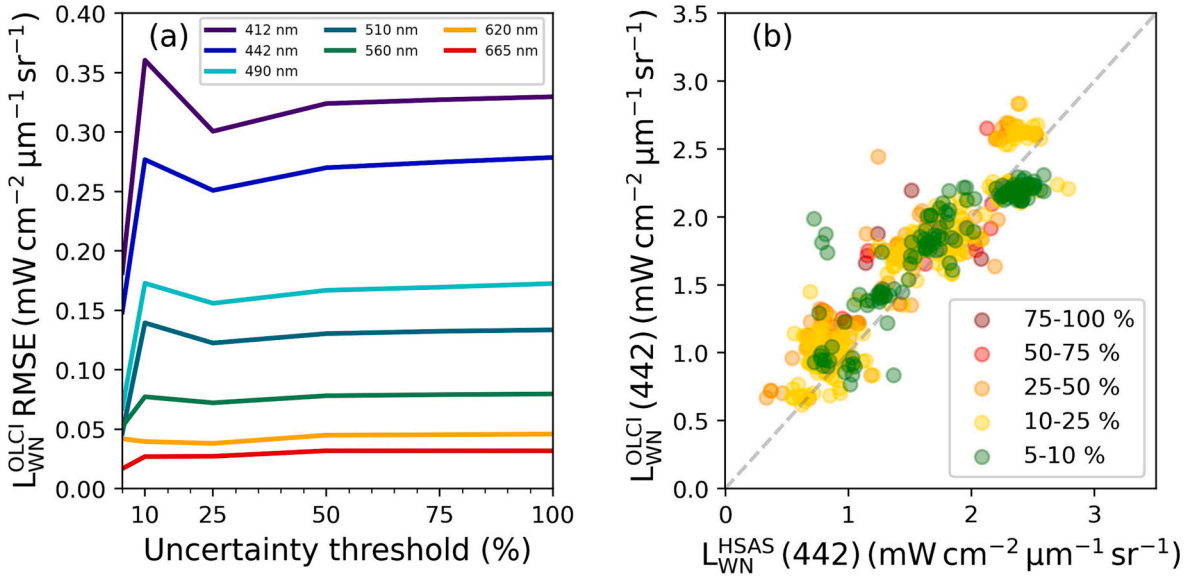



**Fig. 2.** (a) Variation of Ψ with the uncertainty threshold for each of the S-3A OLCI IPF-OL-2 bands, (b) Water-leaving radiance matchups at 442 nm under different uncertainty thresholds for S-3A OLCI IPF-OL-2.

SAS *in situ* dataset acquired during the four AMT campaigns using a ±1 hour window, CV threshold of 0.15 for satellite data, and average uncertainty threshold of 25% for *in situ* data (6% for *in situ* data in the blue-green spectral region). This resulted in 38 S-3A and 20 S-3B passes available for matchups, with an average of 13 matches per pass. The minimum and maximum number of matches per pass are 1 and 36 respectively.

A total of 253 and 101 matchups were obtained in coincident matchup analysis for S-3A and S-3B respectively, with $N_{3A}$=79 for AMT26, $N_{3A}$=119 for AMT27, $N_{3A}$=33, $N_{3B}$=51 for AMT28 and $N_{3A}$=22, $N_{3B}$=50 for AMT29. The typical distance between 2 *in situ* records considered independently for validation analysis was ∼ 500 m. The minimal separation between same-image matches is given by the nominal sensor resolution. Results for the independent matchup analysis per processor, including number of matchups, are provided in the supplementary data in Figs. S4 to S10.

Figs. S2a and S3a show the $L_{WN}$ spectra derived from the HyperSAS instrument concurrent with satellite matchups for the S-3A OLCI and S-3B OLCI eras. As expected for low Chl-a waters, the dominant signals were at blue bands, with peaks at either 442 or 490 nm. Water-leaving radiance values then decreased gradually with wavelength to converge to a set of flat curves with radiance close to zero from 580 nm onward. Values were consistent with those reported for the region in previous cruises, in particular values in the blue-green part of the spectrum were below 3 mW cm$^{-2}$ µm$^{-1}$ sr$^{-1}$ for most of the spectra. Figs. S2b-g and S3b-g show the $L_{WN}$ spectra observed by the different sensors and processors for the S-3A OLCI and S-3B OLCI eras. POLYMER exhibited a similar spectra to the *in situ* measurements, while IPF-OL-2 and l2gen

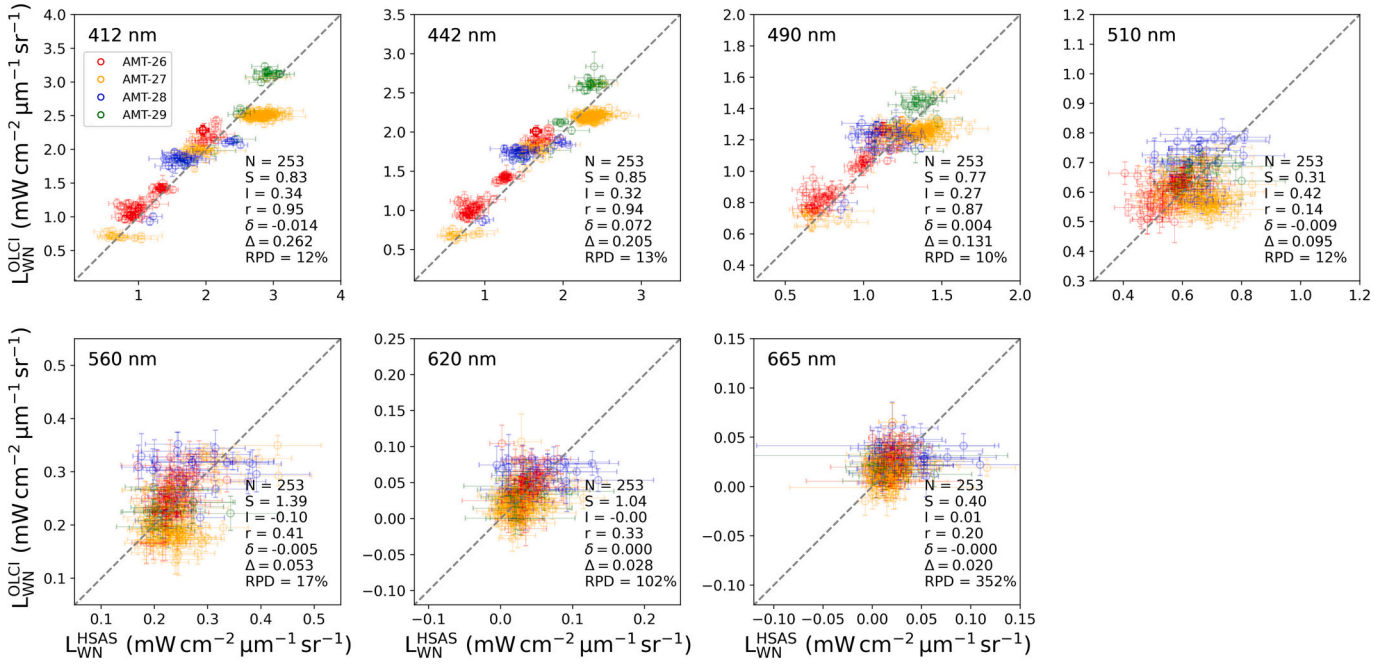



**Fig. 3.** Scatter plots of the S-3A OLCI IPF-OL-2 $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line.

were notably different in the green and red parts of spectrum. POLYMER and l2gen showed a wider range of values than IPF-OL-2 at 412 nm. As with the *in situ* data, the values obtained were $L_{WN}$ < 3 mW cm$^{-2}$ $\mu$m$^{-1}$ sr$^{-1}$, showing that the data were not affected by sun glint, whitecaps, or spray. Differences in the spectral shape obtained for Suomi-VIIRS (Figs. S2e, S3e) and NOAA-20 VIIRS (Figs. S2f, S3f) were partially the result of their reduced band set, although mean $L_{WN}$ values were generally lower than the *in situ* equivalent for all bands, particularly in the green. The spectra observed for MODIS (Figs. S2g, S3g) showed a flatter spectral shape in the blue part of the spectrum, with more than half of the retrievals presenting values of $L_{WN}$ < 1 mW cm$^{-2}$ $\mu$m$^{-1}$ sr$^{-1}$ between 412 nm and 531 nm.

Scatter plots of *in situ* $L_{WN}$ *versus* OLCI AC processors coincident matchups for each band are given in Figs. 3, 4 and 5 for S-3A OLCI, and Figs. 6, 7 and 8 for S-3B OLCI. Each of the three processors showed good accuracy in the retrieval of the 412 nm band, with correlation coefficients ranging from 0.84 to 0.95. Among them, OLCI IPF-OL-2 showed the highest correlation and lowest bias, root mean square difference, and RPD for S-3A OLCI. For S-3B OLCI, the performances of OLCI IPF-OL-2 and POLYMER were nearly identical, except for the intercept values. The type-II regression slopes ranged from 0.83 to 0.99, with the POLYMER and l2gen products performing closest to a slope of 1 for S-3A OLCI and S-3B OLCI respectively. Results show low bias and bias-corrected root mean square differences falling between 0.25 and 0.45. RPD values for IPF-OL-2 and POLYMER S-3A OLCI products were in the range of 12% to 14%, increasing to 16% to 17% for l2gen S-3A OLCI and all S-3B OLCI processors. Overall, IPF-OL-2 S-3A OLCI demonstrated the best performance at 412 nm, achieving an RPD of 12%.

All three processors performed well at 442 nm, with correlation coefficients varying from 0.81 to 0.95, low bias and bias-corrected root mean square difference between 0.2 and 0.4 mW cm$^{-2}$ $\mu$m$^{-1}$ sr$^{-1}$. POLYMER provided the highest correlation and lowest bias, root mean square difference and RPD for this band for both S-3A OLCI and S-3B OLCI. The type-II regression slopes were between 0.85 and 1.04, and closest to 1 for the POLYMER product for S-3A OLCI and for the IPF-OL-2 product for S-3B OLCI. RPD values for the IPF-OL-2 and POLYMER S-3A OLCI products were between 11% and 13%, whereas they were up to 13% to 20% for S-3B OLCI; and around 23% for S-3A OLCI and S-3B OLCI l2gen due to a consistent overestimate compared to the *in situ* $L_{WN}$. Of the two sensors and three AC processors, POLYMER S-3A OLCI performed best overall at 442 nm with an RPD of 11%. Two data clusters were visible in the OLCI IPF-OL-2 results at 442 nm, particularly for S-3B OLCI: one at higher values from the blue waters of the Atlantic gyres, the other at lower values from greener tropical waters.

At 490 nm, the AC processors exhibited a similar trend to the blue bands. The POLYMER AC performed best for S-3A OLCI, having the lowest intercept, bias, root mean square difference and RPD of 9% and a slope closest to 1. Although the slope for the S-3B OLCI l2gen product was the closest to 1, the results were worse for S-3B OLCI for all ACs. This is because there was a persistent overestimation in $L_{WN}$, leading to a high 20% RPD.

At 510 and 560 nm the pattern changed due to the low signal in this part of the spectrum, resulting from the low Chl-a concentrations in the Atlantic waters sampled. For S-3A OLCI IPF-OL-2 the 560 nm band performed well, but the 510 nm band exhibited a large scatter around the 1:1 resulting in a high intercept, and low slope and correlation coefficient. For POLYMER, there was a large offset caused by the clusters of points below the 1:1 line. Consequently there was a decrease in the slope and correlation coefficients, and the intercept and RPD were higher than for the blue bands. The l2gen product also showed a large scatter resulting in a high slope and RPD, and the lowest correlation coefficients of the three ACs for both bands. Overall, POLYMER and IPF-OL-2 performed best at 510 and 560 nm respectively, although improvement is needed for the 510 nm band for all processors.

All three AC processors were poor at 620 nm and 665 nm as a result of oligotrophic waters yielding very low signals in the red part of the spectrum. The resulting scatter plots show large scatter and RPD, the slopes were high and the correlation coefficients were low. The POLYMER processor exhibited the best performance at 620 nm in terms of RPD but the satellite retrievals failed to match the variability in the *in situ* data, as evidenced by an almost flat regression fit. The same applies to the 665 nm, which showed RPD values well above 100%.

The number of matchups for each processor varied considerably when assessed individually (Figs. S4 to S9 supplementary material): $N_{3A}$=353, $N_{3B}$=269 for IPF-OL-2, $N_{3A}$=490, $N_{3B}$=238 for POLYMER and

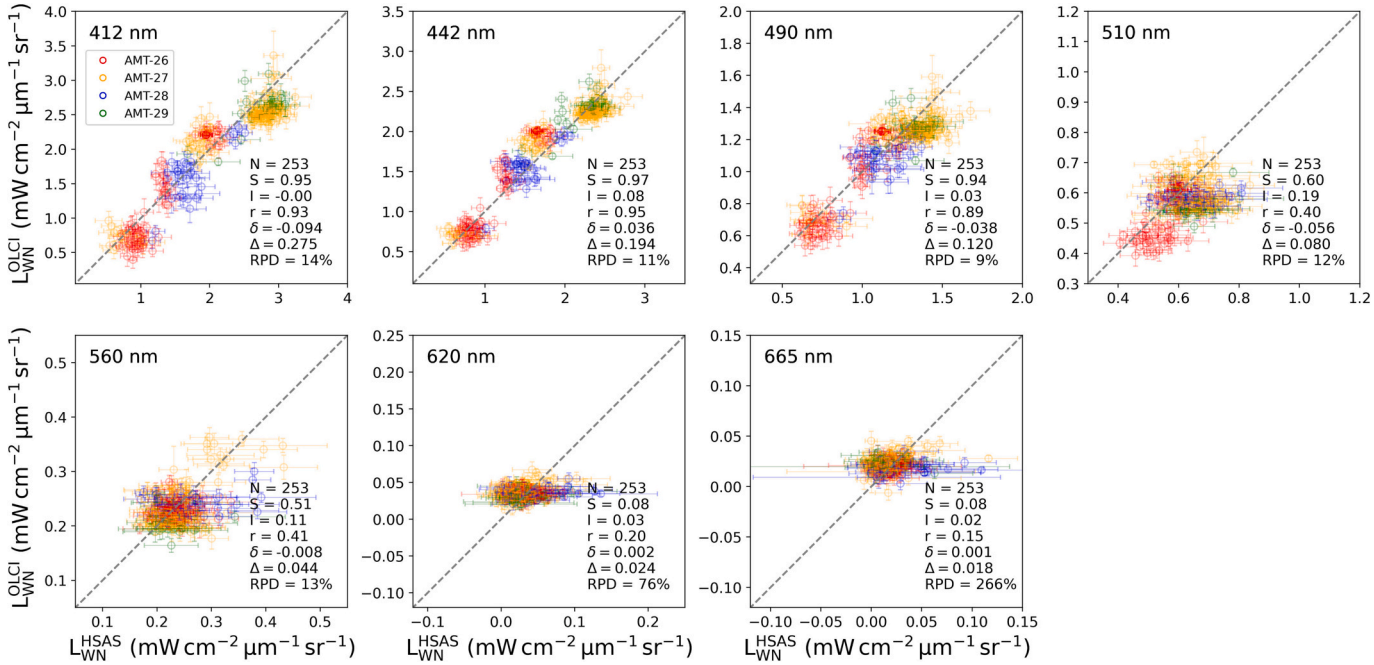



**Fig. 4.** Scatter plots of the S-3A OLCI POLYMER product $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line.

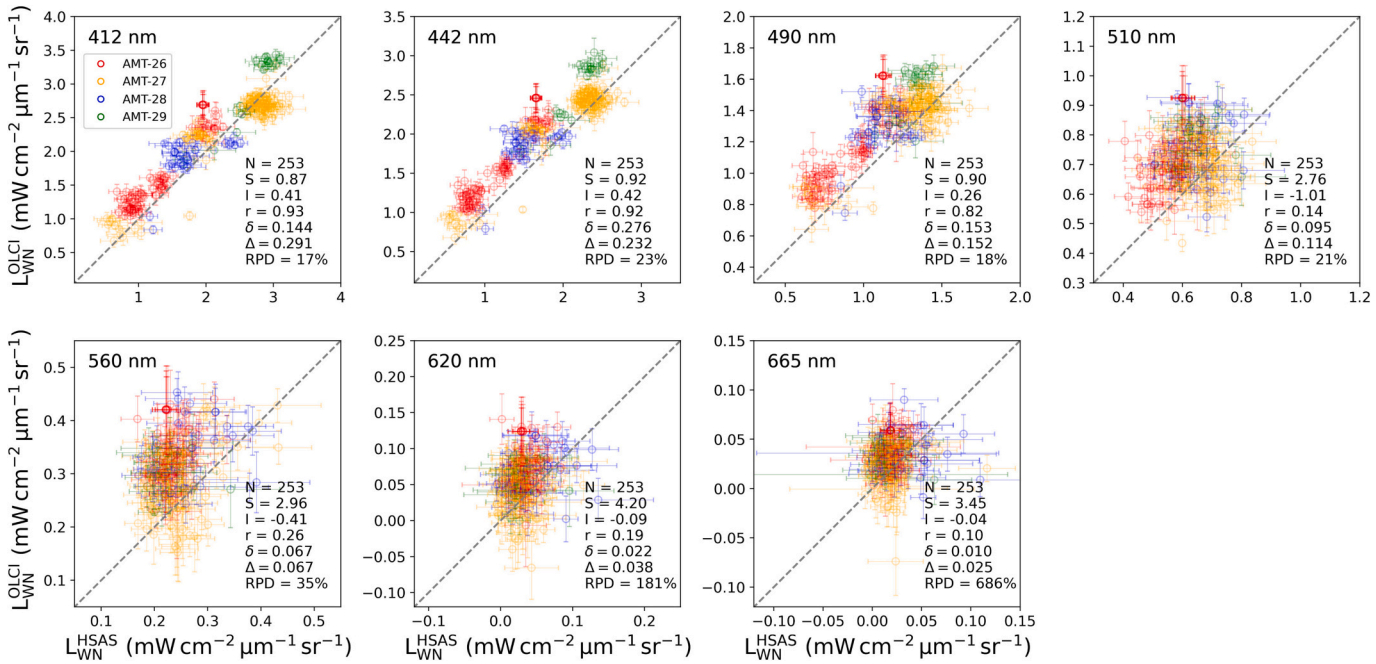



**Fig. 5.** Scatter plots of the S-3A OLCI l2gen product $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line.

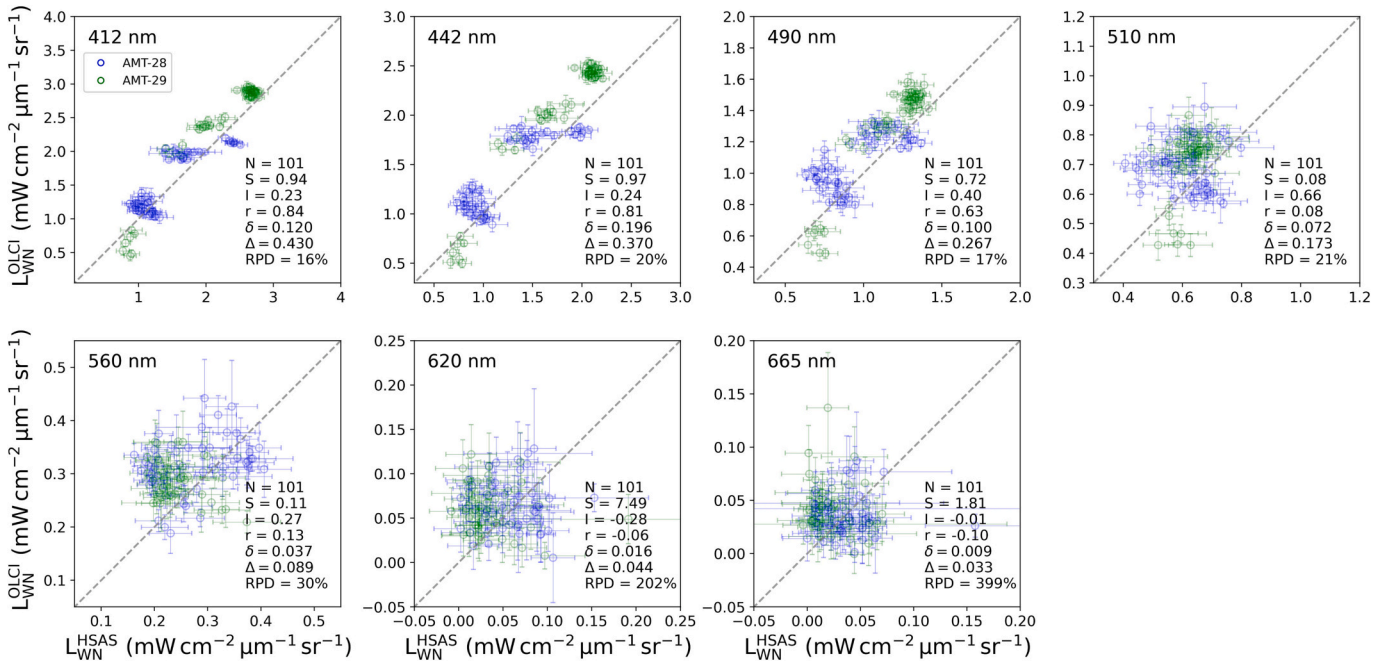



**Fig. 6.** Scatter plots of S-3B OLCI IPF-OL-2 $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line.

$N_{3A}$=453, $N_{3B}$=305 for l2gen. POLYMER performed best at 442–490 nm and, due to the very low radiometric signal, performance was poor at 560–620 nm for all processors. Results from the round-robin ranking (Table 3) show that S-3A OLCI POLYMER performed best overall, with the highest and second highest normalised scores of the two exercises (1.28 and 1.09) and the highest total normalised score (1.24). Scores for the other two S-3A OLCI processors, l2gen and IPF-OL-2, suggest that their performance is comparable with the average for the nine sensors and AC combinations under study, while the performance of the three S-3B OLCI products was slightly worse.

### 3.3. Comparison of MODIS, OLCI and VIIRS

The different overpass times, band centres and spatial resolutions of S-3A/B OLCI, MODIS-Aqua, Suomi-VIIRS and NOAA-20 VIIRS meant that a point-to-point coincident matchup comparison was not possible due to the small number of common matchups. To facilitate comparisons between the three ocean-colour sensors, the time window for the MODIS-Aqua analysis was increased from ±1 h to ±3 h, whilst the uncertainty and CV thresholds were kept the same as for OLCI and VIIRS. This resulted in 40 MODIS-Aqua passes, 32 Suomi-VIIRS passes and 18 NOAA-20 VIIRS passes available for matchups.

MODIS-Aqua (Fig. 9) and NOAA-20 VIIRS (Fig. 11) exhibit a similar

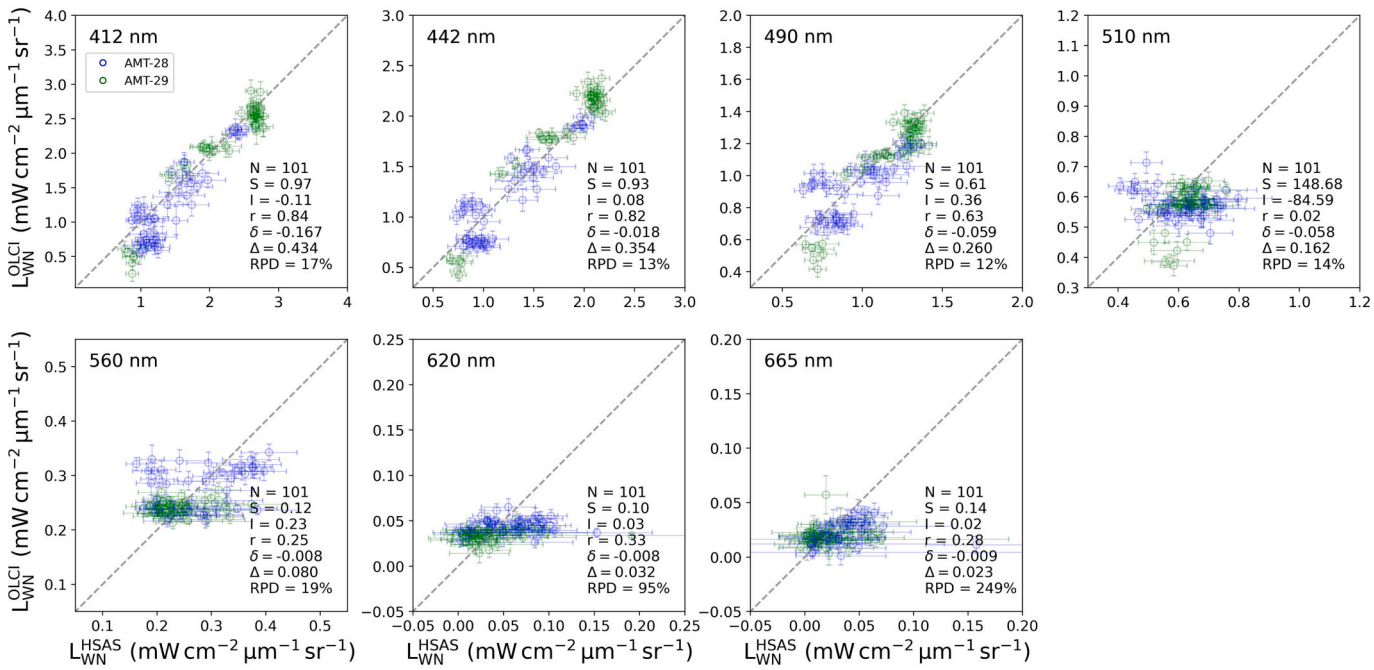



**Fig. 7.** Scatter plots of the S-3B OLCI POLYMER product $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line.

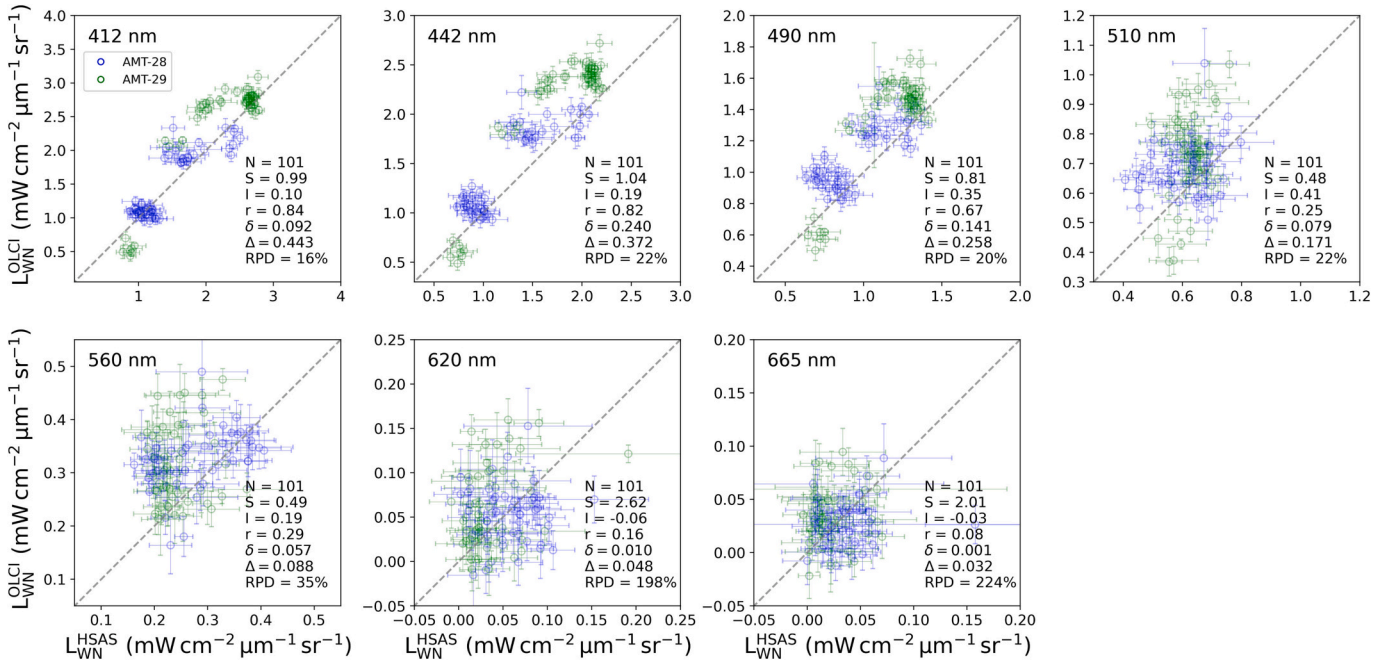



**Fig. 8.** Scatter plots of the S-3B OLCI l2gen product $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line.

behaviour to the S-3B OLCI IPF-OL-2 product, with an overestimate at high $L_{WN}$ and RPD in the 13% to 22% range. In particular for MODIS-Aqua, there was an overestimate of some reflectance values from pixels affected by the 'SEAICE' SeaDAS l2gen quality flag, resulting in a high slope, RPD and root mean square difference (Fig. 9). On the other hand, results for Suomi-VIIRS (Fig. 10) suggest a similar trend to the S-3A OLCI IPF-OL-2, with the overestimates at low $L_{WN}$ driving the high slope and RPD values.

Of the NASA sensors, MODIS-Aqua performed worst for AMT26–27 and best for AMT28–29, with scores of 0.77 and 1.24 respectively (Table 3). NOAA-20 VIIRS performed below average (score of 0.88) and comparably to S-3B OLCI l2gen. Suomi-VIIRS results, with scores of 0.91 and 0.92, suggest an average performance comparable to S-3A OLCI l2gen.

### 3.4. Sentinel-3A and 3B OLCI tandem phase comparison

For AMT28, there were 14 coincident passes within ±12 h from *in situ* acquisition, resulting in a total of $N = 150$ coincident matchups. Results for the comparison with AMT28 are summarized in Fig. 12. At bands from 412 to 510 nm, S-3A OLCI IPF-OL-2 was more accurate than S-3B OLCI IPF-OL-2, which tended to overestimate at low $L_{WN}$ values. At

**Table 3**
Results from the round-robin processor comparison.

| Processor | AMT26–27 | AMT26–27 | AMT26–29 |
|---|---|---|---|
| S-3A OLCI IPF-OL-2 | 1.12 | 0.78 | 1.02 |
| S-3A OLCI POLYMER | 1.28 | 1.09 | 1.24 |
| S-3A OLCI l2gen | 0.91 | 0.91 | 1.07 |
| S-3B OLCI IPF-OL-2 | – | 0.99 | 0.95 |
| S-3B OLCI POLYMER | – | 0.96 | 0.93 |
| S-3B OLCI l2gen | – | 1.06 | 0.88 |
| Suomi-VIIRS | 0.92 | 0.91 | 0.95 |
| NOAA-20 VIIRS | – | 1.06 | 0.88 |
| MODIS-Aqua | 0.77 | 1.24 | 1.07 |

560 nm we saw a slight switch in this trend with more S-3A OLCI points now below the 1:1, resulting in a higher RPD compared to S-3B OLCI. At red bands, due to the very low $L_{WN}$ signal, both S-3A OLCI and S-3B OLCI showed a high degree of scatter which resulted in low correlation values and high RPD. These results are consistent with the AMT28 trends shown in Figs. S3 and S7 of the supplementary material. AMT28 results for the POLYMER and l2gen processors are included in the supplementary material (Figs. S10 and S11).

To further investigate these results, the 150 S-3A and 3B IPF-OL-2 coincident AMT28 matchups were split in two phases: the Tandem Phase (from the beginning of the AMT28 campaign on 23 September to 16 October 2018), and the Drift Phase (from 17 October to the end of the AMT28 campaign on 29 October 2018). S-3A OLCI and S-3B OLCI IPF-OL-2 $L_{WN}$ were directly compared for both phases with $N = 107$ for the Tandem Phase and $N = 43$ for the Drift Phase. During the Tandem Phase, IPF-OL-2 showed a good agreement at all bands from 412 to 620 nm, with a slope and correlation coefficient close to 1, low bias, intercept, and mean square difference and an RPD value of 8% (Fig. 13 left, Table 4). In contrast, Drift Phase results (Fig. 13 right, Table 5) show a significant decrease in agreement. This is evidenced by increased values for bias, intercept, and root mean square difference, resulting in an RPD value of 22%.

At the band level, Table 4 shows that RPD values for the Tandem Phase were within 2 to 7% between 412 and 510 nm, with slopes within 0.96 to 1.17 and correlation coefficients above 0.5 for the same wavelength range. At 560 nm the RPD increased to 20%. During the Drift Phase (Table 5) S-3B OLCI showed positive biases when compared to S-3A OLCI with RPD values ranging from 15% to 22% between 412 and 560 nm. At 620 nm and 665 nm, results were poor for both phases due to low signal, but S-3A OLCI/S-3B OLCI agreement was better during the Tandem Phase in terms of regression and correlation.

Tandem and Drift Phase results for the POLYMER and l2gen processors are provided in the supplementary material. For POLYMER, the agreement between sensors during the Tandem Phase (Fig. S12 left panel, Table S4) was comparable to IPF-OL-2, although with reduced scatter which results in a lower RPD of 5%. During the Drift Phase, POLYMER S-3A and S-3B OLCI had a smaller RPD of 10% and bias and higher correlation compared to IPF-OL-2 S-3A and S-3B OLCI (Fig. S12 right panel, Table S5). For l2gen, the agreement between the sensors during the Tandem Phase was comparable to IPF-OL-2, although the metrics for the 412–665 nm dataset are slightly worse (Fig. S13 left panel, Table S6). The results for the l2gen Drift Phase (Fig. S13 right panel, Table S7) are not conclusive due to the smaller number of matchups.

## 4. Discussion

### 4.1. Uncertainties in fiducial reference measurements of above-water radiometry and other validation criteria

Fixed-platform sites, such as MOBY (Clark et al., 2003), BOUSSOLE (Antoine et al., 2008) and AERONET-OC (Zibordi et al., 2009b), have traditionally been used to acquire very high quality *in situ* data for vicarious calibration as well as validation. While Bailey et al. (2008) showed that some of the conditions (requirements for hyper-spectral resolution and low Chl-a waters, restrictions on aerosol type and load) can be relaxed without compromising the quality of the vicarious calibration, these platforms are usually bound by stringent requirements (Zibordi et al., 2015). Consequently, the uncertainty at these sites has been well characterised. For AERONET-OC, the estimated uncertainty in normalised water-leaving radiance varies from 5% in the blue to 7–10% in the red (Gergely and Zibordi, 2014). MOBY is even more stringent and varies from 3% in the blue to 5% in the red (Brown et al., 2007). For

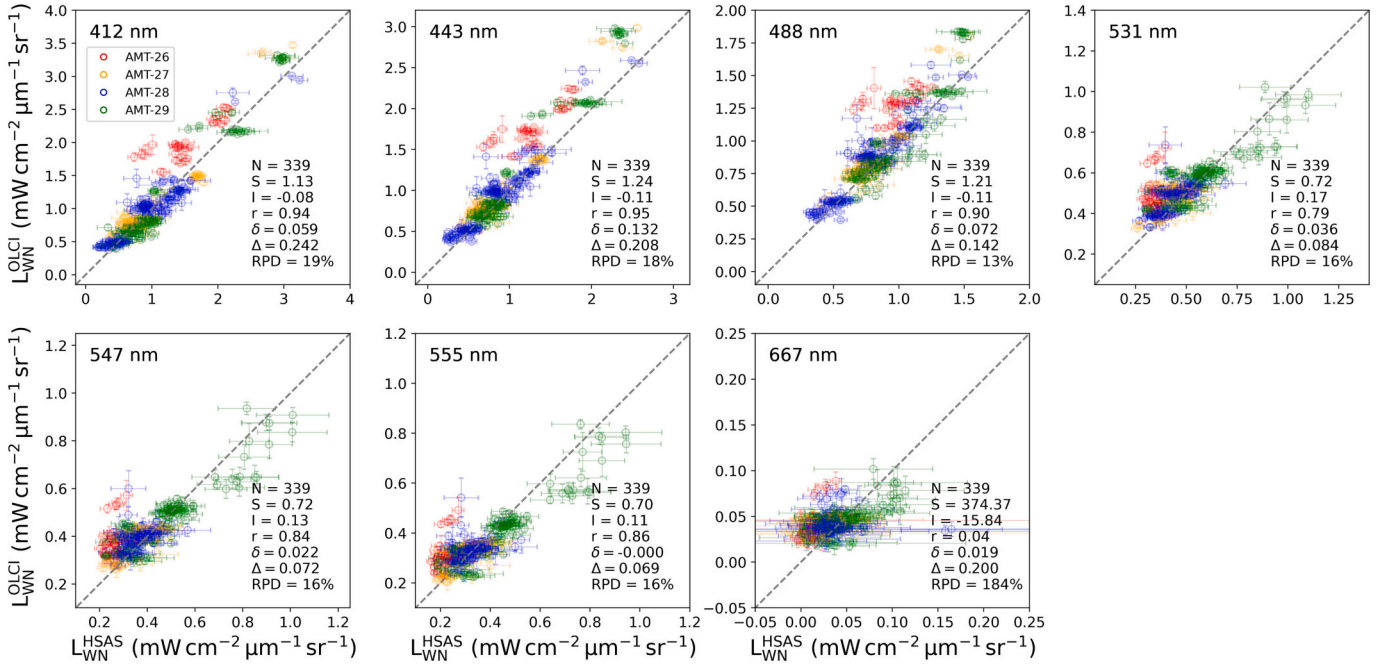



**Fig. 9.** Scatter plots of the MODIS-Aqua $L_{WN}$ product *versus* HyperSAS *in situ* $L_{WN}$ for the common bands in the visible. The dashed line corresponds to the 1:1 line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

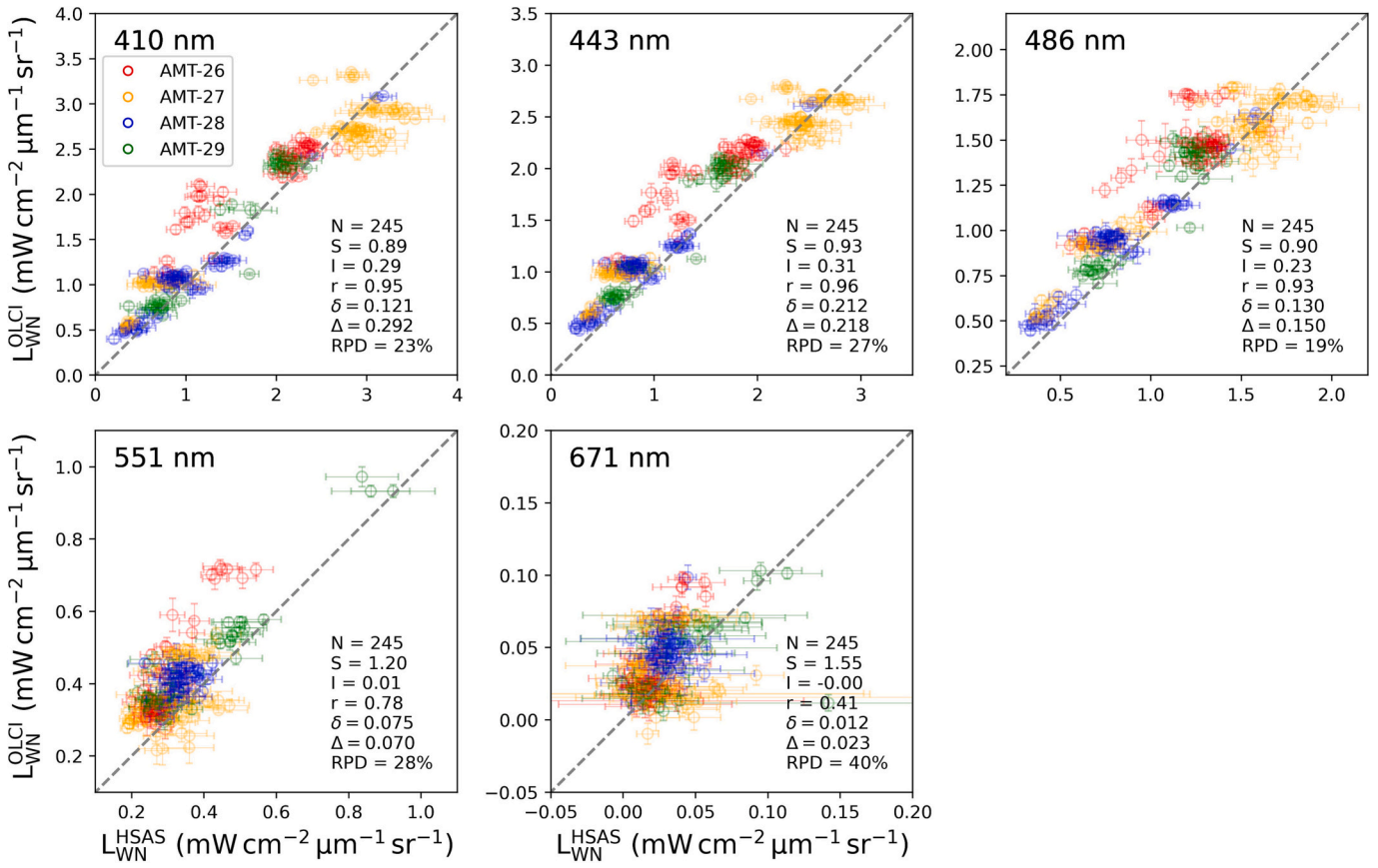



**Fig. 10.** Scatter plots of the Suomi-NPP VIIRS product $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the common bands in the visible. The dashed line corresponds to the 1:1 line.

BOUSSOLE the overall uncertainty in water-leaving reflectance is approximately 6% (Antoine et al., 2008), with a more recent Monte-Carlo assessment showing values below 4% for remote sensing reflectances in the blue and green bands, and approximately 5% in the red channels (Białek et al., 2020). Studies using these data for validation exercises usually assume that the *in situ* dataset has already been acquired under optimal conditions and undergone quality assurance, so no filtering based on an uncertainty threshold or any other quality indicator is performed.

Shipborne platforms have the potential to be used for vicarious calibration (Bailey et al., 2008). Protocols have been proposed to minimise the effect of the ship and environmental conditions (Hooker and Morel, 2003) but, in practice, radiometric data from such sources have been mostly used for radiometric validation (Brewin et al., 2016; Tilstone et al., 2017; Giannini et al., 2021). Uncertainties arising from above-water radiometry on moving platforms are less well characterised than for fixed sites, most likely due to difficulties in defining an uncertainty budget (Zibordi et al., 2012), and are typically checked through intercomparison exercises between radiometers (Vabson et al., 2019; Tilstone et al., 2020; Alikas et al., 2020). Most studies rely on well-controlled acquisition conditions to obtain high-quality, low-uncertainty measurements on ships, with the ship maneuvered into station and sensors pointing at optimal zenith angles (Ruddick et al., 2006). For other sources, such as ships of opportunity, quality control has typically been implemented through post-acquisition filtering based on spectral shapes and typical maximum values (Qin et al., 2017; Tilstone et al., 2022).

While Alikas et al. (2020) had previously computed the uncertainty associated with $R_{rs}$ measured on a moving ship for two radiometer systems (TriOS-RAMSES and Seabird HyperSAS) deployed on the AMT, all the *in situ* data were used to validate S-3A OLCI regardless of uncertainty. To the best of our knowledge, uncertainty estimates have never been used to select *in situ* data measured from a moving ship for validation purposes. Here we have demonstrated that, while our uncertainty threshold is higher than the nominal uncertainties of most fixed platforms, lowering it has no appreciable effect on the validation metrics for the region. In fact, our proposed 25% uncertainty threshold provides the greatest advantage in terms of both data quantity and quality by minimising the value of Ψ while maximising the number of matchups. Most of the fixed platforms that provide high-quality FRM for satellite validation are in coastal and shelf environments which cover eutrophic to mesotrophic ecosystems. Quantifying uncertainties in radiometric data from the AMT highlights the value of using semi-autonomous underway measurements with low uncertainties to enhance the matchup capability in remote oligotrophic regions of the ocean.

Some regions transected by the AMT campaigns (*e.g.*, the subtropical gyres) are known to be relatively homogeneous with low temporal and spatial variability (Brewin et al., 2016). The homogeneity threshold study in Section 3.1 corroborated this by showing that changes in the coefficient of variation threshold had very little impact in the validation statistics. The comparison of different validation methods performed by Concha et al. (2021) showed that the Bailey and Werdell (2006) protocol produced 20% more S-3A OLCI matchups than the Zibordi et al. (2009a) protocol. This indicates that changes in the size of the extraction box and the threshold number of valid pixels within the box (50% valid pixels within a 5 × 5 pixel box for Bailey and Werdell (2006), 100% valid pixels within a 3 × 3 pixel box for Zibordi et al. (2009b) would have a bigger impact than any changes in the CV threshold for our AMT data. Regarding the effect of changing the matchup temporal window, we observed a small improvement in relative differences when increasing the temporal window, with the other statistical metrics remaining

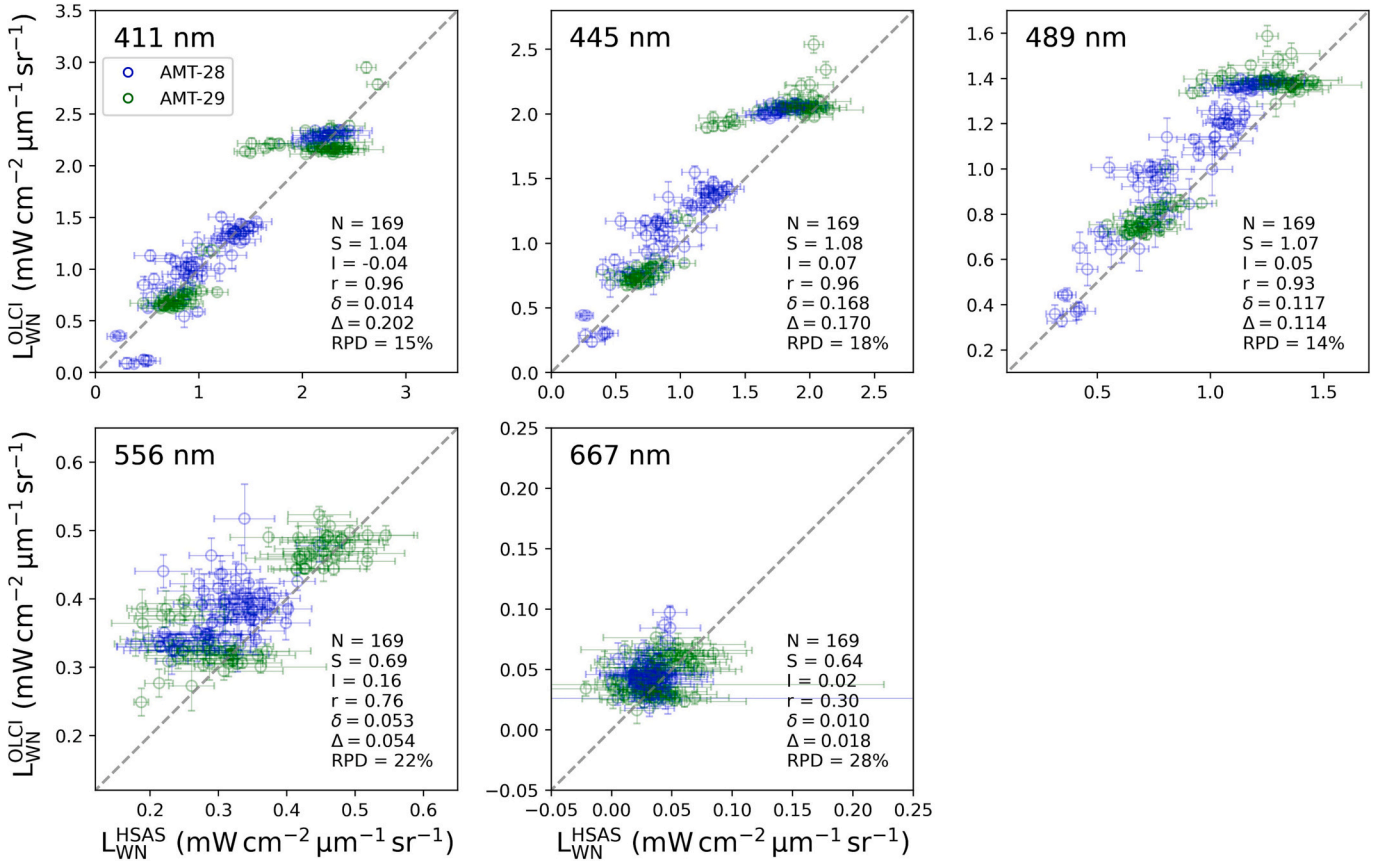



**Fig. 11.** Scatter plots of the NOAA-20 VIIRS product $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the common bands in the visible. The dashed line corresponds to the 1:1 line.

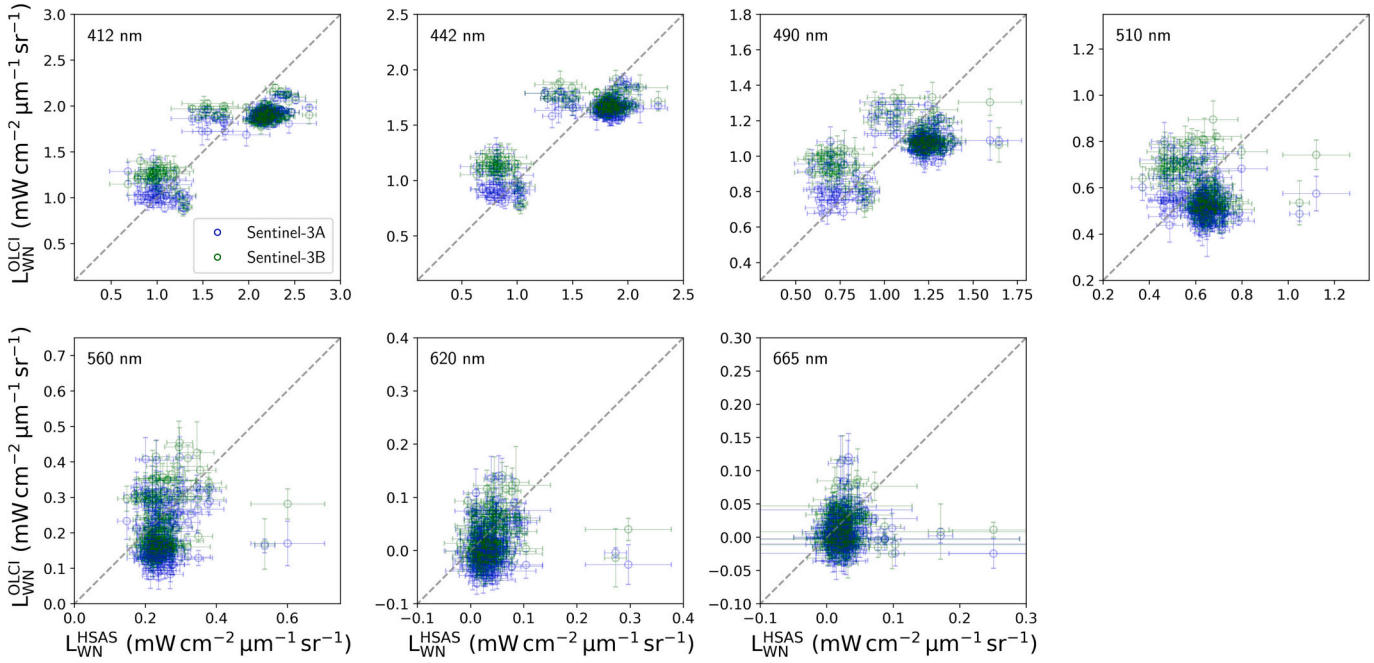



**Fig. 12.** Coincident AMT28 matchup results for S-3A OLCI (blue) and S-3B OLCI (green) IPF-OL-2 $L_{WN}$ *versus* HyperSAS *in situ* $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

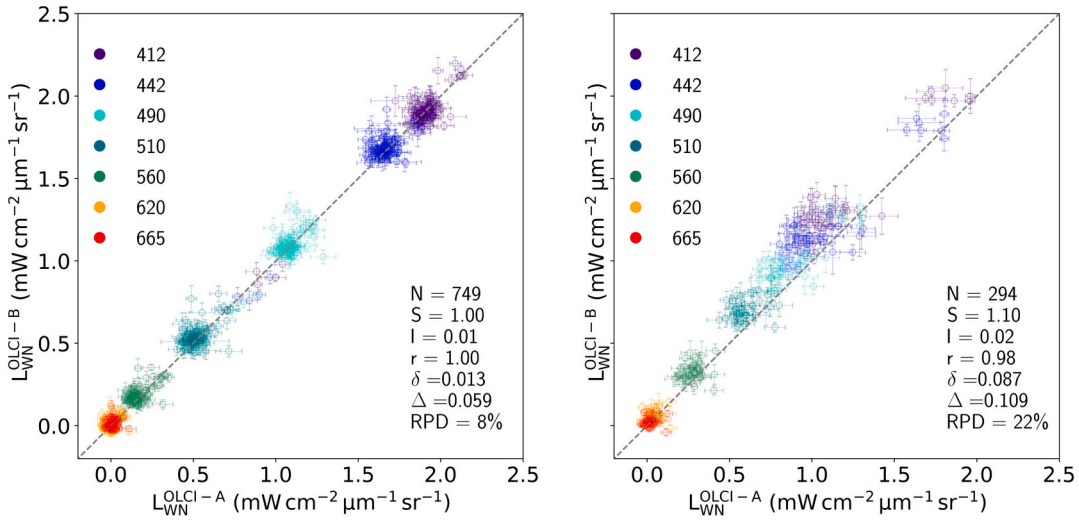



**Fig. 13.** Coincident Tandem Phase (left) and Drift Phase (right) matchup results for S-3B OLCI IPF-OL-2 $L_{WN}$ *versus* S-3A OLCI $L_{WN}$ for the OLCI bands in the visible. The dashed line corresponds to the 1:1 line.

**Table 4**
S-3B *versus* S-3A OLCI IPF-OL-2 tandem phase statistics.

| Band | N | S | I | r | Ψ | δ | Δ | RPD |
|---|---|---|---|---|---|---|---|---|
| 412 | 107 | 1.05 | −0.09 | 0.94 | 0.068 | 0.001 | 0.068 | 2% |
| 442 | 107 | 1.07 | −0.09 | 0.91 | 0.077 | 0.018 | 0.074 | 2% |
| 490 | 107 | 1.17 | −0.17 | 0.73 | 0.065 | 0.011 | 0.065 | 3% |
| 510 | 107 | 0.96 | 0.04 | 0.51 | 0.070 | 0.023 | 0.066 | 7% |
| 560 | 107 | 0.75 | 0.06 | 0.58 | 0.053 | 0.020 | 0.050 | 20% |
| 620 | 107 | 0.94 | 0.01 | 0.35 | 0.042 | 0.014 | 0.039 | 110% |
| 665 | 107 | 1.11 | 0.00 | 0.20 | 0.031 | 0.004 | 0.031 | 127% |

**Table 5**
S-3B *versus* S-3A OLCI IPF-OL-2 drift phase statistics.

| Band | N | S | I | r | Ψ | δ | Δ | RPD |
|---|---|---|---|---|---|---|---|---|
| 412 | 42 | 0.97 | 0.21 | 0.91 | 0.216 | 0.172 | 0.130 | 16% |
| 412 | 42 | 0.90 | 0.26 | 0.91 | 0.192 | 0.148 | 0.122 | 15% |
| 490 | 42 | 0.78 | 0.30 | 0.81 | 0.147 | 0.107 | 0.100 | 15% |
| 510 | 42 | 0.53 | 0.39 | 0.37 | 0.138 | 0.106 | 0.088 | 20% |
| 560 | 42 | 0.70 | 0.13 | 0.20 | 0.086 | 0.051 | 0.069 | 22% |
| 620 | 42 | 0.37 | 0.04 | 0.14 | 0.054 | 0.017 | 0.051 | 77% |
| 665 | 42 | 0.30 | 0.02 | 0.13 | 0.041 | 0.005 | 0.040 | 116% |

stable. These results are consistent with similar tests in other oligotrophic regions, see for example Table 2 in Brewin et al., 2013.

### 4.2. Radiometric performance of OLCI, MODIS and VIIRS

The design of our study follows a conventional validation approach, which solely yields a measure of the differences between *in situ* and satellite data. *A priori*, our findings are not able to identify the source of the divergence or suggest strategies to enhance the accuracy of the satellite outputs. According to Zibordi et al. (2022), discrepancies between satellite and *in situ* radiometric data may stem from errors in atmospheric correction, instrument characterization, and/or the quality of the *in situ* measurements. Instrumentation and experimental setup are among the various field measurement aspects that can introduce systematic variations between *in situ* and satellite retrievals, but their impacts can be mitigated through proper design. The previous section has addressed the role of uncertainties in this context. Furthermore, our *in situ* data meet the minimum spectral resolution requirement of 3 nm for validating OLCI in oligotrophic waters (Zibordi et al., 2017). In order to enhance the satellite atmospheric correction algorithm and gain an understanding of its different aspects, it is essential to develop an error budget as part of the validation process. Although the derivation of an error budget for OLCI radiometric products is beyond the scope of this paper, Li et al. (2019) conducted an exhaustive investigation which revealed that the primary sources of error in the China Sea were the aerosol estimation and the Rayleigh-correction.

Data obtained from AMT covered a wide range of trophic environments. While the shape of the *in situ* spectra (Figs. S2a and S3a) were consistent with those reported by Lee et al. (2015) in hyperoligotrophic waters in the South Pacific gyre, the signal for bands 412–510 nm was, in most cases, substantially lower. Spectra in the lower signal range matched radiometry reported for cruises across the Southern Atlantic and Southeastern Pacific oceans (Garcia et al., 2011; Rudorff et al., 2014). The highest variability in the magnitude of the signal was observed between 412 and 490 nm and can be attributed to the gradient in phytoplankton biomass concentration and composition, and associated co-varying material (*e.g.*, non-algal particles and yellow substances).

The GCOS requirement for the radiometric accuracy of S-3 OLCI is 5% for blue and green bands (GCOS, 2011). Using FRM's from AMT, we found that RPD values for the IPF-OL-2 processor varied from 10% at 490 nm to 17% at 560 nm for S-3A (Fig. 3) and from 16% at 412 nm to 30% at 560 nm for S-3B (Fig. 6). These values match relative differences reported by Zibordi et al. (2022) in the oligotrophic waters of the Casablanca Platform in the Mediterranean Sea. S-3A retrievals of $L_{WN}$(412) provided the best standard product results in terms of regression and correlation, although there was still some scatter around the 1:1 line. The higher RPD value at 412 nm and 442 nm could be due to inadequate identification and quantification of the aerosol concentration and composition in the IPF-OL-2 atmospheric correction, as suggested by Alikas et al. (2020). These results suggest that further work is still required on improving the IPF-OL-2 AC and the system vicarious calibration across bands, particularly for S-3B OLCI.

Using the POLYMER processor (Figs. 4, 7) resulted in improved retrievals for the blue and green bands, with RPD values from 9% at 490 nm to 14% at 412 nm for S-3A (Fig. 4) and from 12% at 490 nm to 19% at 560 nm for S-3B (Fig. 7). While the water types under study are different, the overall results match recent studies (Mograne et al., 2019; Alikas et al., 2020; Tilstone et al., 2022) where POLYMER outperformed C2R-CC in coincident analysis. While the POLYMER team have been proactive in their SVC efforts, the relative band-to-band calibration holds greater weight than the absolute gains for processors, such as POLYMER, that place stringent limitations on water reflectance as part

of their AC model (Steinmetz et al., 2011; Steinmetz and Ramon, 2018). This suggested that POLYMER's superior performance is due to better AC procedures, rather than differences in SVC. Although POLYMER accurately retrieves data contaminated by sun glint, thin clouds, heavy aerosol, plumes or adjacency effects, standard atmospheric correction schemes such as IPF-OL-2 and l2gen often fail to handle these cases correctly, as they rely on extrapolation from the near infrared. Improving the retrievals at blue and green bands is critical for increasing the accuracy of the OLCI case 1 chlorophyll product, as demonstrated by Tilstone et al. (2021) where S-3A OLCI Chl-a from POLYMER performed best in comparison to underway spectrophotometric Chl-a. Nonetheless, it was observed that POLYMER exhibited a tendency to restrict the dynamic range of $L_{WN}$ in the green bands, which could affect the precision of the band ratios and chlorophyll products derived from them. Alternatively, the introduction of a line-height approach, where Chl-a is associated with a band difference instead of a band ratio can provide more stability against these spectral biases (Cazzaniga and Kwiatkowska, 2018; Tilstone et al., 2021).

SeaDAS l2gen OLCI results exhibited a consistent positive bias in the blue and green bands for both S-3A OLCI and S-3B OLCI (Figs. 5, 8). Overall, this overestimation was more prominent than suggested in the validation of $R_{rs}$ against AERONET-OC reported by NASA OBPG (2022g), which was the only SeaDAS l2gen OLCI R2022 study available at the time of writing. These radiometric biases could be attributed to the use of a model-based vicarious calibration rather than an issue with the l2gen atmospheric correction, and would be likely to improve when *in situ*-derived system vicarious calibration gains become available for the l2gen processor.

The comparison of coincident S-3A OLCI and S-3B OLCI IPF-OL-2 products again showed that S-3A OLCI performed better than S-3B OLCI for our AMT28 data during the tandem and drift phases in 2018, with S-3B OLCI displaying a consistent overestimate for the blue bands at low $L_{WN}$ (Fig. 12). It is worth noting that 60% of our matchups for the tandem phase analysis came from the same pair of data granules acquired on 12th October 2018 in South Atlantic tropical waters, which presented active 'BPAC ON' and 'MEGLINT' flags for both S-3A OLCI and 3B. This suggests that differences may be due to the bright pixel and glint correction algorithms.

During the Tandem Phase, S-3A and S-3B OLCI data were acquired almost simultaneously over the same areas, under similar solar and satellite geometries, and marine and atmospheric conditions. All three OLCI processors exhibited good agreement between S-3B and S-3A OLCI $L_{WN}$ during the Tandem Phase, with RPD values ranging from 5% for POLYMER to 10% for l2gen (Figs. 13, S12 and S13). The spectral analysis during this phase revealed relative percentage differences of 2–8% in the 412–510 nm range, increasing towards the yellow and red parts of the spectrum with RPD values within 5–20% for 560 nm, 24–110% for 620 nm and 34–127% for 665 nm (Tables 4, 5, S4, S5, S6 and S7). These results are consistent with the findings of Lamquin et al. (2020b) and Zibordi et al. (2022). To address remaining biases between S-3A and S-3B OLCI, which stem from differences in spectral characterization and radiometric/geometric calibrations, a comprehensive and cross-mission consistent SVC should be implemented.

As the Drift Phase progressed, the two sensors began to drift apart into different orbits, resulting in increasing systematic differences in sensor overpass times for a given location. Upon intercomparing the S-3B and S-3A OLCI data products, we found significant and consistent discrepancies, with an overall RPD value of 22% for IPF-OL-2 and 18% for l2gen. While orbital differences lead to variations in target and light conditions, variability within our time window (± 1 h) should be considered random, if not negligible, and therefore cannot be solely responsible for the observed increase in biases after the Tandem Phase. On the contrary, the analysis conducted by (Zibordi et al., 2022) found that these biases were due to systematic differences in viewing geometry, suggesting a potential dependence of the atmospheric correction on viewing angle. Section 3.4 results indicated that the degradation during the Drift Phase was most and least pronounced for IPF-OL-2 and POLYMER respectively, suggesting that the POLYMER atmospheric correction may be less sensitive to these viewing geometry effects.

At the time of writing, there were no assessments of R2022 MODIS-Aqua, Suomi-VIIRS or NOAA-20 VIIRS radiometric accuracy beyond the matchup analysis from NASA OBPG as part of their R2022 reporting (NASA OBPG, 2022f, 2022h, 2022i). This validation was performed relative to all available matches from SeaBASS and AERONET-OC, so only a subset of the matchups are relevant to the open-ocean type waters of our study.

The assessment of MODIS-Aqua showed a consistent overestimation of $L_{WN}$ in all bands (Fig. 9). This is particularly evident in comparison to AMT26 *in situ* data and resulted in the lowest score for the AMT26–27 round-robin exercise. Studies using global R2018 data from a range of environments report similar (8–22%, Song et al. (2022), Mélin (2022)) and higher (20–45%, Barnes et al. (2019)) RPD values in the blue and green bands. These results could be due to an overcorrection of sensor degradation in the blue bands introduced in the R2018 data reprocessing and carried over to R2022. Although positive biases in the blue and green bands may cause an overestimation of band ratio values, and ultimately lead to the underestimation of Chl-a, Tilstone et al. (2021) demonstrated that MODIS-Aqua performed well and comparably to OLCI. This can be attributed to the use of band difference algorithms, which prioritise spectral shape over the absolute values of remote sensing reflectance.

We found that Suomi-VIIRS performed well in the blue and green bands, and less so in the red (Fig. 10). This is consistent with previous findings by Wang et al. (2015, 2016), who also reported that Suomi-VIIRS was accurate in the open ocean especially for blue and green bands. Large standard deviation values for bands above 600 nm due to low signal in the red from clear waters have also been reported in various Suomi-VIIRS assessments (*e.g.*, Wang and Jiang (2018), using MOBY data).

As with l2gen OLCI, the system vicarious calibration for NOAA-20 VIIRS was performed using an ocean reflectance model-based approach. This fact seemed to affect its performance when compared to the other NASA sensors, resulting in the lowest score of the AMT28–29 round-robin exercise. While NOAA-20 VIIRS performed worse than Suomi-VIIRS overall, there was a reasonable agreement between the two VIIRS sensors during the mission's overlap, as reported by Mélin (2021, 2022). Based on our data, the overall NOAA-20 VIIRS performance was comparable with the l2gen OLCI products.

Our findings indicate notable variations across the three NASA missions, despite utilising the same atmospheric correction approach. These differences may originate from multiple factors, including variances in sensor design, operation, and spectral properties (Barnes et al., 2019; Lyapustin et al., 2023), which can propagate through the processing stages (Li et al., 2019; Tilstone et al., 2021; Park et al., 2022). Additionally, the atmospheric correction process may respond diversely to distinct observation and illumination geometries, aerosol types and water types (Gordon, 2021; Li et al., 2022; Wang et al., 2022). Moreover, temporal and spatial disparities due to overpass times that are a few hours apart and the mapping of dissimilar pixel sizes and shapes could impact the resultant products.

The results shown in section 3.2 highlight an overestimate in the OLCI IPF-OL-2 product at blue and green bands in the low Chl-a waters of the Atlantic Ocean. OLCI $L_{WN}$ could be improved through use of the POLYMER atmospheric correction and through further updates of the system vicarious calibration of OLCI. The derivation of mission-wide gain factors for S-3 OLCI allows for a stable, single-baseline reprocessed dataset and it is key for the incorporation of OLCI data into climate records such as the ESA Ocean Colour Climate Change Initiative (OC-CCI) dataset. Though planning for the European Commission Copernicus Ocean Colour System Vicarious Calibration infrastructure (OC-SVC) is underway, its potential start of operations in 2027 is well beyond current mission timescales.

The AMT provides optimal conditions for the collection of radiometric FRM for the sustained validation of OLCI and other ocean-colour sensors over a range of optical conditions in the Atlantic Ocean, and especially in open-ocean regions. The AMT also has the potential to contribute to improve current atmospheric correction algorithms in open-ocean waters (Frouin et al., 2019). For example, Gossn et al. (2019) have shown that a baseline residual approach exploiting the new OLCI SWIR band at 1016 nm in conjunction with the 865 nm band reduces the correlation between water and aerosol signal, improving the atmospheric correction over extremely turbid waters. Similar methods could boost the AC performance regionally over North Atlantic gyre waters, where input from dust can be considerable (Pabortsava et al., 2017). The acquisition of collocated radiometry and Chl-a from both HPLC and underway absorption measurements during AMT campaigns has proved valuable to develop and validate bio-optical algorithms, as done by Brewin et al. (2016) for MODIS, MERIS and VIIRS, and further demonstrated by Tilstone et al. (2021) for S-3 OLCI.

## 5. Conclusions

We have shown how the AMT platform can provide fiducial reference measurements in the open ocean using a method that maximizes the quantity and quality of matchups. We have implemented a robust matchup procedure that takes into account the homogeneity, the quality, and the temporal and spatial variability characteristics of both *in situ* and satellite data. We have used this method together with *in situ* fiducial reference measurements and associated uncertainties, collected over four AMT campaigns, to evaluate three ocean-colour instruments (MODIS, VIIRS and OLCI) in Atlantic open-ocean waters.

The methodology presented here illustrates the need for moving platforms, such as ships, to perform careful and rigorous calculation of the uncertainties of *in situ* radiometric data to quality-control measurements for use in validation. This paper has defined thresholds for these uncertainties, which will serve as a useful guideline for autonomous measurements for future seaborne validation campaigns. We have demonstrated that the use of an uncertainty threshold can reduce bias in validation results, and that the validation statistics are not particularly sensitive to the use of varying temporal or spatial homogeneity matchup criteria, at least for the region under study.

Using these FRMs for the Atlantic Ocean we have shown that IPF-OL-2 RPD values were within 10%–21% in blue and blue-green bands. In the yellow and red, where the water leaving radiance signal is low and within the radiometric noise of the sensor, RPD values for IPF-OL-2 increased further. We found that the use of POLYMER with S-3 OLCI improves the retrieval performance of $L_{WN}$ at most wavelengths. For SeaDAS l2gen, we found a systematic bias between the *in situ* radiometric data, possibly due to vicarious gains.

FRMs acquired for AMT28 also provided a direct comparison between S-3A and S-3B OLCI $L_{WN}$ during the Tandem and Drift Phases. This analysis demonstrated the consistency of OLCI products across missions when flying in tandem. Outside of the Tandem Phase, we provided evidence for substantial systematic differences between the two OLCI sensors, which can be attributed to viewing angle interdependencies in the OLCI standard atmospheric correction IPF-OL-2.

Suomi-VIIRS $L_{WN}$ products RPD values were within 19% and 28% for the common bands (410, 443, 486, 551 and 671 nm), whereas MODIS-Aqua and NOAA20-VIIRS were closer to *in situ* with relative difference between 13% and 22%. The results highlighted the level of accuracy of the S-3 OLCI data streams in comparison with those from legacy missions. Of the ocean-colour sensors tested, S-3A OLCI provided the best performance for radiometric products.

## Author contributions

Concepts and methods were derived by SP, GT, RB and GDO and implemented by SP; data were collected by FN, GT, RB, GDO (*in situ*) and SP (satellite); data were processed by JL, GDO, HEK (*in situ*) and SP (satellite); analysis was performed by SP; draft preparation was done by SP and GT; review and editing was done by SP, GT, RB, GDO; funds to support the research were sought by CD, TC, GT, GDO, RB.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgements

The authors would like to thank the captain and crews of RRS Discovery on AMT27 and 29, and James Clark Ross on AMT26 and 28. SP, GDO, GT, JL, RB, HEK, were supported by the AMT4SentinelFRM (contract number ESRIN/RFQ/3-14457/16/I-BG), AMT4OceanSatFlux (contract number 000125730/18/NL/FF/gp) and AMT4CO2Flux (contract number 4000136286/21/NL/FF/ab) contracts funded by the European Space Agency. FN, RB, GDO and GT were also supported by the UK Natural Environment Research Council (NERC) National Capability funding to Plymouth Marine Laboratory for the Atlantic Meridional Transect. RB and GDO were also supported by the National Centre for Earth Observation (NCEO). RB was also supported by a UKRI Future Leader Fellowship (MR/V022792/1). We would also like to thank the Natural Environment Research Council Earth Observation Data Acquisition and Analysis Service (NEODAAS) for their role in the acquisition of satellite imagery and use of their Linux cluster for running the satellite models. This is contribution number 341 of the AMT program.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2023.113844.

## References

Alikas, K., Vabson, V., Ansko, I., Tilstone, G.H., Dall'Olmo, G., Nencioli, F., Vendt, R., Donlon, C., Casal, T., 2020. Comparison of above-water Seabird and TriOS radiometers along an Atlantic Meridional Transect. Remote Sens. 12, 1669. https://doi.org/10.3390/rs12101669.

Antoine, D., Guevel, P., Desté, J.F., Bécu, G., Louis, F., Scott, A., Bardey, P., 2008. The BOUSSOLE buoy - a new transparent-to-swell taut mooring dedicated to marine optics: design, tests, and performance at sea. J. Atmos. Ocean. Technol. 25, 968–989. https://doi.org/10.1175/2007JTECHO563.1.

Bailey, S.W., Werdell, P.J., 2006. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sens. Environ. 102, 12–23. https://doi.org/10.1016/j.rse.2006.01.015.

Bailey, S., Hooker, S., Antoine, D., Franz, B., Werdell, J., 2008. Sources and assumptions for the vicarious calibration of ocean color satellite observations. Appl. Opt. 47, 2035–2045. https://doi.org/10.1364/AO.47.002035.

Barnes, B.B., Cannizzaro, J.P., English, D.C., Hu, C., 2019. Validation of VIIRS and MODIS reflectance data in coastal and oceanic waters: an assessment of methods. Remote Sens. Environ. 220, 110–123. https://doi.org/10.1016/j.rse.2018.10.034.

Białek, A., Vellucci, V., Gentil, B., Antoine, D., Gorroño, J., Fox, N., Underwood, C., 2020. Monte Carlo–based quantification of uncertainties in determining ocean remote sensing reflectance from underwater fixed-depth radiometry measurements. J. Atmos. Ocean. Technol. 37, 177–196. https://doi.org/10.1175/JTECH-D-19-0049.1.

Brewin, R.J.W., Raitsos, D.E., Pradhan, Y., Hoteit, I., 2013. Comparison of chlorophyll in the Red Sea derived from MODIS-Aqua and in vivo fluorescence. Remote Sens. Environ. 136, 218–224. https://doi.org/10.1016/j.rse.2013.04.018.

Brewin, R.J.W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.Y., Devred, E., Doerffer, R., Fomferra, N., Franz, B., Grant, M., Groom, S., Horseman, A., Hu, C., Krasemann, H., Lee, Z., Maritorena, S., Mélin, F., Peters, M., Platt, T., Regner, P., Smyth, T., Steinmetz, F., Swinton, J., Werdell, J., White, G.N., 2015. The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. Remote Sens. Environ. 162, 271–294. https://doi.org/10.1016/j.rse.2013.09.016.

Brewin, R.J.W., Dall'Olmo, G., Pardo, S., van Dongen-Vogels, V., Boss, E.S., 2016. Underway spectrophotometry along the Atlantic Meridional Transect reveals high performance in satellite chlorophyll retrievals. Remote Sens. Environ. 183, 82–97. https://doi.org/10.1016/j.rse.2016.05.005.

Brockmann, C., Doerffer, R., Peters, M., Stelzer, K., Embacher, S., Ruescas, A., 2016. Evolution of the C2RCC Neural Network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. In: Living Planet Symposium, p. 54. URL: https://www.brockmann-consult.de/wp-content/uploads/2017/11/sco1_12.

Brown, S., Flora, S., Feinholz, M., Yarbrough, M., Houlihan, T., Peters, D., Kim, Y., Mueller, J., Johnson, B., Clark, D., 2007. The Marine Optical BuoY (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration. In: Proceedings of SPIE - The International Society for Optical Engineering, 6744. https://doi.org/10.1117/12.737400.

Cazzaniga, I., Kwiatkowska, E., 2018. Sentinel-3 OLCI Chlorophyll Index switch for low-chlorophyll waters. Technical Report. URL: https://www.eumetsat.int/media/47752 (accessed November 2022).

Clark, D.K., Yarbrough, M.A., Feinholz, M., Flora, S., Broenkow, W., Kim, Y.S., Johnson, B.C., Brown, S.W., Yuen, M., Mueller, J.L., 2003. MOBY, A radiometric buoy for performance monitoring and vicarious calibration of satellite ocean color sensors: Measurement and data analysis protocols. SeaWiFs Postlaunch Technical Report Series 4. URL: https://ntrs.nasa.gov/api/citations/20030063145/downloads/20030063145.pdf.

Clerc, S., Donlon, C., Borde, F., Lamquin, N., Hunt, S.E., Smith, D., McMillan, M., Mittaz, J., Woolliams, E., Hammond, M., Banks, C., Moreau, T., Picard, B., Raynal, M., Rieu, P., Guérou, A., 2020. Benefits and lessons learned from the sentinel-3 tandem phase. Remote Sens. 12 https://doi.org/10.3390/rs12172668.

Concha, J.A., Bracaglia, M., Brando, V.E., 2021. Assessing the influence of different validation protocols on Ocean Colour match-up analyses. Remote Sens. Environ. 259, 112415 https://doi.org/10.1016/j.rse.2021.112415.

Donlon, C., 2011. Sentinel-3 Mission Requirements Traceability Document (MRTD). Technical Report EOP-SM/2184/CD-cd.. European Space Agency. URL: https://sentinel.esa.int/documents/247904/1848151/Sentinel-3-Mission-Requirements-Traceability.

Dutkiewicz, S., Hickman, E., Jahn, O., Henson, S., Beaulieu, C., Monier, E., 2019. Ocean colour signature of climate change. Nat. Commun. 10, 578. https://doi.org/10.1038/s41467-019-08457-x.

EUMETSAT, 2021a. Sentinel-3 OLCI L2 Report for Baseline Collection OL_L2M_003. Technical Report. URL: https://www.eumetsat.int/media/47794.

EUMETSAT, 2021b. Sentinel-3 OLCI Marine User Handbook OL_L2M_003. Technical Report. URL: https://www-cdn.eumetsat.int/files/2021-03/Sentinel3%20OLCI%20Marine%20User%20Handbook.pdf.

Franz, B.A., Bailey, S.W., Werdell, P.J., McClain, C.R., 2007. Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. Appl. Opt. 46, 5068–5082. https://doi.org/10.1364/ao.46.005068.

Frouin, R.J., Franz, B.A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., Chowdhary, J., Dierssen, H.M., Tan, J., Dubovik, O., Huang, X., Davis, A.B., Kalashnikova, O., Thompson, D.R., Remer, L.A., Boss, E., Coddington, O., Deschamps, P.Y., Gao, B.C., Gross, L., Hasekamp, O., Omar, A., Pelletier, B., Ramon, D., Steinmetz, F., Zhai, P.W., 2019. Atmospheric correction of satellite ocean-color imagery during the PACE era. Front. Earth Sci. 7, 145. https://doi.org/10.3389/feart.2019.00145.

Garcia, C., Ferreira, A., Dogliotti, A., Garcia, V., Group-GOAL, H, 2011. Phytoplankton blooms in the Patagonian shelf-break and vicinities: bio-optical signature and performance of ocean color algorithms. In: AGU Fall Meeting Abstracts, 1633. https://doi.org/10.1029/2010JC006595.

GCOS, 2011. Systematic Observation Requirements from Satellite-Based Data Products for Climate 2011 Update. Supplemental Details to the Satellite-Based Component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC". URL: https://library.wmo.int/idurl/4/48411.

Gergely, M., Zibordi, G., 2014. Assessment of AERONET-OC LWN uncertainties. Metrologia 51, 40–47. https://doi.org/10.1088/0026-1394/51/1/40.

Giannini, F., Hunt, B.P., Jacoby, D., Costa, M., 2021. Performance of OLCI Sentinel-3A satellite in the Northeast Pacific coastal waters. Remote Sens. Environ. 256, 112317 https://doi.org/10.1016/j.rse.2021.112317.

Glover, D.M., Jenkins, W.J., Doney, S., 2011. Modeling Methods for Marine Science, 1st ed. Cambridge University Press. URL. http://www.cambridge.org/9780521867832.

Gordon, H.R., 2021. Evolution of ocean color atmospheric correction: 1970–2005. Remote Sens. 13, 5051. https://doi.org/10.3390/rs13245051.

Gossn, J.I., Ruddick, K.G., Dogliotti, A.I., 2019. Atmospheric correction of OLCI imagery over extremely turbid waters based on the red, NIR and 1016 nm bands and a new baseline residual technique. Remote Sens. 11, 220. https://doi.org/10.3390/rs11030220.

Gregg, W.W., Casey, N.W., 2004. Global and regional evaluation of the SeaWiFS chlorophyll data set. Remote Sens. Environ. 93, 463–479. https://doi.org/10.1016/j.rse.2003.12.012.

Gregg, W.W., Rousseaux, C.S., Franz, B.A., 2017. Global trends in ocean phytoplankton: a new assessment using revised ocean colour data. Remote Sens. Lett. 8, 1102–1111. https://doi.org/10.1080/2150704X.2017.1354263.

Harshada, D., Raman, M., Jayappa, K., 2021. Evaluation of the operational Chlorophyll-a product from global ocean colour sensors in the coastal waters, south-eastern Arabian Sea. Egypt. J. Remote Sens. Space Sci. 24, 769–786. https://doi.org/10.1016/j.ejrs.2021.09.005.

Hooker, S.B., Maritorena, S., 2000. An evaluation of oceanographic radiometers and deployment methodologies. J. Atmos. Ocean. Technol. 17, 811–830. https://doi.org/10.1175/1520-0426(2000)017<0811:AEOORA>2.0.CO;2.

Hooker, S.B., Morel, A., 2003. Platform and environmental effects on above-water determinations of water-leaving radiances. J. Atmos. Ocean. Technol. 20, 187–205. https://doi.org/10.1175/1520-0426(2003)020<0187:PAEEOA>2.0.CO;2.

Hooker, S.B., Lazin, G., McLean, S., Lazin, G., Zibordi, G., McLean, S., 2002. An evaluation of above- and in-water methods for determining water-leaving radiances. J. Atmos. Ocean. Technol. 19, 486–515. https://doi.org/10.1175/1520-0426(2002)019%3C0486:AEOAAI%3E2.0.CO;2.

JCGM, 2008. Guide to the Expression of Uncertainty in Measurement, Joint Committee for Guides in Metrology–JCGM 100 (ISO/IEC Guide 98–3). URL. http://www.iso.org/sites/JCGM/GUM-introduction.htm.

Kratzer, S., Plowey, M., 2021. Integrating mooring and ship-based data for improved validation of OLCI chlorophyll-a products in the Baltic Sea. Int. J. Appl. Earth Obs. Geoinf. 94, 102212 https://doi.org/10.1016/j.jag.2020.102212.

Kulk, G., Platt, T., Dingle, J., Jackson, T., Jönsson, B.F., Bouman, H.A., Babin, M., Brewin, R.J.W., Doblin, M., Estrada, M., Figueiras, F.G., Furuya, K., González-Benítez, N., Gudfinnsson, H.G., Gudmundsson, K., Huang, B., Isada, T., Kovač, A., Lutz, V.A., Marañón, E., Raman, M., Richardson, K., Rozema, P.D., Poll, W.H.V.D., Segura, V., Tilstone, G.H., Uitz, J., Dongen-Vogels, V.V., Yoshikawa, T., Sathyendranath, S., 2020. Primary production, an index of climate change in the ocean: satellite-based estimates over two decades. Remote Sens. 12, 826. https://doi.org/10.3390/rs12050826.

Kyryliuk, D., Kratzer, S., 2019. Evaluation of Sentinel-3A OLCI products derived using the case-2 regional CoastColour processor over the Baltic Sea. Sensors 19, 3609. https://doi.org/10.3390/s19163609.

Lamquin, N., Clerc, S., Bourg, L., Donlon, C., 2020a. OLCI A/B tandem phase analysis, part 1: level 1 homogenisation and harmonisation. Remote Sens. 12 https://doi.org/10.3390/rs12111804.

Lamquin, N., Déru, A., Clerc, S., Bourg, L., Donlon, C., 2020b. OLCI A/B tandem phase analysis, part 2: benefits of sensors harmonisation for level 2 products. Remote Sens. 12, 2702. https://doi.org/10.3390/rs12172702.

Lee, Z., Wei, J., Voss, K., Lewis, M., Bricaud, A., Huot, Y., 2015. Hyperspectral absorption coefficient of "pure" seawater in the range of 350–550 nm inverted from remote sensing reflectance. Appl. Opt. 54, 546–558. https://doi.org/10.1364/AO.54.000546.

Li, J., Jamet, C., Zhu, J., Han, B., Li, T., Yang, A., Guo, K., Jia, D., 2019. Error budget in the validation of radiometric products derived from OLCI around the China Sea from open ocean to coastal waters compared with MODIS and VIIRS. Remote Sens. 11 https://doi.org/10.3390/rs11202400.

Li, Q., Jiang, L., Chen, Y., Wang, L., Wang, L., 2022. Evaluation of seven atmospheric correction algorithms for OLCI images over the coastal waters of Qinhuangdao in Bohai Sea. Reg. Stud. Mar. Sci. 56, 102711 https://doi.org/10.1016/j.rsma.2022.102711.

Lin, J., Dall'Olmo, G., Tilstone, G.H., Brewin, R.J.W., Vabson, V., Ansko, I., Evers-King, H., Casal, T., Donlon, C., 2022. Derivation of uncertainty budgets for continuous above-water radiometric measurements along an Atlantic Meridional Transect. Opt. Express 30, 45648–45675. https://doi.org/10.1364/OE.470994.

Liu, Y., Röttgers, R., Ramírez-Pérez, M., Dinter, T., Steinmetz, F., Nöthig, E.M., Hellmann, S., Wiegmann, S., Bracher, A., 2018. Underway spectrophotometry in the Fram Strait (European Arctic Ocean): a highly resolved chlorophyll a data source for complementing satellite ocean color. Opt. Express 26, A678–A696. https://doi.org/10.1364/OE.26.00A678.

Lyapustin, A., Wang, Y., Choi, M., Xiong, X., Angal, A., Wu, A., Doelling, D.R., Bhatt, R., Go, S., Korkin, S., Franz, B., Meister, G., Sayer, A.M., Roman, M., Holz, R.E., Meyer, K., Gleason, J., Levy, R., 2023. Calibration of the SNPP and NOAA 20 VIIRS sensors for continuity of the MODIS climate data records. Remote Sens. Environ. 295, 113717 https://doi.org/10.1016/j.rse.2023.113717.

Martinez, E., Antoine, D., D'Ortenzio, F., Gentili, B., 2009. Climate-driven basin-scale decadal oscillations of oceanic phytoplankton. Science 326, 1253–1256. https://doi.org/10.1126/science.1177012.

Masoud, A.A., 2022. On the retrieval of the water quality parameters from Sentinel-3/2 and Landsat-8 OLI in the Nile Deltaś coastal and inland waters. Water 14, 593. https://doi.org/10.3390/w14040593.

Mazeran, C., Ruescas, A., 2021. Ocean Colour System Vicarious Calibration Tool. Technical Report. URL: https://www.eumetsat.int/media/47502. Accessed November 2022.

McClain, C.R., Feldman, G.C., Hooker, S.B., 2004. An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. Deep-Sea Res. II Top. Stud. Oceanogr. 51, 5–42. https://doi.org/10.1016/j.dsr2.2003.11.001.

Mélin, F., 2021. From validation statistics to uncertainty estimates: application to VIIRS ocean color radiometric products at European coastal locations. Front. Mar. Sci. 8 https://doi.org/10.3389/fmars.2021.790948.

Mélin, F., 2022. Validation of ocean color remote sensing reflectance data: analysis of results at European coastal sites. Remote Sens. Environ. 280, 113153 https://doi.org/10.1016/j.rse.2022.113153.

Mograne, M.A., Jamet, C., Loisel, H., Vantrepotte, V., Mériaux, X., Cauvin, A., 2019. Evaluation of five atmospheric correction algorithms over French optically-complex waters for the Sentinel-3A OLCI ocean color sensor. Remote Sens. 11, 668. https://doi.org/10.3390/rs11060668.

Moore, T.S., Campbell, J.W., Dowell, M.D., 2009. A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. Remote Sens. Environ. 113, 2424–2430. https://doi.org/10.1016/j.rse.2009.07.016.

Morel, A., Antoine, D., Gentili, B., 2002. Bidirectional reflectance of oceanic waters: accounting for raman emission and varying particle scattering phase function. Appl. Opt. 41, 6289–6306. https://doi.org/10.1364/AO.41.006289.

Müller, D., Krasemann, H., Brewin, R.J.W., Brockmann, C., Deschamps, P.Y., Doerffer, R., Fomferra, N., Franz, B., Grant, M., Groom, S., Mélin, F., Platt, T., Regner, P., Sathyendranath, S., Steinmetz, F., Swinton, J., 2015. The Ocean Colour Climate Change Initiative: I. A methodology for assessing atmospheric correction processors based on in-situ measurements. Remote Sens. Environ. 162, 242–256. https://doi.org/10.1016/j.rse.2013.11.026.

NASA OBPG, 2022a. NOAA-20 Visible and Infrared Imager/Radiometer Suite (VIIRS) Ocean Color Data. https://doi.org/10.5067/NOAA-20/VIIRS/L2/OC/2022. Reprocessing. accessed: 2022-10-20.

NASA OBPG, 2022b. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Ocean Color Data. https://doi.org/10.5067/AQUA/MODIS/L2/OC/2022. Reprocessing. accessed: 2022-10-20.

NASA OBPG, 2022c. SENTINEL-3A Ocean and Land Colour Imager (OLCI) Ocean Color Data. https://doi.org/10.5067/SENTINEL-3A/OLCI/L2/EFR/OC/2022. Reprocessing. accessed: 2022-10-20.

NASA OBPG, 2022d. SENTINEL-3B Ocean and Land Colour Imager (OLCI) Ocean Color Data. https://doi.org/10.5067/SENTINEL-3B/OLCI/L2/EFR/OC/2022. Reprocessing. accessed: 2022-10-20.

NASA OBPG, 2022e. SUOMI-NPP Visible and Infrared Imager/Radiometer Suite (VIIRS) Ocean Color Data. https://doi.org/10.5067/SUOMI-NPP/VIIRS/L2/OC/2022. Reprocessing. accessed: 2022-10-20.

NASA OBPG, 2022f. MODIS-Aqua Ocean Color Reprocessing 2022.0. Technical Report. URL: https://oceancolor.gsfc.nasa.gov/reprocessing/r2022/aqua/. Accessed: 2022-10-20.

NASA OBPG, 2022g. OLCI Ocean Color Reprocessing 2022.0. Technical Report. URL: https://oceancolor.gsfc.nasa.gov/reprocessing/r2022/olci/. Accessed: 2022-10-20.

NASA OBPG, 2022h. VIIRS-NOAA20 Ocean Color Reprocessing 2022.0. Technical Report. URL: https://oceancolor.gsfc.nasa.gov/reprocessing/r2022/noaa20/. Accessed: 2022-10-20.

NASA OBPG, 2022i. VIIRS-SNPP Ocean Color Reprocessing 2022.0. Technical Report. URL: https://oceancolor.gsfc.nasa.gov/reprocessing/r2022/snpp/. Accessed: 2022-10-20.

Pabortsava, K., Lampitt, R.S., Benson, J., Casciotti, K., Crowe, C., McLachlan, R., Le Moigne, F., Moore, C.M., Pebody, C., Provost, P., Rees, A., Tilstone, G., Woodward, E., 2017. Carbon sequestration in the deep Atlantic enhanced by Saharan dust. Nat. Geosci. 10, 189–194. https://doi.org/10.1038/ngeo2899.

Park, M.S., Lee, S., Ahn, J.H., Lee, S.J., Choi, J.K., Ryu, J.H., 2022. Decadal measurements of the first geostationary ocean color satellite (GOCI) compared with MODIS and VIIRS data. Remote Sens. 14 https://doi.org/10.3390/rs14010072.

Qin, P., Simis, S.G., Tilstone, G.H., 2017. Radiometric validation of atmospheric correction for MERIS in the Baltic Sea based on continuous observations from ships and AERONET-OC. Remote Sens. Environ. 200, 263–280. https://doi.org/10.1016/j.rse.2017.08.024.

Renosh, P.R., Doxaran, D., Keukelaere, L.D., Gossn, J.I., 2020. Evaluation of atmospheric correction algorithms for Sentinel-2-MSI and Sentinel-3-OLCI in highly turbid estuarine waters. Remote Sens. 12, 1285. https://doi.org/10.3390/rs12081285.

Ruddick, K.G., De Cauwer, V., Park, Y.J., Moore, G., 2006. Seaborne measurements of near infrared water-leaving reflectance: the similarity spectrum for turbid waters. Limnol. Oceanogr. 51, 1167–1179. https://doi.org/10.4319/lo.2006.51.2.1167.

Rudorff, N., Frouin, R., Kampel, M., Goyens, C., Meriaux, X., Schieber, B., Mitchell, B.G., 2014. Ocean-color radiometry across the Southern Atlantic and Southeastern Pacific: accuracy and remote sensing implications. Remote Sens. Environ. 149, 13–32. https://doi.org/10.1016/j.rse.2014.03.029.

Salama, M.S., Spaias, L., Poser, K., Peters, S., Laanen, M., 2022. Validation of Sentinel-2 (MSI) and Sentinel-3 (OLCI) water quality products in turbid estuaries using fixed monitoring stations. Front. Remote Sens. 2 https://doi.org/10.3389/frsen.2021.808287.

Siegel, D., Franz, B., 2010. Oceanography: century of phytoplankton change. Nature 466, 569–571. https://doi.org/10.1038/466569a.

Siemer, J.P., Machín, F., González-Vega, A., Arrieta, J.M., Gutiérrez-Guerra, M.A., Pérez-Hernández, M.D., Vélez-Belchí, P., Hernández-Guerra, A., Fraile-Nuez, E., 2021. Recent trends in SST, Chl-a, productivity and wind stress in upwelling and open ocean areas in the upper eastern North Atlantic subtropical gyre. J. Geophys. Res. Oceans 126. https://doi.org/10.1029/2021JC017268 e2021JC017268.

Song, Q., Chen, S., Hu, L., Wang, X., Shi, X., Li, X., Deng, L., Ma, C., 2022. Introducing two fixed platforms in the Yellow Sea and East China Sea supporting long-term satellite ocean color validation: preliminary data and results. Remote Sens. 14, 2894. https://doi.org/10.3390/rs14122894.

Steinmetz, F., Ramon, D., 2018. Sentinel-2 MSI and Sentinel-3 OLCI consistent ocean colour products using POLYMER. In: Remote Sensing of the Open and Coastal Ocean and Inland Waters, International Society for Optics and Photonics. SPIE, pp. 46–55. https://doi.org/10.1117/12.2500232.

Steinmetz, F., Deschamps, P.Y., Ramon, D., 2011. Atmospheric correction in presence of sun glint: application to MERIS. Opt. Express 19, 9783–9800. https://doi.org/10.1364/OE.19.009783.

Thuillier, G., Hersé, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., Simon, P.C., Mandel, H., 2003. The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the Atlas and Eureca missions. Sol. Phys. 214, 1–22. https://doi.org/10.1023/A:1024048429145.

Tilstone, G., Mallor-Hoya, S., Gohin, F., Couto, A.B., Sá, C., Goela, P., Cristina, S., Airs, R., Icely, J., Zühlke, M., Groom, S., 2017. Which ocean colour algorithm for MERIS in North West European waters? Remote Sens. Environ. 189, 132–151. https://doi.org/10.1016/j.rse.2016.11.012.

Tilstone, G., Dall'Olmo, G., Hieronymi, M., Ruddick, K., Beck, M., Ligi, M., Costa, M., D'Alimonte, D., Vellucci, V., Vansteenwegen, D., Bracher, A., Wiegmann, S., Kuusk, J., Vabson, V., Ansko, I., Vendt, R., Donlon, C., Casal, T., 2020. Field intercomparison of radiometer measurements for ocean colour validation. Remote Sens. 12, 1587. https://doi.org/10.3390/rs12101587.

Tilstone, G.H., Pardo, S., Dall'Olmo, G., Brewin, R.J.W., Nencioli, F., Dessailly, D., Kwiatkowska, E., Casal, T., Donlon, C., 2021. Performance of Ocean Colour Chlorophyll a algorithms for Sentinel-3 OLCI, MODIS-Aqua and Suomi-VIIRS in open-ocean waters of the Atlantic. Remote Sens. Environ. 260, 112444 https://doi.org/10.1016/j.rse.2021.112444.

Tilstone, G.H., Pardo, S., Simis, S.G.H., Qin, P., Selmes, N., Dessailly, D., Kwiatkowska, E., 2022. Consistency between satellite ocean colour products under high coloured dissolved organic matter absorption in the Baltic Sea. Remote Sens. 14, 89. https://doi.org/10.3390/rs14010089.

Vabson, V., Kuusk, J., Ansko, I., Vendt, R., Alikas, K., Ruddick, K., Ansper, A., Bresciani, M., Burmester, H., Costa, M., D'Alimonte, D., Dall'Olmo, G., Damiri, B., Dinter, T., Giardino, C., Kangro, K., Ligi, M., Paavel, B., Tilstone, G., Van Dommelen, R., Wiegmann, S., Bracher, A., Donlon, C., Casal, T., 2019. Field intercomparison of radiometers used for satellite validation in the 400–900 nm range. Remote Sens. 11, 1129. https://doi.org/10.3390/rs11091129.

Vanhellemont, Q., Ruddick, K., 2018. Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. Remote Sens. Environ. 216, 586–597. https://doi.org/10.1016/j.rse.2018.07.015.

Vanhellemont, Q., Ruddick, K., 2021. Atmospheric correction of Sentinel-3 OLCI data for mapping of suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. Remote Sens. Environ. 256, 112284 https://doi.org/10.1016/j.rse.2021.112284.

Wang, M., Jiang, L., 2018. VIIRS-derived ocean color product using the imaging bands. Remote Sens. Environ. 206, 275–286. https://doi.org/10.1016/j.rse.2017.12.042.

Wang, M., Liu, X., Jiang, L., Son, S., Sun, J., Shi, W., Tan, L., Naik, P., Mikelsons, K., Wang, X., et al., 2015. VIIRS ocean color research and applications. In: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, pp. 2911–2914. https://doi.org/10.1109/IGARSS.2015.7326424.

Wang, M., Jiang, L., Liu, X., Son, S., Sun, J., Shi, W., Tan, L., Mikelsons, K., Wang, X., Lance, V., 2016. VIIRS ocean color products: A progress update. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 5848–5851. https://doi.org/10.1109/IGARSS.2016.7730528.

Wang, J., Wang, Y., Lee, Z., Wang, D., Chen, S., Lai, W., 2022. A revision of NASA SeaDAS atmospheric correction algorithm over turbid waters with artificial neural networks estimated remote-sensing reflectance in the near-infrared. ISPRS J. Photogramm. Remote Sens. 194, 235–249. https://doi.org/10.1016/j.isprsjprs.2022.10.014.

Werdell, J., Bailey, S., 2005. An improved bio-optical data set for ocean color algorithm development and satellite data product variation. Remote Sens. Environ. 98, 122–140. https://doi.org/10.1016/j.rse.2005.07.001.

Werdell, P.J., Bailey, S.W., Franz, B.A., Morel, A., McClain, C.R., 2007. On-orbit vicarious calibration of ocean color sensors using an ocean surface reflectance model. Appl. Opt. 46, 5649–5666. https://doi.org/10.1364/AO.46.005649.

Werdell, J., McKinna, L.I., Boss, E., Ackleson, S.G., Craig, S.E., Gregg, W.W., Lee, Z., Maritorena, S., Roesler, C.S., Rousseaux, C.S., Stramski, D., Sullivan, J.M., Twardowski, M.S., Tzortziou, M., Zhang, X., 2018. An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. Prog. Oceanogr. 160, 186–212. https://doi.org/10.1016/j.pocean.2018.01.001.

Windle, A.E., Evers-King, H., Loveday, B.R., Ondrusek, M., Silsbe, G.M., 2022. Evaluating atmospheric correction algorithms applied to OLCI Sentinel-3 data of Chesapeake Bay waters. Remote Sens. 14, 1881. https://doi.org/10.3390/rs14081881.

Zibordi, G., Voss, K.J., 2014. Chapter 3.1 - in situ optical radiometry in the visible and near infrared. In: Zibordi, G., Donlon, C.J., Parr, A.C. (Eds.), Optical Radiometry for Ocean Climate Measurements, Volume 47 of Experimental Methods in the Physical Sciences. Academic Press, pp. 247–304. https://doi.org/10.1016/B978-0-12-417011-7.00010-6.

Zibordi, G., Berthon, J.F., Mélin, F., D'Alimonte, D., Kaitala, S., 2009a. Validation of satellite ocean color primary products at optically complex coastal sites: Northern Adriatic Sea, Northern Baltic Proper and Gulf of Finland. Remote Sens. Environ. 113, 2574–2591. https://doi.org/10.1016/j.rse.2009.07.013.

Zibordi, G., Mélin, F., Berthon, J.F., Holben, B., Slutsker, I., Giles, D., D'Alimonte, D., Vandemark, D., Feng, H., Schuster, G., Fabbri, B.E., Kaitala, S., Seppälä, J., 2009b. AERONET-OC: a network for the validation of ocean color primary products. J. Atmos. Ocean. Technol. 26, 1634–1651. https://doi.org/10.1175/2009JTECHO654.1.

Zibordi, G., Ruddick, K., Ansko, I., Moore, G., Kratzer, S., Icely, J., Reinart, A., 2012. In situ determination of the remote sensing reflectance: an inter-comparison. Ocean Sci. 8, 567–586. https://doi.org/10.5194/os-8-567-2012.

Zibordi, G., Mélin, F., Voss, K.J., Johnson, B.C., Franz, B.A., Kwiatkowska, E., Huot, J.P., Wang, M., Antoine, D., 2015. System vicarious calibration for ocean color climate change applications: requirements for in situ data. Remote Sens. Environ. 159, 361–369. https://doi.org/10.1016/j.rse.2014.12.015.

Zibordi, G., Talone, M., Voss, K.J., Johnson, B.C., 2017. Impact of spectral resolution of in situ ocean color radiometric data in satellite matchups analyses. Opt. Express 25, A798–A812. https://doi.org/10.1364/OE.25.00A798.

Zibordi, G., Mélin, F., Berthon, J.F., 2018. A regional assessment of OLCI data products. IEEE Geosci. Remote Sens. Lett. 15, 1490–1494. https://doi.org/10.1109/lgrs.2018.2849329.

Zibordi, G., Kwiatkowska, E., Mélin, F., Talone, M., Cazzaniga, I., Dessailly, D., Gossn, J. I., 2022. Assessment of OLCI-A and OLCI-B radiometric data products across European seas. Remote Sens. Environ. 272, 112911 https://doi.org/10.1016/j.rse.2022.112911.