Remote Sensing of Environment 270 (2022) 112848

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

An ensemble neural network atmospheric correction for Sentinel-3 OLCI over coastal waters providing inherent model uncertainty estimation and sensor noise propagation

Thomas Schroeder [a,*], Michael Schaale [b], Jennifer Lovell [c], David Blondeau-Patissier [a]

[a] CSIRO Oceans & Atmosphere, Brisbane, QLD 4001, Australia
[b] Freie Universität Berlin, Institute for Space Sciences, D-12165 Berlin, Germany
[c] CSIRO Oceans and Atmosphere, Hobart, TAS 7000, Australia

ARTICLE INFO

Keywords:
Sentinel-3
Atmospheric correction
Neural networks
Uncertainty propagation
Coastal waters

ABSTRACT

Accurate atmospheric correction (AC) is a prerequisite for quantitative ocean colour remote sensing and remains a challenge in particular over coastal waters. Commonly AC algorithms are validated by establishing a mean retrieval error from match-up analysis, which compares the satellite-derived surface reflectance with concurrent ground radiometric observations. Pixel-based reflectance uncertainties however, are rarely provided by AC algorithms and those for the operational Ocean and Land Colour Instrument (OLCI) marine reflectance product are not yet recommended for use. AC retrieval errors and uncertainties directly determine the quality with which ocean colour products can be estimated from the marine surface reflectance. Increasingly there is also the need for reflectance uncertainty products to be used as data assimilation inputs into biogeochemical models.

This paper describes the development of a new coastal AC algorithm for Sentinel-3 OLCI that provides pixel-based estimation of the inherent model inversion uncertainty and sensor noise propagation. The algorithm is a full-spectral model-based inversion of radiative transfer (RT) simulations in a coupled atmosphere–ocean system using an ensemble of artificial neural networks (ANN) that were initialized differently during the training process, but composed of the same network architecture. The algorithm has been validated against in-situ radiometric observations across a wide range of optical water types, and has been compared with the latest EUMETSAT operational Level 2 processor IPF-OL-2 v7.01.

In this analysis we found that the ensemble ANN showed improved performance over the operational Level 2 processor with a band-averaged (412–708 nm) mean absolute percentage error (MAPE) of 16% compared to 37% and a four-times lower band-averaged bias of -0.00045 sr^{-1}.

In the ensemble inversion process we account for three uncertainty components: (1) the total model variance that describes the variance of the data from the different ANNs, (2) the prediction variance of the mean, which is based on calculations of the RT simulations and (3) the instrument noise variance of the mean by propagating the OLCI spectral signal-to-noise ratios (SNR). To study algorithm performance and to quantify the contribution of the different uncertainty components to the total uncertainty, we applied the algorithm to an optically complex full resolution (FR) test scene covering coastal waters of the Great Barrier Reef, Australia. The uncertainties associated with the instrument noise variance were found to be two orders of magnitude lower than the uncertainty components of the prediction and total model variances. The overall largest uncertainty component in our uncertainty framework is attributed to the total model inversion error from averaging the responses of the slightly different adapted networks in the ensemble. The algorithm is made publicly available as a Python/C plugin for the Sentinel Application Platform (SNAP).

* Corresponding author.
E-mail address: Thomas.Schroeder@csiro.au (T. Schroeder).

https://doi.org/10.1016/j.rse.2021.112848
Received 24 May 2021; Received in revised form 2 December 2021; Accepted 6 December 2021
Available online 21 December 2021

1. Introduction

After more than 30 years of operational ocean colour remote sensing, atmospheric correction (AC) still remains a challenge, in particular over optically complex coastal waters. Algorithms are constantly refined, improved and validated. Validation of AC algorithms is important because their accuracy directly determines the accuracy with which satellite ocean color products can be derived from subsequent processing using atmospherically corrected reflectance spectra as input.

AC is the process of removing the effects of Rayleigh and Mie scattering as well as absorption of atmospheric gases and the influence of certain types of aerosols from remotely sensed imagery. Eq. (1) summarizes the different radiance sources that contribute to the total remotely sensed signal over water at top of atmosphere (TOA) (Gordon, 1997). By neglecting contributions from whitecaps and sun-glint the total radiance L_{TOA} can be written as a sum of the radiance contributions from molecular scattering L_R, aerosol scattering L_A, any interactions between aerosols and molecules L_{RA}, and from the radiance leaving the water L_W:

$$L_{TOA} = L_R + L_A + L_{RA} + tL_W \quad (1)$$

where t is the diffuse transmittance along the pixel-to-sensor path (height, spectral and geometric dependencies were omitted in this notation).

Over the ocean the remotely sensed signal at TOA is dominated by atmospheric path radiance generated by scattered photons that have not interacted with the ocean's body (terms 1–3 of Eq. 1). The shorter spectral bands (400–700 nm) are strongly affected by atmospheric scattering, which is inversely proportional to the wavelength – with dependencies ranging from λ^{-4} for molecular (Rayleigh) scattering to $\lambda^{-0.5}$–$\lambda^{-2.5}$ for aerosol (Mie) scattering. To enable multitemporal image analysis these contributions have to be removed from the imagery to obtain the water-leaving radiance (term 4 of Eq. 1) that provides information about the optically significant water constituents.

For quantitative image analysis AC has been recognized as an important processing step since the very first operational ocean colour measurements by the Coastal Zone Color Scanner in 1978 (Gordon et al., 1983). Since then, many empirical or radiative transfer-based approaches have been developed for the correction of atmospheric scattering and absorption. Empirical image-based approaches, such as the empirical line method or dark object subtraction (Karpouzli and Malthus, 2003; Nazeer et al., 2014) have proven to be of limited use in automated multi-temporal image studies as they rely on invariant-in-time surface reflectance calibration targets or measurements concurrent to the satellite or airborne image acquisition and are therefore not further discussed. Removal of atmospheric effects over turbid coastal waters is generally more complex than over open ocean waters where atmospheric path radiance can be separated under the assumption of negligible water-leaving radiance in the near-infrared (NIR) spectral region (>700 nm) (IOCCG, 2010). This assumption is commonly referred to as the black pixel assumption (Siegel et al., 2000) and allows decoupling of atmospheric and oceanic signals, since the TOA radiance is then entirely due to atmospheric path radiance in the NIR. Rayleigh scattering is well understood (Wang, 2016) and the atmospheric path radiance can be further corrected for molecular scattering using auxiliary information about the surface barometric pressure. The aerosol optical properties can be estimated from the remaining decoupled and Rayleigh-corrected NIR signal by analyzing the spectral slopes using at least two NIR bands. Shorter wavelengths (400–700 nm) can be corrected by extrapolation of the estimated aerosol optical properties obtained from inversion of pre-computed look-up tables (Gordon and Wang, 1994; Antoine and Morel, 1999).

However, the black pixel assumption becomes invalid above turbid, highly scattering waters (Lavender et al., 2005; Morel and Bélanger, 2006). Various additional corrections have been developed that attempt to remove the NIR water leaving radiance contributions from the TOA signal in order to re-establish the black pixel assumption (Siegel et al., 2000; Ruddick et al., 2000; Lavender et al., 2005; Bailey et al., 2010; Moore and Lavender, 2010; Vanhellemont, 2020). Other methods try to overcome these limitations by entirely shifting the aerosol retrieval into the shortwave infrared (SWIR) spectral region (~1000–3000 nm), where the absorption of pure water dominates the in-water optical properties and the black pixel assumption becomes valid again including for moderately turbid waters (Gao et al., 2007; Wang and Shi, 2007; Wang, 2007; Gossn et al., 2019; Ibrahim et al., 2019). Yet, not all ocean colour satellites offer SWIR spectral bands and researchers therefore have explored alternative correction methods that utilize shorter bands in the ultraviolet (UV) spectral region (≤400 nm) or a combination of UV-NIR bands to re-establish the black pixel assumption over waters with strong absorption of detritus and coloured dissolved organic matter (CDOM) (Wang, 2007; He et al., 2012; Singh et al., 2019).

In recent years several alternative AC methods were developed that are more suitable for application above turbid and optically complex coastal waters. These were summarized by the International Ocean Color Coordinating Group (IOCCG) in report number 10 (IOCCG, 2010), and include methods such as spectral matching or spectral optimization (Kuchinke et al., 2009; Steinmetz et al., 2011) and artificial neural networks (Jamet et al., 2005; Brajard et al., 2006; Schroeder et al., 2007a; Doerffer, 2008; Brockmann et al., 2016; Fan et al., 2017). The recent study by Frouin et al. (2019) provides a comprehensive overview of the history and progress of atmospheric correction during the last two decades. It also provides a good overview of parametric algorithms for atmospheric correction such as neural networks.

The reliability of an AC algorithm is critical for any subsequent application such as the processing of the satellite-derived reflectance through bio-optical models. In addition, mission targets, e.g. 5% relative uncertainty in L_w for blue-green bands in open ocean waters (Hooker and McClain, 2000; GCOS, 2016), are required to ensure the detection of chlorophyll with an acceptable accuracy across a wide range of concentrations (Hooker et al., 1992). The overall accuracy of an AC algorithm is usually determined by comparison of the satellite-derived reflectance against in-situ radiometric observations. This match-up approach quantifies the mean retrieval accuracy ideally across a large number of concurrent in-situ and satellite observations. For optically complex waters AC retrieval errors are substantially higher than in open ocean waters. Recent validation studies comparing various AC algorithms applied to Sentinel-3 OLCI over coastal and inland waters show mean absolute percentage retrievals errors in the range of 30–440% in the blue spectral region at 412.5 nm, 7–52% in the green at 560 nm and 11–78% in the red at 708 nm (Bi et al., 2018; Li et al., 2019; Alikas et al., 2020; Shen et al., 2020; Giannini et al., 2021; Vanhellemont and Ruddick, 2021).

In addition to a good overall retrieval accuracy certain applications, such as the assimilation of marine reflectance into biogeochemical models (BGCM) (Baird et al., 2016; Jones et al., 2016; Steven et al., 2019), also require information about the pixel-by-pixel retrieval uncertainty. A new study of the impact of measurement uncertainties has been presented recently by McKinna et al. (2021) who demonstrated how standard uncertainties might be included into ocean colour validation by correcting their validation results of an empirical ocean colour algorithm for model and observation uncertainties. Accounting for these uncertainties in their validation metric generally improved their validation results.

An often recommended document for expressing uncertainty in measurements is provided in the *Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement* (GUM) (JCGM, 2008), whose nomenclature is increasingly adopted by the ocean colour community (Białek et al., 2020; McKinna et al., 2019; McKinna et al., 2021). The metrological terms used in GUM, such as uncertainty, are defined in the *International vocabulary of basic and general terms in metrology* (VIM) (Metrology, 2017), while basic statistical terms used in GUM were

sourced from the International Standard ISO 3534-1. The term uncertainty in GUM is defined as a "parameter associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand". As such, the variance or its positive square root being the standard deviation of a measurement can be used to describe the dispersion and hence its uncertainty.

IOCCG report 18 (IOCCG, 2019) provides an overview of methods and the current state of research for uncertainty propagation in ocean colour remote sensing. First-order Taylor series approximation, often referred to as the *law of propagation of uncertainty* (Taylor and Kuyatt, 1994) provides a computationally-efficient framework to study uncertainty propagation analytically. Several ocean colour studies have demonstrated analytical solutions for the propagation of uncertainties associated with the marine reflectance to its derived bio-optical quantities (Lee et al., 2010; Maritorena et al., 2010; McKinna et al., 2019; McKinna et al., 2021). Another approach for uncertainty propagation is the model-based approach using Monte Carlo (MC) RT-simulations (JCGM S1, 2008). The MC-approach has been successfully used to evaluate uncertainty approximations from propagating radiometric uncertainties through ocean colour and bio-optical models (Milutinović and Bertino, 2011; McKinna et al., 2019) or through specific measurement functions such as for estimating in-situ above water radiometric quantities (Białek et al., 2020).

Often a more complete per-pixel uncertainty budget is required for satellite ocean colour observations taking into account the uncertainty of the TOA radiometric signal and the uncertainty associated with the AC algorithm. Per-pixel uncertainty estimates however, are rarely provided by AC algorithms as their calculation can be computationally expensive especially if the underlying remote sensing problem is highly non-linear. A Bayesian approach to AC, and quantification of uncertainty in the water reflectance has been implemented by Frouin and Pelletier (2015) and applied to SeaWiFS imagery, which takes into account a TOA noise term that incorporates measurement errors and modeling uncertainty. A computationally-efficient first order Taylor approximation has been applied by Gillis et al. (2018) to simulated hyper-spectral data to investigate how uncertainty in sensor TOA radiance is propagated through to the uncertainty in the derived remote sensing reflectance. McKinna et al. (2019) implemented a similar approach applied to SeaWiFS data, but took into account the spectral covariance of the TOA radiances, which were estimated over clear open ocean waters. This approach of analytical uncertainty propagation of radiometric noise in marine reflectance has also been implemented operationally for Sentinel-3 in the OLCI clear water branch and is outlined in detail in ACRI-ST (2013) and does not consider systematic errors from AC. Provision of pixel-by-pixel uncertainty estimation is a key requirement for all Sentinel-3 OLCI ocean colour products. However, four years into the Sentinel-3 mission the use of these uncertainty products is still not recommended by EUMETSAT due to lack of verification (EUMETSAT, 2019).

This study presents the development and validation of an AC algorithm for Sentinel-3 OLCI over optically complex waters, including the development of a framework to provide pixel-by-pixel uncertainty estimates of the derived marine reflectance. The algorithm is implemented as an ensemble of artificial neural networks (ANN) that accounts for the propagation of Sentinel-3 OLCI TOA random radiometric noise and the estimation of the AC inversion model uncertainty. Rather than calculating the per-pixel variance from a single observation, the ANN ensemble approach computes the variance of the mean as an uncertainty measure from the arithmetic mean of the individual network responses for a given observation. In detail, we account for three uncertainty components: (1) the total model variance that describes the mean variance of the data from the different ANNs, (2) the prediction variance of the mean, which is computed from calculations of the RT simulations that were used to train the ANN models and (3) the instrument noise variance of the mean by propagating the OLCI spectral signal-to-noise ratios (SNR). Radiometric uncertainties associated with absolute calibration errors of OLCI (bias) are not taken into account. This model-based approach does not calculate the uncertainties via explicit analytical uncertainty propagation as outlined in GUM but from the parametric ANN ensemble methodology. Ensemble modeling is commonly used in weather forecasting or climate predictions where models with different initial conditions are used to quantify the sensitivity to small errors (Parker, 2013). The larger the spread of the ensemble predictions the greater the uncertainty. The same concept is explored in our framework where the different initial conditions are provided by ANNs with varying internal weight configurations.

Validation of the AC algorithm is achieved using in-situ reflectance observations. To illustrate the algorithm performance across different water types, ranging from clear blue open ocean waters to highly scattering turbid and absorbing productive coastal waters, we apply the algorithm to an optically complex full resolution image of the central Great Barrier Reef, Australia. The contribution of the different uncertainty components across these water types is discussed, followed by a discussion of the algorithm limitations and future directions.

2. Methods

The ANN algorithm developed in this study was adapted from an approach previously implemented by Schroeder et al. (2007a, 2007b) for MERIS, but on the basis of a different learning algorithm and is now significantly enhanced by providing pixel-based uncertainty estimation and extended out-of-scope testing of input and output spectra. The algorithm class is a full spectral physics-based inversion method that, in contrast to standard atmospheric correction algorithms, does not attempt to decouple atmospheric and oceanic light fields (Siegel et al., 2000). The algorithm performs a pixel-by-pixel direct inversion of the TOA signal into spectral remote sensing reflectance $R_{rs}(\lambda)=L_W(\lambda)/E_d(\lambda)$ [sr^{-1}] defined as the ratio of the water-leaving radiance $L_W(\lambda)$ and the downwelling irradiance $E_d(\lambda)$, with associated sensor and inverse model uncertainties at mean sea level. The sun and observing geometry dependence of these quantities have been omitted for brevity.

2.1. Radiative transfer simulations

The parameterization of the forward RT simulations are described in detail in Schroeder (2005) and Schroeder et al. (2007a, 2007b), and briefly summarized here. As a forward model we used a scalar version of the Matrix-Operator-MOdel (MOMO) (Fell and Fischer, 2001; Fischer and Grassl, 1984; Hollstein and Fischer, 2012), which was extensively used for look-up table (LUT) computations in the MERIS Ground Segment (ESA, 2013). MOMO solves the radiative transfer equation numerically based on the Matrix Operator Method (Plass et al., 1973) and has been validated with other radiative transfer codes and spectral radiation observations (Santer et al., 2005).

Here MOMO was used to simulate the light field in a coupled ocean-atmosphere system and to build a large database of more than 20 million reflectance spectra at the bottom of the atmosphere (BOA) and at TOA. This database was established from mono-chromatic azimuthally-resolved simulations for a variety of different sun and observing angles as well as variations of the environmental conditions associated with different concentrations of oceanic and atmospheric constituents briefly summarized in Table 1. Model inputs consist of inherent optical properties (IOPs) of the atmospheric and oceanic constituents that were derived from Mie computations and in-water bio-optical model approximations.

All simulations were performed for a US Standard Atmosphere and eight aerosol assemblages, each composed of a maritime and a continental aerosol model in the boundary layer (0–2 km) and troposphere (2–12 km), and a sulfuric acid background aerosol model in the stratosphere (12–50 km). The spectral aerosol extinction coefficients, at relative humidities between 70 and 99%, were provided by variable aerosol optical thickness (*AOT*) values ranging between 0.03 and 1 at

Table 1
Variations of the sun and observing geometry and environmental conditions accounted for in the forward RT simulation resulting in an ANN training database of 24,640,000 spectra.

Geometry (7,700 cases)	14	Sun zenith angles [0–76°]
	11	Observer zenith angles [0–59°]
	25	Relative azimuth angles [0–180°]
	2	Height levels [0 km (BOA), 50 km (TOA)]
Environmental conditions (3,200 cases)	1	US Standard Atmosphere
	8	Aerosol assemblages (Mix of maritime, continental and H_2SO_4 at rel. humidity 70–99%)
	5	*AOT*s at 550 nm [0.03–1]
	2	Barometric pressure values [980, 1040 hPa]
	2	Wind speeds [1.5, 7.2 m s^{-1}]
	20	Random concentration triplicates of: chlorophyll-a [0.05–50 mg m^{-3}], TSM [0.05–50 g m^{-3}] and CDOM [0.005–1 m^{-1}] at 443 nm
Sensor bands	12	412.5, 442.5, 490, 510, 560, 620, 665, 708.75, 753.75, 778.75, 865, and 885 nm

550 nm. A rough sea surface characterized by wind speeds of 1.5 and 7.2 m s^{-1} and barometric air pressure variations of 980 and 1040 hPa were considered for the atmospheric parameterization.

The oceanic model was parameterized by randomly varying the concentrations of optically-active water constituents such as of chlorophyll-a, total suspended matter (TSM) and CDOM absorption at 443 nm. The underlying bio-optical model providing the corresponding IOPs for these concentration ranges was derived from European observations (Babin, 2000). Inelastic scattering and white caps were not considered in the simulations, which were subsequently used to adapt the inverse ANN algorithm. In the radiative transfer simulations the satellite response functions are accounted for in the process of calculating the atmospheric transmission functions using a modified k-distribution method (Bennartz and Fischer, 2000). In this study, due to the opto-mechanical and imaging design similarity of the MERIS and OLCI sensors, we reused the MERIS simulations but adapted the inverse model to the different observing geometry of the OLCI sensor.

2.2. ANN algorithm development

The ANN was implemented as a three-layer perceptron (Bishop, 1995) and in this application represents a multi-variate nonlinear function mapping between the TOA spectral reflectance (input) and the BOA spectral reflectance (output). Within such a network each layer consists of neurons – which are the basic, linear or non-linear, processing nodes. Each neuron is connected with each neuron of the next layer by a weight. The weights, in statistical terms the free parameters, were initialized by small random numbers using a fixed seed, and estimated during a supervised learning procedure in which the network "learned" to associate an input vector $\underline{X}$ with a given output vector $\underline{Y}$. The weights between two layers can be expressed as a matrix $\boldsymbol{W}$ and the complete analytic function represented by a three-layer network is then given by Eq. (2).

$$\underline{Y} = S_2 \times \left\{ \boldsymbol{W}_2 \times S_1 \left(\boldsymbol{W}_1 \times \underline{X} \right) \right\} \tag{2}$$

For the ANN architecture presented in this study, the activation function is linear for the output layer (S_2) and non-linear (logistic) for the hidden layer (S_1). Training of a network consisted of minimizing the sum of squared deviations between the desired (training values) and the network's recall values by adapting the weight matrices ($\boldsymbol{W}_1$, $\boldsymbol{W}_2$) iteratively using a Limited Memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm (Liu and Nocedal, 1989).

The training data was extracted randomly from the simulated database. In detail, we extracted N_T=100,000 training and N_V=100,000 independent validation spectra at BOA and TOA, of which one 18-dimensional input vector $\underline{X}$ (Eq. (3)) consisted of the full TOA spectral reflectance $\rho_{\mathrm{TOA}}(\lambda)$ for Sentinel-3 bands Oa2–8, 11–12, 16–18 (λ=[412.5–885] nm, Table 3), the surface wind speed *ws*, the surface barometric pressure *p*, the cosine of the sun zenith $\cos(\theta_0)$ and the angular information of the observing geometry transformed into Cartesian coordinates *x, y, z*.

$$\underline{X} = [\rho_{\mathrm{TOA}}(\lambda), ws, p, cos(\theta_0), x, y, z]^T \tag{3}$$

The geometry transformation into Cartesian coordinates is an image processing artifice that helps to avoid image discontinuities that would result from the inversion due to the discontinuity of the azimuth difference and the observer zenith across the scan-lines. The training and validation subset sizes (N_T=N_V=100,000 samples) were based on the demand for a low error of the sample standard deviation $\Delta s = s/\sqrt{2(n-1)}$ (Brandt, 2014). Using $\Delta s/s$ to derive relative errors of those standard deviations and by choosing an error margin below 1% by also aiming for a confidence level of 99.9% (3.29 σ) one arrives at a sample size N_T=N_V=100,000.

The associated 12-dimensional output vector $\underline{Y}$ (Eq. (4)) contained the log-transformed remote sensing reflectance R_{rs} [sr^{-1}] at BOA for Sentinel-3 bands Oa2–8 and 11 (λ=[412.5–708.75] nm) and the aerosol optical thickness (*AOT*) at four AERONET wavelengths (440, 550, 670 and 870 nm), which were not further evaluated in this study.

$$\underline{Y} = \left[log_{10}\left(R_{\mathrm{rs,BOA}}(\lambda)\right), AOT(\lambda) \right]^T \tag{4}$$

As there is no direct pathway to obtain the optimum network architecture, a series of different networks were trained by varying the number of neurons in the hidden layer between 20 and 120 by accounting for different input transformations (e.g., PCA) and added noise levels. Training was stopped for each network configuration after 1000 iterations over the full training data set of 100,000 input vectors (Eq. (3)) and monitored by the Mean Squared Error (MSE). The best performing network architecture associated with the lowest MSE was selected by comparing the inversion performance on an independent test data set of 100,000 spectra that was not used during the training process. For this network architecture four additional networks were trained following the same processing framework but using a different weight initialization by selecting a random seed. These five networks were then pooled together into a network ensemble for the subsequent development of the uncertainty framework and the R_{rs} validation against in-situ radiometric observations. As the computation time significantly increases with the number of networks, especially for the computation of the uncertainties, the ensemble size was limited to five networks (N_a=5). The uncertainties will be computed from the different responses of the five networks to the input data set. It is important to emphasize that the individual network architectures of the ensemble (e.g. number of hidden layer neurons) are identical and that only the response for a given input will be different due to the different training initialization and final weight configuration.

Detection of out-of-scope input and output spectra is achieved by implementation of an extended validation test filter (eVT) (Schaale and Schroeder, 2013) that extends a simple minimum-maximum data validity test by a density test to ensure that all data inversion points fall in the regions of finite data density. The inverse model development flow diagram provides and overview of these processing steps (Fig. 1).

2.3. Uncertainty estimation

When utilizing just one ANN it is difficult and very time-consuming to calculate uncertainties of the retrieved values (Aires, 2004a, 2004b, 2004c; Bishop, 1995). Here the results of an ensemble of five networks (see section 2.2) are used to estimate retrieval uncertainties, which are based on three vectorial variances of a mean value: (1) the model variance of the mean $\overline{\sigma^2_{ANN}}$, (2) the prediction variance of the mean $\overline{\sigma^2_P}$ and (3) the instrument noise variance of the mean $\overline{\sigma^2_N}$, in addition to a

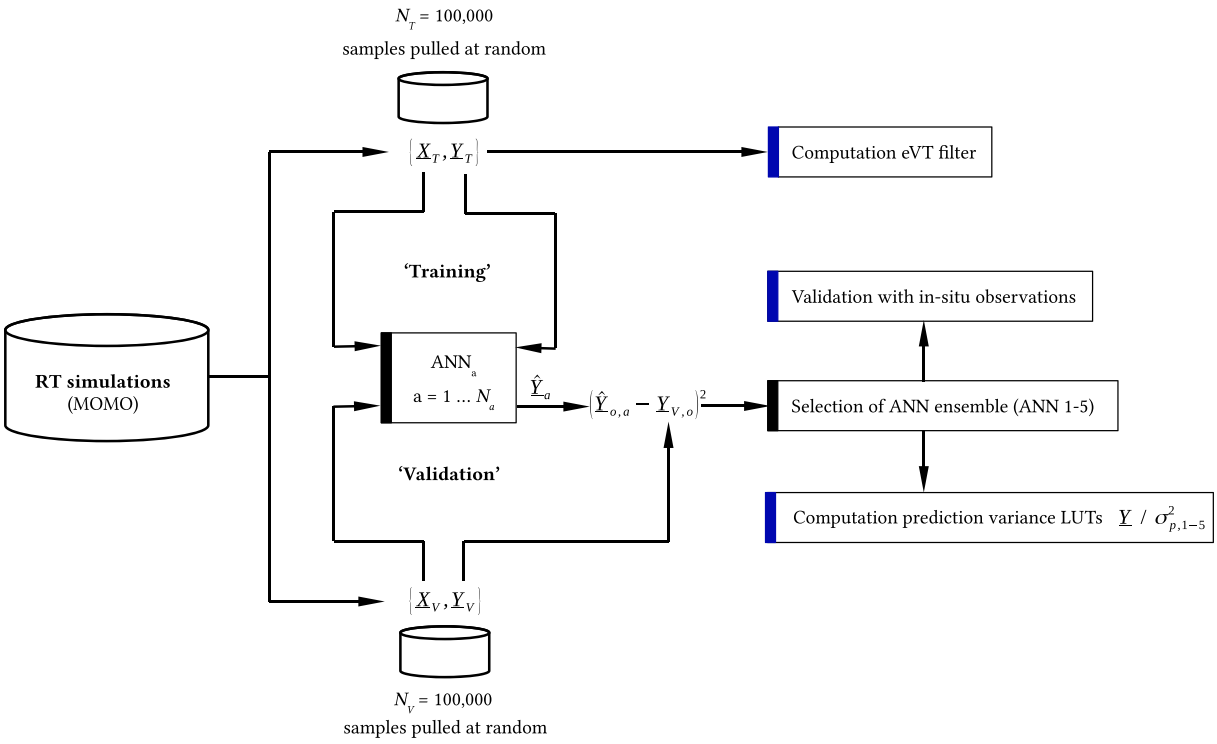

Fig. 1. Data flow for ANN training and validation and subsequent computation steps marked in blue.

'non-vanishing' model variance term $\underline{\sigma^2}_{ANN}$ (see Fig. 2).

The 'non-vanishing' model variance term was added to this framework as the three variances of the mean, which are divided by the number of networks N_a, (in theory) will vanish for large N_a ($N_a \to \infty$). The square root of these variance components, which will be introduced in the following sections, is what we define as the total uncertainty in this framework.

2.3.1. Model variance of the mean

The model variance $\underline{\sigma^2}_{ANN}$ is the most simple term and describes the variance of the data around the mean of the different ANN results. After the recall with test data $\underline{X}_V$, five response vectors $\hat{\underline{Y}}_a$ (a=1,...,5) are obtained. The mean response $\overline{\hat{\underline{Y}}}$ for N_a=5 networks is given by Eq. (5).

$$\overline{\hat{\underline{Y}}} = \frac{1}{N_a}\sum_{a=1}^{N_a}\hat{\underline{Y}}_a \tag{5}$$

The variance per dimension 'o' of the results is then given by Eq. (6).

$$\sigma^2_{ANN,o} = \sum_{i=1}^{N_a}\left(\hat{Y}_{o,i} - \overline{\hat{Y}}_{o,i}\right)^2 \Big/ (N_a - 1) \tag{6}$$

The $\sigma^2_{ANN,o}$ form the components of the vector of output model variances $\underline{\sigma^2}_{ANN}$. The strict utilization of the uncertainty propagation mechanism under the assumption that each $\hat{\underline{Y}}_{a,o}$ has the same standard uncertainty $\sqrt{\sigma^2_{ANN,o}}$ leads to the uncertainty of the mean expressed by Eq. (7).

$$\overline{\sigma^2_{ANN,o}} = \sigma^2_{ANN,o} \Big/ N_a \tag{7}$$

The $\overline{\sigma^2_{ANN,o}}$ in turn, form the components of the vector of output variances of the mean $\overline{\underline{\sigma^2}_{ANN}}$. Thus $\underline{\sigma^2}_{ANN}$ is an N_a-dimensional vector holding the model variances while the vector $\overline{\underline{\sigma^2}_{ANN}}$ contains the model variances of the means (see Fig. 2).

2.3.2. Prediction variance of the mean

The estimates of the prediction variance, which is in general the magnitude of the model variance, are based on calculations with the simulated validation data set (N_V=100,000) rather than the training data set because this introduces a response of the networks to unknown data. The validation data set produces (N_V=100,000) outputs for each network ANN$_a$ (a=1,...,5) (Fig. 3).

The data pairs $\underline{Y}_V^{(n)}/\hat{\underline{Y}}_a^{(n)}$ being available after the application of an input $\underline{X}_V^{(n)}$ to an ANN$_a$ are used to define the squared error difference $se_a^{(n)} = (\hat{\underline{Y}}_a^{(n)} - \underline{Y}_V^{(n)})^2$ as an objective error measure. The total sum of the squared differences $se_{T,a} = (\|\hat{\underline{Y}}_{a,T} - \underline{Y}_T\|)^2$ is the so-called (scalar) cost function, which is minimized during the ANN training process. In turn the sum of the squared differences $se_a = se_{V,a} = (\|\hat{\underline{Y}}_a - \underline{Y}_V\|)^2$ (with $\hat{\underline{Y}}_a = \hat{\underline{Y}}_{a,V}$) is also a scalar term, which will deviate from the cost function value because an ANN$_a$ is an approximation and that deviation indicates the level of generalization achieved when comparing this value against the value of the cost function. The sum of squared differences is a sum over all dimensions of an output vector and the number of samples. Therefore, it is important to split this sum into the dimension components ($se_{a,o}$) and to form small intervals of binned values for each output dimension 'o', because it is unlikely that the overall sum value $se_a = (\|\hat{\underline{Y}}_a - \underline{Y}_V\|)^2$ results from a sum of squared error values being equally distributed in dimensions and in magnitude over the whole range of output values.

Prior to the histogram generation its bins have to be fixed in size and position (Fig. 4, Eq. (8))

$$\begin{aligned} &b_0 = b_{min} \\ &b_{n_b} = b_{max} \\ &\Delta b = b_{j+1} - b_j \Rightarrow n_b = (b_{max} - b_{min})/\Delta b \end{aligned}$$

$$b_j = b_{min} + j\cdot\Delta b \text{ with } j = 0, \ldots, n_b \tag{8}$$

The lower bounds b_{min}, upper bounds b_{max} and bin sizes Δb depend on the retrieval under consideration and are given for the log-transformed R_{rs} and linear AOT in Table 2.

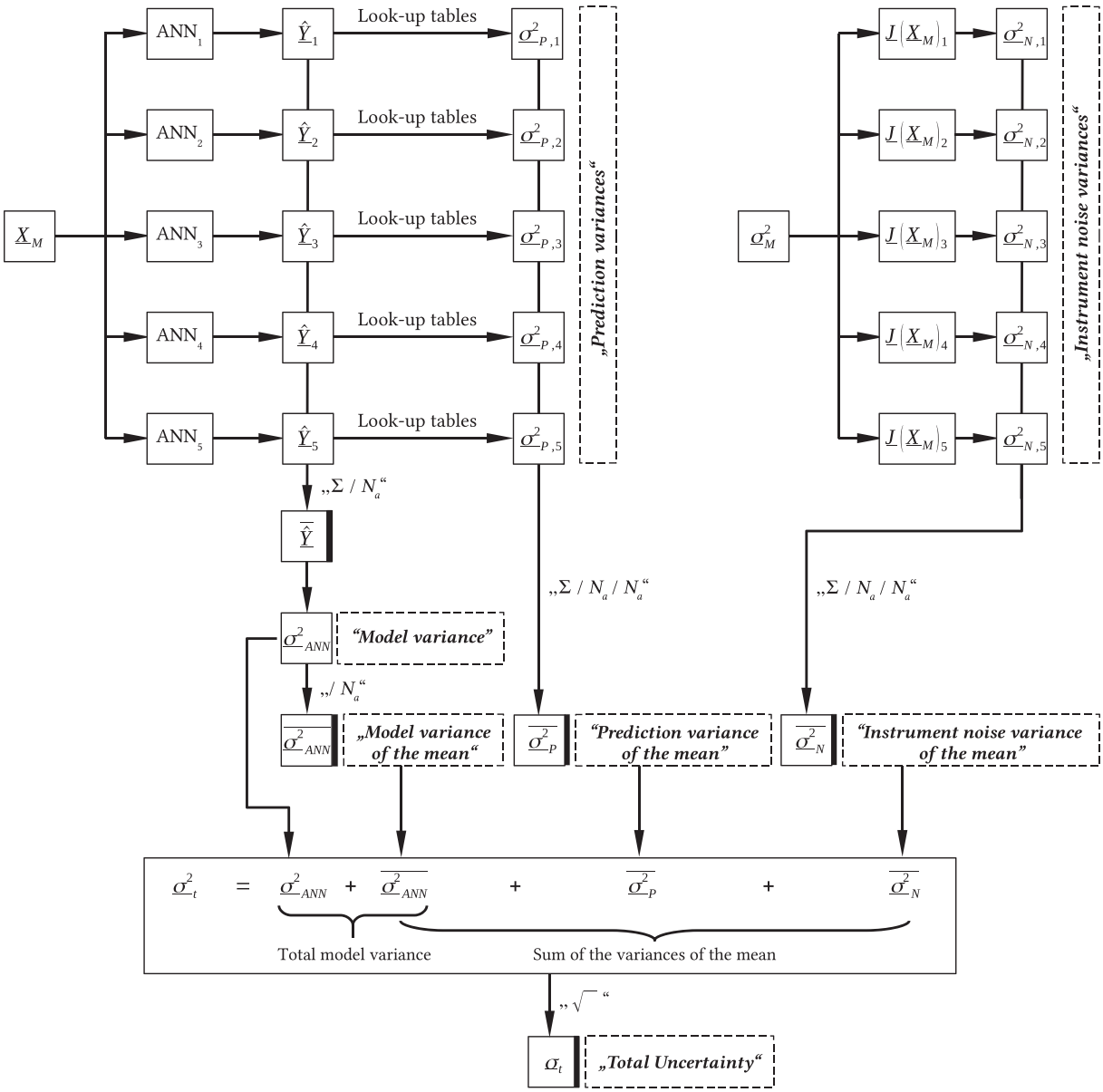

Fig. 2. Components of the inversion variances contributing to the total uncertainty. Final products stored in the ANN output file are marked-up by the dashed boxes.

Fig. 3. ANN application.

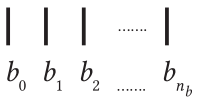

Fig. 4. Histogram sketch.

A network with 'N_o' output dimensions needs 'N_o' histograms each having 'j=(n_b + 1)' bins. Those bins in turn contain 'hit' counters ($sec_{a,o,j}$) as well as sums of error squares ($se_{a,o,j}$) which are both set to zero prior to the analysis.

Table 2
Histogram parameters to estimate the prediction variance of the mean.

Compound	b_{min}	b_{max}	Δb	n_b
$\log_{10}(R_{rs}(\lambda))$	-9.00	+1.50	0.01	1050
$AOT(\lambda)$	0.00	1.10	0.01	110

After allocating the histogram all N_V values $\underline{X}_V$ contained in the validation data set are presented sequentially to a trained ANN_a. Each response $\underline{\hat{Y}}_a^{(n)}$ is compared against the desired result $\underline{Y}_V^{(n)}$ by calculating the squared difference $se_{a,o}$ for each dimension separately. The squared differences for each dimension are summed to the $se_{a,o,j\mathrm{bin}}$ of the correct bin whose index j_{bin} is identified by checking a double inequality (Eq. (9)). In addition the bin hit counter, $sec_{a,o,j\mathrm{bin}}$ is incremented:

$$b_j - \Delta b/2 \le \hat{Y}_{a,o}^{(n)} < b_j + \Delta b/2$$
$$\text{with } j = 0, \ldots, n_b \Rightarrow j_{bin}$$

$$\Rightarrow se_{a,o,j_{\text{bin}}} += \left(\widehat{Y}_{a,o}^{(n)} - Y_{V,o}^{(n)}\right)^2 \text{ and } sec_{a,o,j_{\text{bin}}} += 1 \tag{9}$$

After looping over all data points N_V the $se_{a,o,j}$ values are normalized by $sec_{a,o,j}$ (Eq. (10)) to obtain the final multi-dimensional tables $\left\{\underline{\widehat{Y}}_a / \underline{\sigma^2_{P,a}}\right\}$ for each ANN$_a$ (a=1,...,5).

$$mse_{a,o,j} = se_{a,o,j} \Big/ sec_{a,o,j} \overset{!}{=} \sigma^2_{P,a,o,j} \text{ if } sec_{a,o,j} \neq 0 \tag{10}$$

With the help of those LUTs we estimate an associated prediction variance $\underline{\sigma^2_{P,a}}$ for each response $\underline{\hat{Y}_a}$ of network 'a' by cubic spline interpolation from the pre-calculated LUTs whose generation requires this effort just once (Fig. 2).

The next step is, as in the preceding section, to calculate an average of all ANN responses by taking their - in this case - individual uncertainties into account. The strict utilization of the uncertainty propagation mechanism leads to the final value of the prediction variance of the mean $\overline{\underline{\sigma^2_{P,a}}}$ that is given by Eq. (11).

$$\overline{\underline{\sigma^2_P}} = \sum_{a=1}^{N_a} \underline{\sigma^2_{P,a}} \Big/ N_a \Big/ N_a \tag{11}$$

2.3.3. Instrument noise variance

The OLCI instrument consists of five 'identical' optical camera blocks (EUMETSAT, 2018). The push-broom field of view of those blocks are optically aligned across-track to form a seamless scan-line. The camera sensors are charge-coupled devices (CCD) and their signals are converted to digital signals by 14-bit analog to digital converters. This processing chain introduces noise from different sources to the measured signal, which is characterized by the signal-to-noise ratio (SNR) values that depend on the camera as well as the wavelength. This measure can be used to add an uncertainty term to the measured signals, which is then transferred through the inversion process (here: ANNs) to form a contribution to the output uncertainty that is due to the instrument noise variance. Camera-averaged SNR values from radiometric calibration data for each OLCI channel/wavelength SNR$_i$ have been tabulated by EUMETSAT (2018) (Table 3). Assuming the sensor noise is dominated by shot noise, which is usually proportional to the square root of the radiance signal, and with additional information of the reference calibration radiance level L_{ref}, these SNR can be scaled to values at observation level L_{obs} using a multiplicative factor of $\sqrt{L_{\text{obs}}/L_{\text{ref}}}$. In this study, for simplicity, we assumed that the SNR is invariant of the TOA radiance signal and used the unscaled values tabulated in Table 3.

From the general definition of the SNR it follows that the wavelength dependent SNR$_i$ is the inverse of the relative signal error err_i. The multiplication of the relative errors err_i with the measured signals $X_{M,i}$ (input dimensions i=1,...,12, see Eq. 3) directly yields signal uncertainties $\sigma_{M,i}$ and the square of those values form the diagonal of a covariance matrix by leaving the co-variances zero (Eq. (12)). With the help of the network's Jacobian $\underline{\boldsymbol{J}}$ matrix (Eq. (13)) which compiles the sensitivities of each output 'o' with respect to each input 'i' ('ΔOutput / ΔInput') the ANN input uncertainty can be propagated to the ANN output (Eq. (14)) (Arras, 1998).

$$\begin{aligned} &err_i = 1/\text{SNR}_i \Rightarrow \Delta X_{M,i} = \sigma_{M,i} = err_i \cdot X_{M,i} \\ &\Rightarrow \sigma^2{}_{M,i} = \Delta X^2{}_{M,i} \qquad \text{for } i = 1, \ldots, 12 \\ &\Rightarrow \sigma^2{}_{M,i} = 0 \qquad\quad \text{for } i = 13, \ldots, 18 \end{aligned} \tag{12}$$

$$\text{with } \underline{\boldsymbol{J}}\left(\underline{\boldsymbol{X}}_M^{(n)}\right) = \begin{pmatrix} \partial Y_1^{(n)} / \partial X_{M,1}^{(n)} & \cdots & \cdots & \cdots & \partial Y_1^{(n)} / \partial X_{M,i}^{(n)} \\ \vdots & \ddots & & \iddots & \vdots \\ \vdots & & \partial Y_o^{(n)} / \partial X_{M,i}^{(n)} & & \vdots \\ \vdots & \iddots & & \ddots & \vdots \\ \partial Y_o^{(n)} / \partial X_{M,1}^{(n)} & \cdots & \cdots & \cdots & \partial Y_O^{(n)} / \partial X_{M,I}^{(n)} \end{pmatrix} \tag{13}$$

$$diag\left(\sigma^2_{N,1}, \sigma^2_{N,2}, \ldots, \sigma^2_{N,O}\right) = \underline{\boldsymbol{J}}\left(\underline{\boldsymbol{X}}_M^{(n)}\right) \cdot diag\left(\sigma^2_{M,1}, \sigma^2_{M,2}, \ldots, \sigma^2_{M,I}\right) \cdot \underline{\boldsymbol{J}}^t\left(\underline{\boldsymbol{X}}_M^{(n)}\right) \tag{14}$$

In parallel to the prediction variances the square root of the components of the vector of instrument noise variances $\underline{\sigma^2_{N,a}}$ are the contributions to the total uncertainty of this component for the given effect. The next step, as in the two preceding sections, is to calculate an average of all ANN responses by taking their individual uncertainties into account. The strict utilization of the uncertainty propagation mechanism leads to the value of the instrument noise variance of the mean $\overline{\underline{\sigma^2_N}}$ that is given by Eq. (15).

$$\overline{\underline{\sigma^2_N}} = \sum_{a=1}^{N_a} \underline{\sigma^2_{N,a}} \Big/ N_a \Big/ N_a \tag{15}$$

The numerical construction of the network's Jacobian $\underline{\boldsymbol{J}}$ from approximations by centered finite difference derivatives makes the calculation of this term (see Fig. 2, right panel) the most time-consuming step. This term usually contributes only very small values to the total variance and its calculation can thus be neglected in standard analyses.

2.3.4. Total model variance and total uncertainty

Finally, all mean variance components (Eqs. 7, 11, 15) are added to yield the total variance of the mean. As the mean variance components are divided by the number of networks N_a, this sum (in theory) tends to vanish for large N_a ($N_a \rightarrow \infty$) it is necessary to add a non-vanishing variance term for a retrieval of the networks ensemble (see Fig. 2, bottom panel). As this component is not known *a priori* we decided to add the model variance $\underline{\sigma^2_{ANN}}$ as such a term, because it best reflects the variance remaining as a lower limit even for large N_a. This idea follows the general procedures applied in other regression models (Draper and Smith, 1998; Penny and Roberts, 1998; Hedderich and Sachs, 2016). The model variance is calculated from the model variance of the mean $\overline{\underline{\sigma^2_{ANN}}}$ by using Eq. (7). The total model variance $\underline{\sigma^2_{t,ANN}}$ is the sum of model variance and the model variance of the mean (Eq. 16).

$$\underline{\sigma^2_{t,ANN}} = \underline{\sigma^2_{ANN}} + \overline{\underline{\sigma^2_{ANN}}} \tag{16}$$

The value $\sqrt{\underline{\sigma^2_t}}$ is what we define as the total uncertainty $\underline{\sigma_t}$ (Eq. 17).

$$\underline{\sigma_t} = \sqrt{\underbrace{\underline{\sigma^2_{ANN}}}_{\text{Model variance}} + \underbrace{\overline{\underline{\sigma^2_{ANN}}} + \overline{\underline{\sigma^2_P}} + \overline{\underline{\sigma^2_N}}}_{\text{Sum of the variances of the mean}}} \tag{17}$$

Table 3
Camera averaged SNR values for used OLCI channels (EUMETSAT, 2018).

Band	2	3	4	5	6	7	8	11	12	16	17	18
λ center (nm)	412.5	442.5	490	510	560	620	665	708.75	753.75	778.75	865	885
$\overline{\text{SNR}_i}$	2388	2182	2000	1985	1798	1607	1553	1423	1128	1513	1238	819

2.4. Linearisation of outputs

The spectral remote sensing reflectance $R_{rs}(\lambda)$ are handled as (decimal) logarithmic values. The non-linear transformation is motivated by the large dynamic range of those values and the ability to avoid non-physical negative values. As a consequence the corresponding variances are also logarithmic values. To meet the end user's need for easy data interpretation we transformed all logarithmic values into linear values which are stored in the output file. The linearisation of the values themselves is straightforward while the transformation of the associated variances needs to be further explained:

Using a second order Taylor polynomial for $f(x)$ (Eq. 18) e.g. the functional relationship $f(R_{rs}) = \log_{10}(R_{rs})$ and omitting the wavelength dependence leads to the desired transformation equation (Eq. 19).

$$\Delta^2 f(x) \approx \left(\frac{\partial f(x)}{\partial x}\right)^2 \cdot \Delta^2 x \tag{18}$$

$$\Delta^2 log_{10}(R_{rs}) = \sigma^2_{log_{10}(R_{rs})} = \left(\frac{\partial log_{10}(R_{rs})}{\partial R_{rs}}\right)^2 \cdot \Delta^2 R_{rs} = \left(-\frac{1}{ln(10)\cdot R_{rs}}\right)^2 \cdot \sigma^2_{R_{rs}}$$

$$\Rightarrow \sigma^2_{log_{10}(R_{rs})} = \frac{\sigma^2_{R_{rs}}}{(ln(10)\cdot R_{rs})^2} \tag{19}$$

The rearrangement of this equation together with the trivial linearisation equation of the basic property finally yields the two equations for the uncertainty linearisation:

$$R_{rs} = 10^{log_{10}(R_{rs})} \text{and } \sigma^2_{R_{rs}} = (ln(10)\cdot R_{rs})^2 \cdot \sigma^2_{log_{10}(R_{rs})} \tag{20}$$

2.5. Out-of-scope detection

Out-of-scope input and output data have to be masked, or flagged, to avoid extrapolation of the ANN ensemble. In this application we implemented an extended validation test (eVT), which is based on a combination of a simple convexity test (CT) and a validation test (VT) that uses a density map of the simulated data space (Fig. 5). The CT ensures that each input and output dimension is between the minimum and maximum of the corresponding dimension of the ANN training data. This assumption however, is rather coarse as it assumes a homogeneously filled data space. Often the data are not distributed homogeneously and thus a CT is an incomplete and unsatisfactory check. The additional VT provides an approximation of the data density $\gamma(x)$ by a kernel density estimator, which consists of a sum of one-dimensional normalized Gaussian kernels. The kernel centres are estimated from a vector quantization algorithm like the Neural Gas algorithm (Martinetz and Schulten, 1991), while the spreads of the centres are estimated from nearest neighbor heuristics. Finally a cut-off density $\gamma(x)_{limit}$ is established from the 1-5% percentile of the integrated training data's density histogram. A more detailed description and application to MERIS data is provided in Schaale and Schroeder (2013). Within the ensemble approach very conservative masking has been implemented, where a given input vector or pixel is masked as 'eVT failed' if that input fails either the CT or VT test for at least one of the ensemble ANNs.

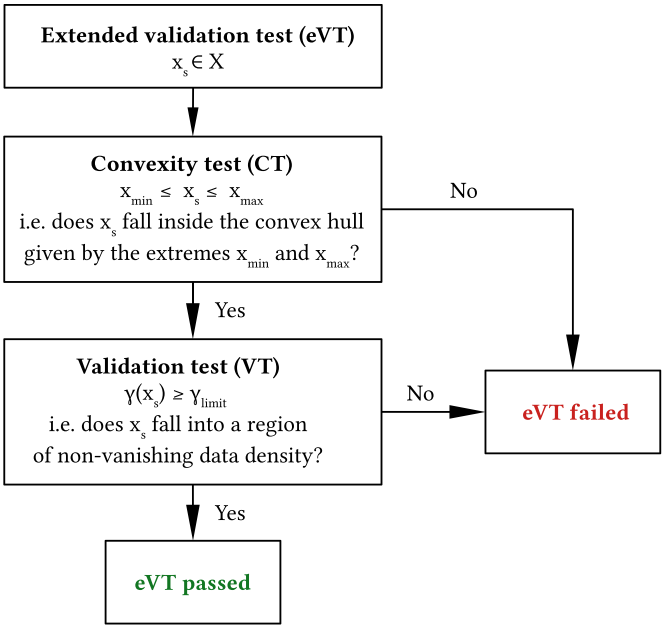

Fig. 5. Flow diagram of the extended validation test (eVT) to filter and flag out-of-scope data.

3. Data sets

3.1. In-situ radiometric data

Hyper-spectral radiometric data for validation were collected with the "Dynamic Above-water Radiance (L) and Irradiance (E) Collector" (DALEC) instrument developed by In-situ Marine Optics Pty Ltd, Australia. The instrument was deployed on an ad-hoc basis between 2016 and 2018 in fixed-mode at the Lucinda Jetty Coastal Observatory (located at 18.52°S, 146.39°E), and in transect-mode in collaboration with the Australian Institute of Marine Science (AIMS) from the RV Solander (Fig. 6).

The DALEC sensor head contains three compact hyper-spectral spectroradiometers (Carl Zeiss Monolithic Miniature Spectrometers), as well as a GPS unit, in addition to pitch and roll sensors, and is designed to be mounted on a boom positioned over water, typically off the ship's bow (Fig. 6c). Each spectroradiometer records 200 channels (400–1050 nm) with a spectral resolution of 10 nm spaced at ~3.3 nm intervals. The radiometer system is calibrated annually at the manufacturer against a NIST traceable standard.

The validation quantity of interest is the spectral remote sensing reflectance $R_{rs}(\lambda)$, which is defined by the ratio of the water-leaving radiance $L_W(\lambda)$ to the downwelling hemispherical irradiance $E_d(\lambda)$. The remote sensing reflectance can be calculated from simultaneous measurements of the total upwelling sea surface radiance $L_u(\lambda)$, the downwelling sky radiance $L_{sky}(\lambda)$ and the downwelling hemispherical irradiance:

$$R_{rs} = \frac{L_W}{E_d} = \frac{L_u - \xi L_{sky}}{E_d} - \epsilon \tag{21}$$

where ξ is the wind speed dependent air-sea interface reflection coefficient and ϵ the residual sunglint and skyglint term. The wavelength and sun and observing geometry dependence of these quantities have been omitted for convenience. The water-leaving radiance cannot be measured directly but can be calculated from the total upwelling radiance by removing the contributions of the diffuse sky light and residual sun-glint. To calculate R_{rs} with Eq. (21) from the simultaneous above water measurements, the appropriate value for ξ (depending on the sun and observing geometry and wind speed) was retrieved from tabulated values by Mobley (2015) using an interpolated look-up table approach. ϵ were estimated iteratively adjusting R_{rs} to the Ruddick et al. (2006) similarity spectrum in the 700–800 nm spectral range.

Maintenance of the DALEC instrument was performed on daily basis during the ship-borne deployments, while the fixed installation deployments at the Lucinda Jetty Coastal Observatory were maintained on fortnightly basis. Five minute interval high resolution hemispherical webcam imagery at Lucinda site allowed for additional quality control of the data.

A recent study by Antoine et al. (2021) found the DALEC above-water radiometric measurements to be in excellent agreement with more commonly used in-water radiometer systems such as the

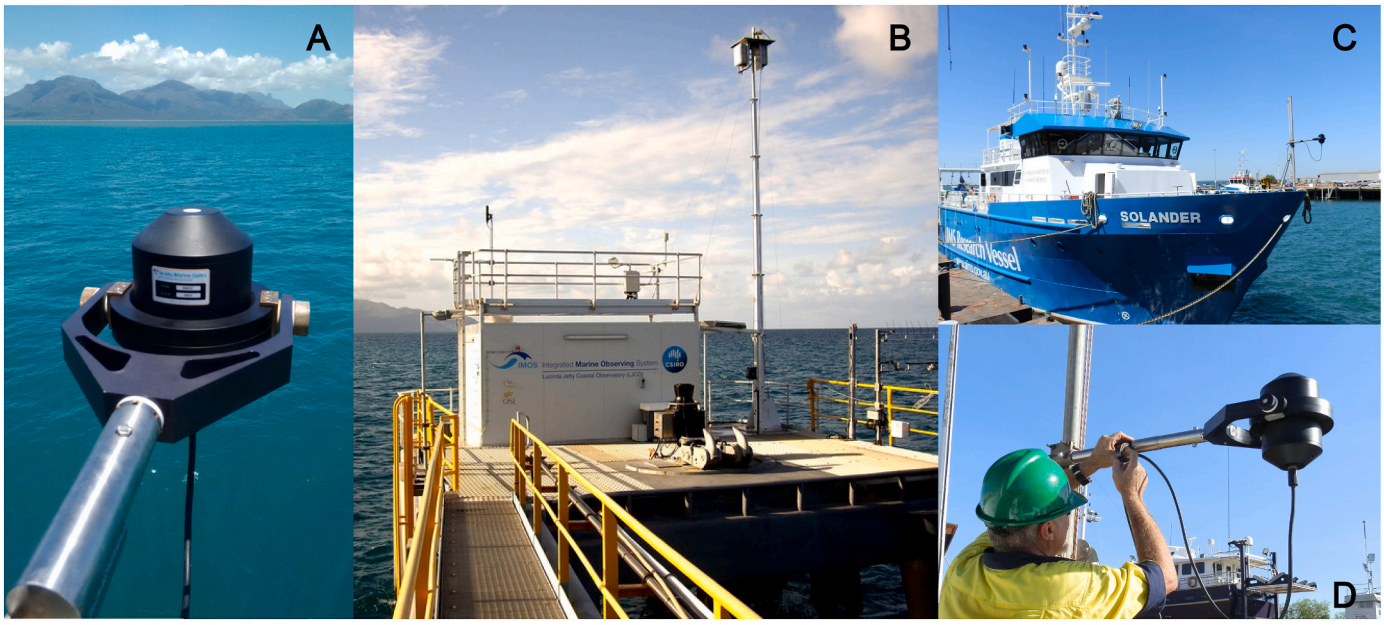

Fig. 6. DALEC hyper-spectral deployment setup at the Lucinda Jetty Coastal Observatory (A, B) and on the AIMS RV Solander (C, D). Image credit: CSIRO.

Biospherical C-OPS and the SeaBird HyperPro II. In oligotrophic waters the DALEC-derived R_{rs} <600 nm showed a dispersion of ~8% compared to the C-OPS and HyperPro systems with no bias and a dispersion of ~25–50% for wavelengths >600 nm as expected for open ocean waters due to the extremely low NIR signal.

The DALEC measurements used for validation in this study, publicly available from the Australian Ocean Data Network AODN (https://portal.aodn.org.au/), cover a wide range of coastal and open ocean waters and were convolved with the Sentinel-3 OLCI mean spectral response functions with nominal center bands listed in Table 3. A more detailed description of data processing protocol and quality control can be found in Brando et al. (2016).

3.2. Satellite data

A regional Sentinel-3A data archive covering the wider Austral-Asian region ([10°N,80°E]-[60°S,180°E]) has been built since mission start from daily FTP downloads of the EUMETSAT rolling Online Data Archive (ODA). The regional archive compromises of Level 1b and 2 reduced (RR) and full-resolution (FR) data in 1200 m and 300 m nadir resolution respectively. The Level 1b data provides calibrated top-of-atmosphere radiances for each OLCI channel, plus annotation data associated to OLCI pixels, while the Level 2 data provides derived geophysical products. The earlier data for the period April 2016 to November 2017 have been replaced with the most recent reprocessing of Processing Baseline 2.23 acquired from the Copernicus Online Data Access Reprocessing Repository (CODARep). All OLCI Level 1b data used in this study are in FR quality and associated with Processing Baseline 2.23 or higher. The comparison with the operational processor (OP-PROC) has been performed using the latest Level 2 ocean colour processing implemented by EUMETSAT on 16 Feb 2021 with processor version IPF-OL-2 v7.01. This data belongs to Processing Baseline Collection 3 (OL_L2M.003) and the main updates include a revised bright pixel atmospheric correction and updated system vicarious calibration (EUMETSAT, 2021). This processor is currently deployed for a full mission reprocessing of OLCI Level 2 Sentinel-3 A/B scheduled to be completed during the second half of 2021. A detailed overview of the evolution of the OLCI Processing Baseline and associated changes can be found online (Sentinel-3 OLCI Processing Baseline, n.d.).

4. Results and discussion

4.1. Match-up comparison and validation

The performance of the ANN atmospheric correction algorithm was quantified by match-up analysis, extracting 3 x 3 satellite pixels from the processed Sentinel-3A FR images at the date and location of the in-situ measurements and comparing the median of the pixel reflectances with the in-situ reflectance spectra within a maximum time window (ΔT) of ±30 min to the satellite overpasses. Such a short time difference was selected to minimize reflectance variability due to tidal effects at the Lucinda Jetty Coastal Observatory. Further, the match-up location of the Lucinda site was shifted by two FR pixels across-track and along-track to the north and east to avoid platform perturbations in the high resolution data. Match-ups were retrieved using spherical geometry by calculating the shortest distance between in-situ and satellite recorded pixel coordinates on the surface of the Earth (great-circle distance). A valid match-up required the location difference to be less than 0.004 degree in both latitude and longitude dimensions. This maximum distance accuracy criterion is at pixel-level given the 300 m spatial resolution of the OLCI FR ocean colour data. The geolocation accuracy of Sentinel-3A data itself is at sub-pixel level, with a RMSE of 0.1 pixel (EUMETSAT, 2019). For our application a valid ANN match-up required all pixels to be unflagged that were evaluated by analyzing the Level 1 LAND, COASTLINE, BRIGHT, SUN GLINT RISK, INVALID and DUBIOUS quality flags, in addition to the ANN out-of-scope flags (See Sec. 2.5). For comparison with the operational processor (OP-PROC) the corresponding Level 2 water-leaving reflectance products were extracted applying the same Level 1 flags in addition to the Level 2 AC-FAIL flag.

Fig. 7 shows all in-situ DALEC spectra (N=87) matching unflagged Sentinel-3A data within ±30 min. The spectra acquired at the Lucinda Jetty Coastal Observatory are mainly influenced by non-algal particulates and CDOM resulting in greenish turbid waters, while the ship-borne transect data also covers clear open ocean blue waters. Due to

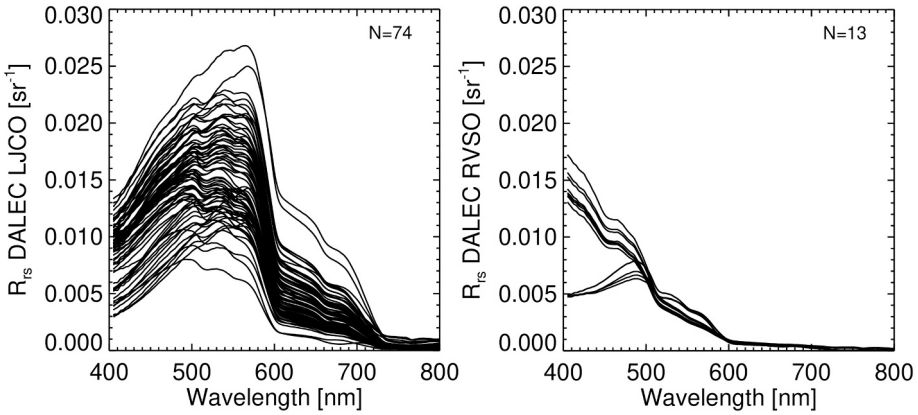

Fig. 7. Spectral variability of all DALEC above-water R_{rs} observations within ±30 min to Sentinel-3A. The left panel displays the observations collected at the Lucinda Jetty Coastal Observatory (LJCO), while the right panel displays the ship-borne measurements acquired on the AIMS RV Solander (RVSO).

the strong tidal variability at the Lucinda site an additional quality control measure was applied to the satellite data to remove match-ups with a standard deviation of more than 40% of the median R_{rs}.

Due to the cloud masking limitations using the Level 1 flags an additional visual inspection of all true-color match-up macro regions (e.g. 300 x 300 pixels) was performed to filter images with scattered cloud and haze conditions. A total of 50 high quality match-ups were used for the validation exercise.

The algorithm performance through match-up analysis was evaluated based on four statistical values (Eqs. 22–25); the root mean squared error (RMSE), the mean absolute percentage error (MAPE), the bias, and the Pearson correlation coefficient (r), which were calculated on linear scale as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x-y)^2} \tag{22}$$

$$\mathrm{MAPE} = \frac{100}{N}\frac{\sum_{n=1}^{N}\left|(x-y)\right|}{y} \tag{23}$$

$$\mathrm{bias} = \frac{1}{N}\sum_{n=1}^{N}(x-y) \tag{24}$$

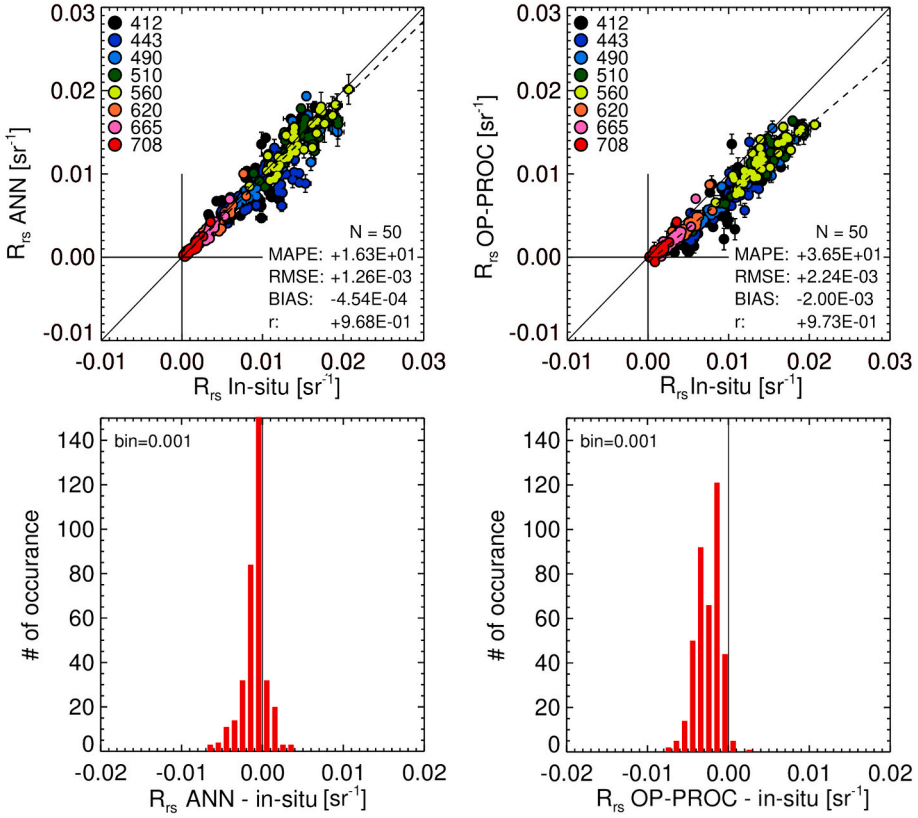

Fig. 8. Scatter plots of in-situ versus satellite estimated above-water remote sensing reflectance with the ANN (left) and OP-PROC (right) including histograms of the absolute difference for all data. The solid line is the 1:1 line while the dashed line is the linear regression line.

Table 4
Spectral and band-averaged error statistics of the match-up analysis.

Wavelength [nm]		412.50	442.50	490.00	510.00	560.00	620.00	665.00	708.75	412.50–708.75
MAPE [%]	ANN	19.5	15.9	10.1	9.1	9.1	19.3	20.4	27.3	16.3
	OP-PROC	40.4	30.4	23.7	24.1	24.4	41.0	47.8	60.4	36.5
RMSE [10^{-3}sr^{-1}]	ANN	2.13	2.46	1.71	1.39	1.09	0.62	0.38	0.29	1.26
	OP-PROC	3.66	3.19	2.99	2.79	2.54	1.21	0.92	0.60	2.24
bias [10^{-3}sr^{-1}]	ANN	-0.17	-0.13	-0.55	-0.51	-0.49	-0.27	-0.17	-0.19	-0.45
	OP-PROC	-3.14	-2.91	-2.82	-2.60	-2.23	-1.02	-0.77	-0.51	-2.00
r	ANN	0.83	0.70	0.91	0.96	0.98	0.96	0.97	0.96	0.97
	OP-PROC	0.84	0.90	0.97	0.98	0.98	0.95	0.93	0.90	0.97
N (-)	ANN	0	0	0	0	0	0	0	0	0
	OP-PROC	0	0	0	0	0	0	1	7	8
Slope	ANN	0.952	0.642	0.916	0.943	0.975	0.991	1.010	1.033	0.946
	OP-PROC	0.867	0.920	0.899	0.871	0.854	0.839	0.873	0.935	0.822
Offset [10^{-3}]	ANN	0.270	2.446	0.504	0.164	-0.223	-0.241	-0.195	-0.223	-0.039
	OP-PROC	-1.928	-2.081	-1.556	-1.060	-0.6123	-0.489	-0.501	-0.435	-0.626

$$r = \frac{\sum_{n=1}^{N}\left(x-\underline{x}\right)\left(y-\underline{y}\right)}{\sqrt{\sum_{n=1}^{N}\left(x-\underline{x}\right)^2\sum_{n=1}^{N}\left(y-\underline{y}\right)^2}} \quad (25)$$

where x=$R_{rs,sat}$ is the satellite observation and y=$R_{rs,in\text{-}situ}$ the in-situ measurement and n the number of match-ups. In addition slope and offset values were calculated to evaluate the quality of the linear regression.

Fig. 8 shows the scatter plots of the ANN and OP-PROC derived R_{rs} in comparison to the in-situ R_{rs}. The validation results demonstrate that the ANN shows improved performance compared to the OP-PROC with a band-averaged (412–708 nm) MAPE of only 16% compared to 37%. The ANN band-averaged absolute error with a RMSE of 0.0013 sr^{-1} is also significantly lower than the OP-PROC, which showed a RMSE of 0.0022 sr^{-1}. The band-averaged bias of the ANN of -0.00045 sr^{-1} is four times lower compared to the operational atmospheric correction and the absolute histogram differences, with respect to the in-situ data are better centred around the zero value with higher slope values of the linear regression. Table 4 provides the spectral and band-averaged error statistics of both algorithms including slope and offset values. Highest

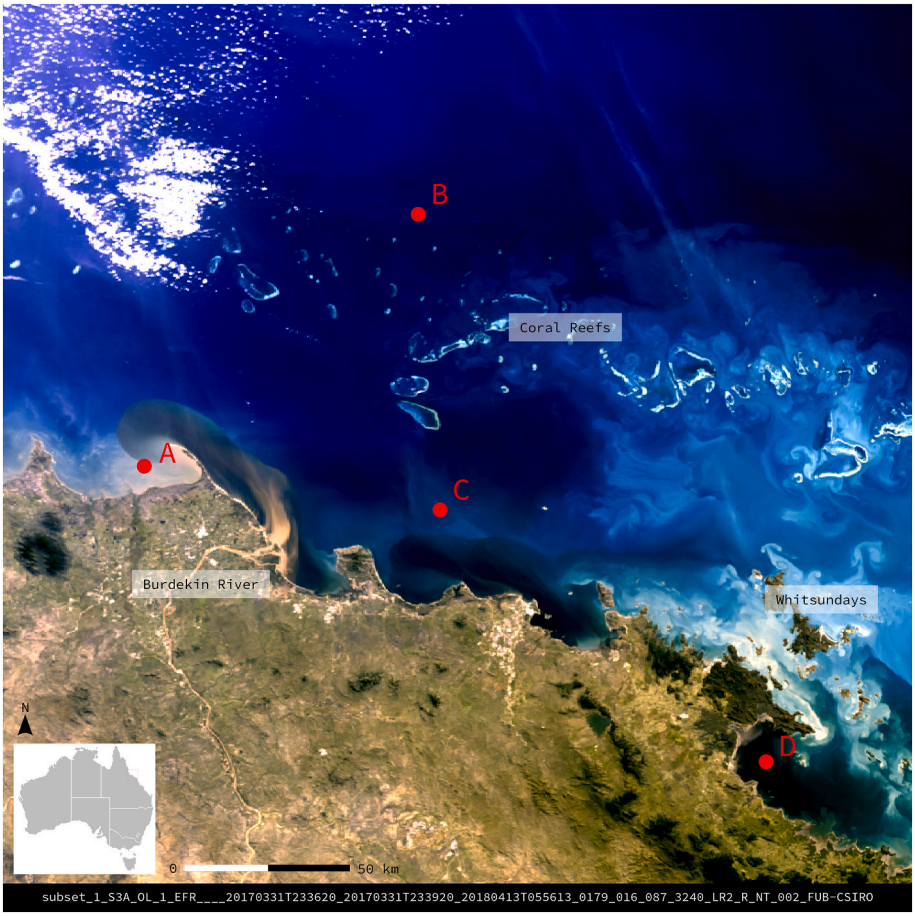

Fig. 9. Sentinel-3A FR TOA true colour composite of the optically complex test scene acquired over the Great Barrier Reef on 31 Mar 2017 with four selected locations (A–D) across different water types for spectral data extraction.

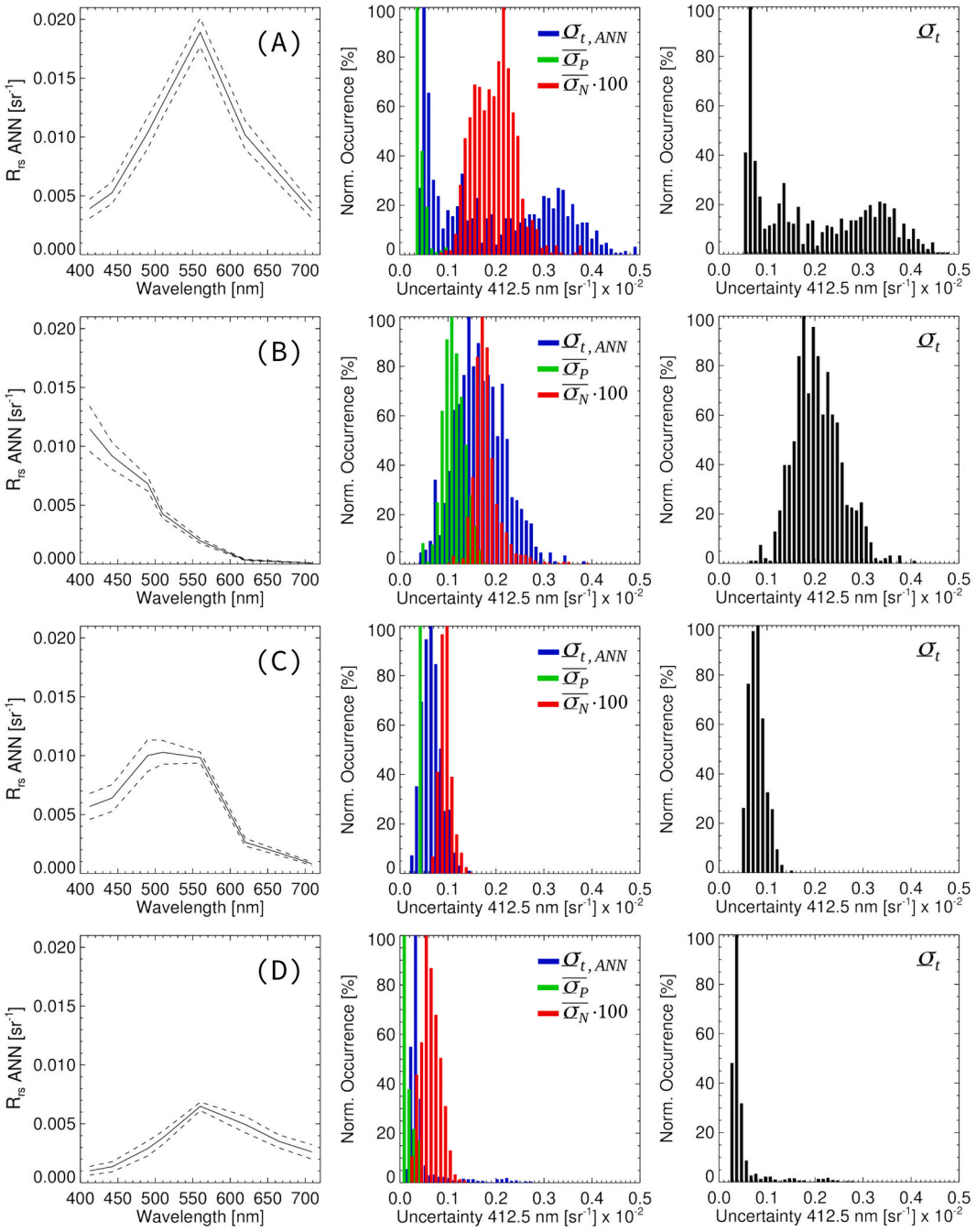

Fig. 10. Spectral uncertainties of four water types extracted at locations A-D in Fig. 9 (left column). (A) Highly turbid sediment-dominated, (B) offshore blue waters, (C) mixed-coastal phytoplankton dominated waters, and (D) high CDOM waters with corresponding uncertainty components at 412.5 nm; $\overline{\sigma_N}$ mean sensor noise uncertainty, $\overline{\sigma_P}$ mean prediction uncertainty and $\underline{\sigma_{t,ANN}}$ total model uncertainty (center column). The mean sensor noise uncertainty was multiplied by a factor 100 to allow plotting all uncertainty components on the same x-axis. $\underline{\sigma_t}$ denotes the total uncertainty (right column).

absolute errors are observed in the blue while highest percentage errors occur in the NIR with a general decrease of the bias from the blue to the NIR. The band-averaged Pearson correlation coefficient of 0.97 is similar for both algorithms, the operational processor however shows improved correlation of the first four bands compared to the ANN. Retrieval of negative reflectance values N(-) occur with the operational processor at 665 and 708.75 nm (Table 4).

The performance of the ANN algorithm was further evaluated using a Sentinel-3A FR test image acquired on 31 March 2017 over coastal waters of the central Great Barrier Reef, Australia (Fig. 9). The image covers the most complex optical conditions that can occur in these waters within one image. Water types range from extremely turbid sediment-dominated flood waters, to highly-absorbing CDOM-dominated waters, mixed productive waters and clear blue open ocean waters.

Fig. 10 shows ANN reflectance spectra and derived uncertainties at four selected locations corresponding to these water types (A=sediment-dominated, B=clear, blue waters, C=mixed-coastal waters and D=high CDOM waters). The total uncertainty (Eq. 17) provides reasonable uncertainty envelopes for all reflectance spectra, which in the presented case are slightly larger at shorter wavelengths for the mixed-productive and blue waters. Four 15 x 15 pixel boxes were extracted at these locations to analyze the different uncertainty components at 412.5 nm, which are presented as histograms in Fig. 10. The uncertainty components were computed as the total model uncertainty $\underline{\sigma}_{t,ANN}$ (Eq. 26), the mean prediction uncertainty $\overline{\sigma_P}$ (Eq. 27) and the mean instrument noise uncertainty $\overline{\sigma_N}$ (Eq. 28) from the individual variances of Eq. 17. The total uncertainty $\underline{\sigma}_t$ is the sum of the individual uncertainty components (Eq. 29).

$$\underline{\sigma}_{t,ANN} = \sqrt{\underbrace{\underline{\sigma^2}_{ANN}}_{\text{Model variance}} + \underbrace{\overline{\underline{\sigma^2}_{ANN}}}_{\text{Variance of the mean}}} \tag{26}$$

$$\overline{\sigma_P} = \sqrt{\underbrace{\overline{\sigma^2_P}}_{\text{Variance of the mean}}} \tag{27}$$

$$\overline{\sigma_N} = \sqrt{\underbrace{\overline{\sigma^2_N}}_{\text{Variance of the mean}}} \tag{28}$$

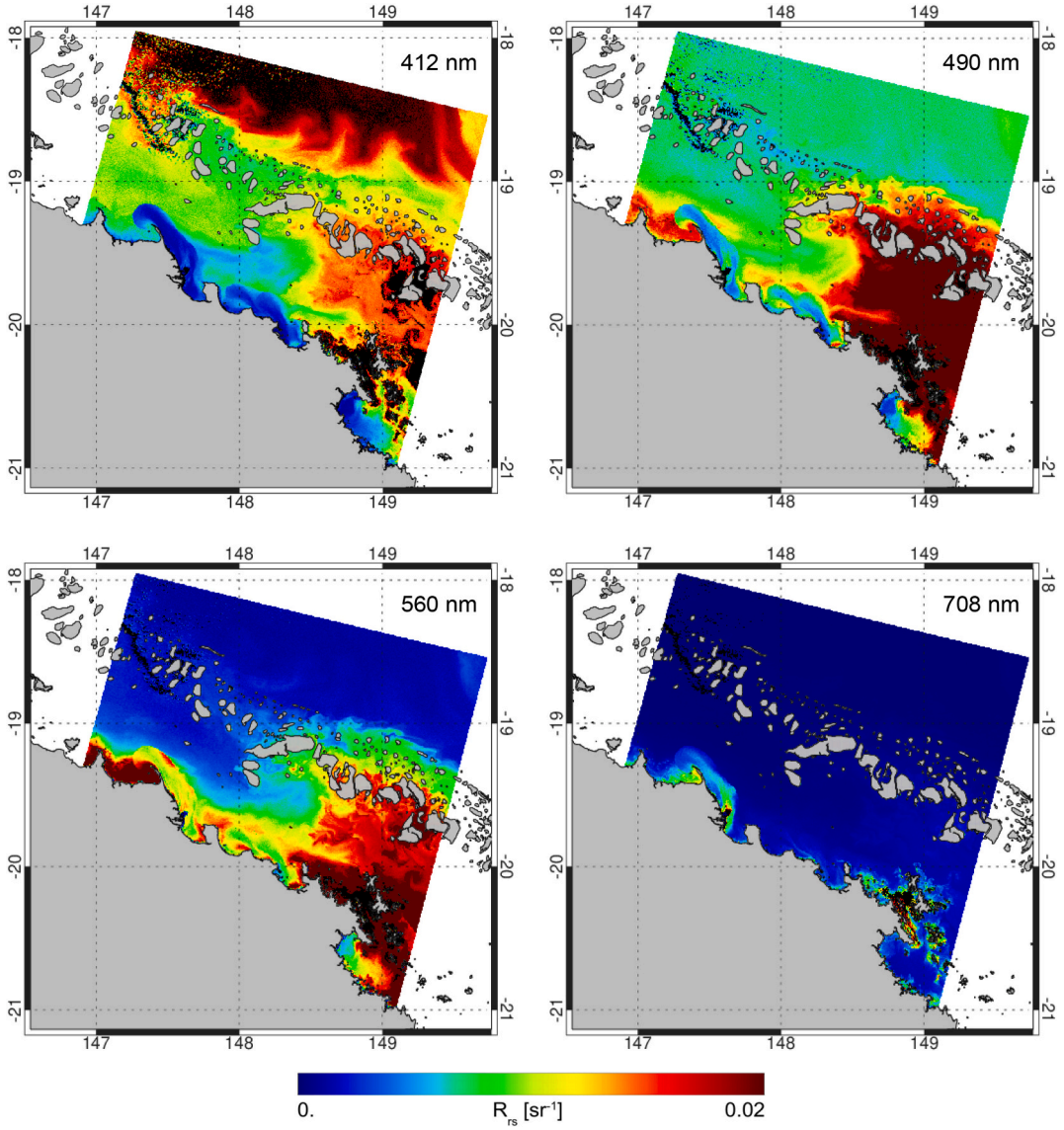

Fig. 11. ANN-derived remote sensing reflectance at four spectral bands. Land and reefs are masked in grey. Algorithm out-of-range conditions and masked clouds (top left corner) are shown in black.

$$\underline{\sigma_t} = \underline{\sigma_{t,ANN}} + \overline{\sigma_P} + \overline{\sigma_N} \quad (29)$$

The uncertainties associated with sensor noise variance (red histogram components) are two orders of magnitude lower than the uncertainty components of the prediction (green) and total model variances (blue), which justifies the assumption made in this study that the sensor SNR is invariant of the TOA radiance signal (Sec. 2.3.3). For standard applications the sensor noise uncertainty component can therefore be neglected, giving a computational saving in image processing time of a factor of ~20. In this application we did not characterize or include the TOA radiance covariance terms (off-diagonal terms in Eq 12), which have been shown to have only a marginal impact on the retrieved uncertainty as the ocean colour product uncertainties are mainly driven by errors in modeling and not the remote sensing sensor itself as demonstrated by ACRI-ST (2013) for the uncertainty propagation in the OLCI clear water branch and applied to MERIS. Further ACRI-ST (2013) concluded that not accounting of the TOA radiance covariance terms will lead to an overestimation of the retrieved marine reflectance uncertainty. In that respect our approach provides a very conservative estimate of the sensor noise uncertainty.

The second largest uncertainty component in our framework is the model prediction uncertainty. This measure is a very conservative uncertainty estimate as the LUT interpolation for a given reflectance value (bin) includes all possible sun and observing geometry as well as marine and atmospheric conditions that were considered in the RT simulations. The overall largest uncertainty component is the total model inversion uncertainty from averaging the responses of the slightly different adapted networks of the ensemble.

4.2. Spatial analysis of R_{rs} and uncertainty products

The inversion of the test scene (Fig. 9) illustrates the spatial performance of the ANN algorithm at four reflectance bands from the blue to the NIR as presented in Fig. 11. Overall the derived spatial reflectance patterns are smooth with no inherent sensor noise visible. The spatial distribution of these patterns are as expected showing high reflectance

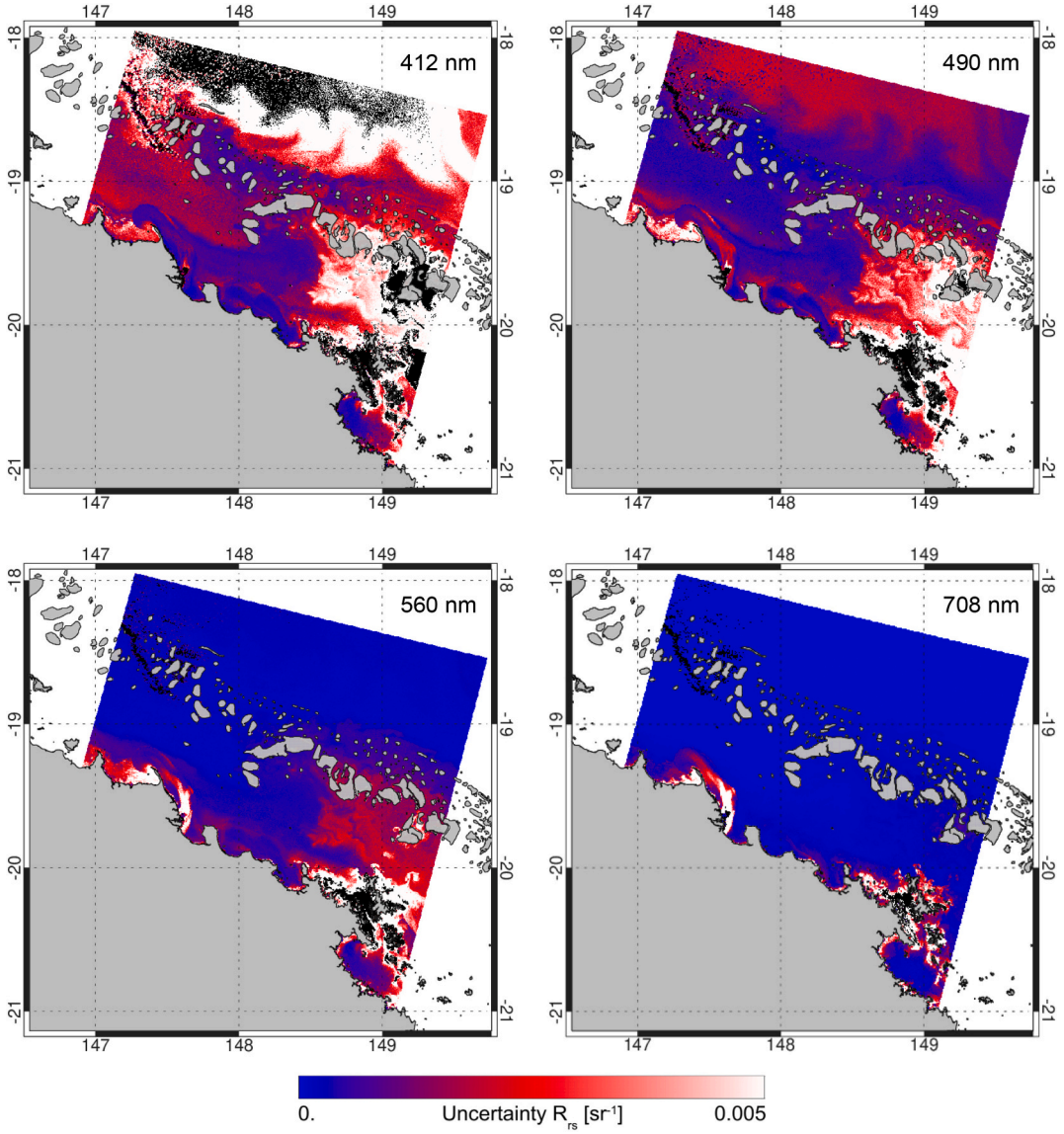

Fig. 12. ANN-derived absolute uncertainty at four spectral bands corresponding to Fig. 11. Land and reefs are masked in grey. Algorithm out-of-range conditions and masked clouds (top left corner) are shown in black.

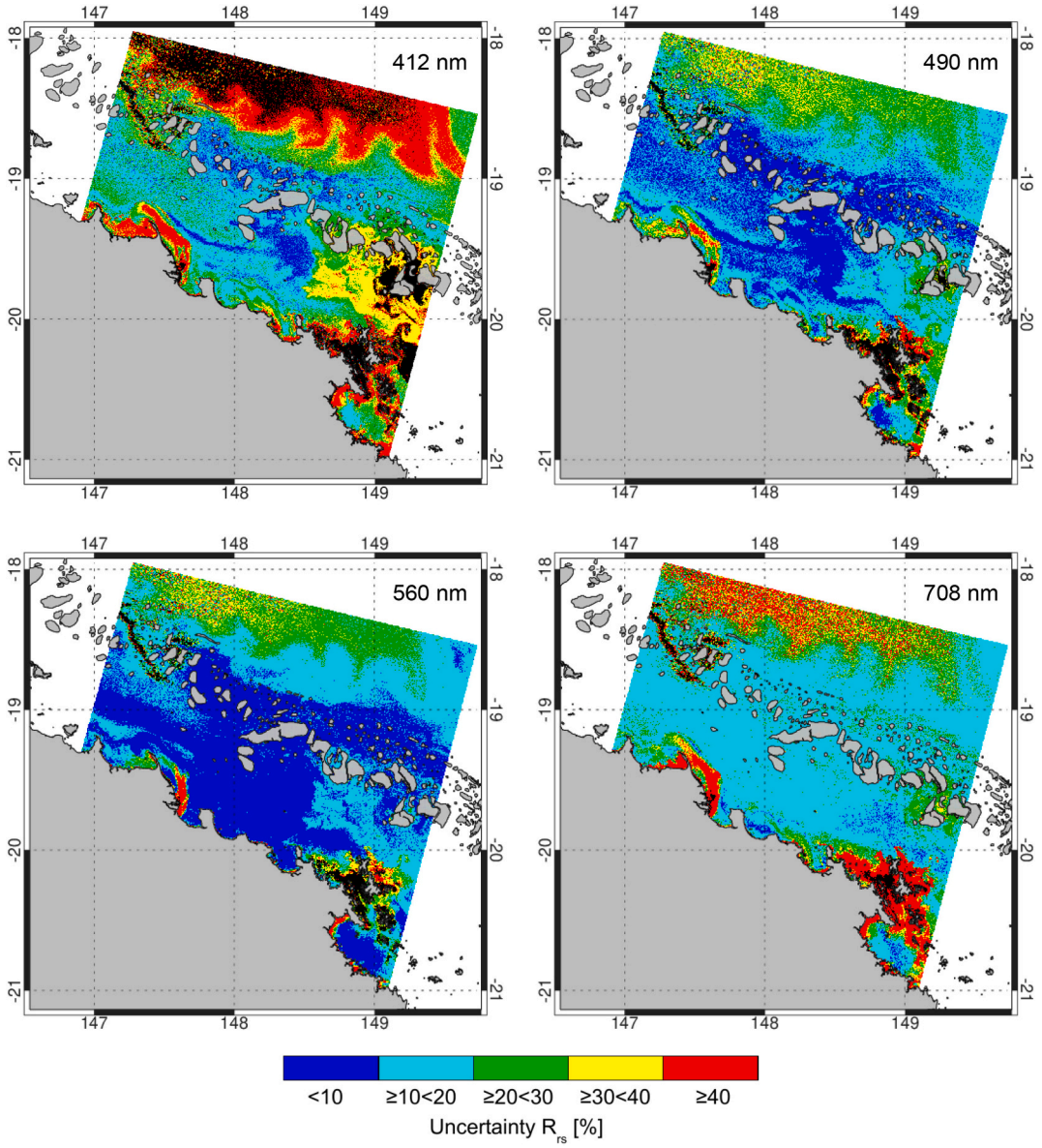

Fig. 13. ANN-derived relative uncertainty at four spectral bands corresponding to Fig. 11. Land and reefs masked in grey. Algorithm out-of-range conditions and masked clouds (top left corner) are shown in black.

values in the blue (412 nm) for the clear off-shore waters and low values along the coastal region due to the absorption of particulate and dissolved components. At 560 and 708 nm, areas with high reflectance correspond to high sediment concentrations. A few out-of-range conditions were detected in the blue spectral region at 412 nm for the clear off-shore waters, the near-shore Burdekin flood plume and the very bright sandy and optically shallow areas of the Whitsundays as indicated by the black color-coded pixels. The derived reflectance in these areas exceeded the maximum reflectance values of the forward RT simulations and therefore resulted in out-of-range conditions. The maximum water quality concentrations considered in the RT simulations were TSS = 50 g m^{-3}, chlorophyll-a = 50 mg m^{-3} and CDOM absorption at 443 nm = 1 m^{-1} (Schroeder et al., 2007a).

The sediment-rich river discharge of the Burdekin River likely exceeds the simulated TSS maximum of 50 g m^{-3}. The corresponding absolute uncertainty maps are shown in Fig. 12, which are highly correlated with the reflectance products. A higher reflectance is usually associated with a higher absolute uncertainty and vise versa. At 412 nm for example we observe high absolute uncertainties of >0.005 sr^{-1}, which are associated with the clear offshore waters. Reduced performance in blue waters is expected as the ANN algorithm was trained predominantly with simulations of coastal waters.

The uncertainty maps therefore provide a useful tool to guide future algorithm improvements such as expanding simulations for blue open ocean waters or extending the current TSS maximum of 50 g m^{-3}. Fig. 13 provides the corresponding spatial distribution of the standard relative uncertainties expressed as $\left(\sigma_t/R_{rs}\right) \times 100$, which are not necessarily correlated with the magnitude of the reflectance. Higher percentage uncertainties >40% are not only observed for the clear offshore waters

at 412 nm but also across all bands within the sediment-rich river discharge of the Burdekin River and the bright, sandy and optically shallow areas of the Whitsundays. For the majority of the selected image however, the retrieved uncertainties are below 20%.

A comparison with the uncertainties provided by the operational processor could not be performed as their use is still not recommended by EUMETSAT due to lack of verification. Systematic errors associated with uncertainties of sensor calibration (bias) were not considered in this study. These are assumed to be minimized for the operational processor by the system vicarious calibration of OLCI.

4.3. Additional sources of uncertainty

The identification of all sources of uncertainties may not be feasible and need to be balanced with the processing requirements in order to keep the inversion algorithm computationally-efficient. In this study, we focused on the most obvious random uncertainty sources from TOA sensor noise and the AC algorithm itself. Uncertainties associated with the ancillary ANN input data such as the barometric pressure, wind speed or the ozone concentration were not considered in our current uncertainty framework. Additional uncertainties may arise from the limitations of the forward model and bio-optical assumptions therein, specifically physical effects that were not considered in the forward RT simulations such as shallow waters, specific algal blooms, whitecaps, inelastic scattering or extreme in-water conditions where the water quality concentrations exceed the simulation ranges. Certain types of aerosols such as absorbing ones or polarization were also not considered in the forward modeling and may lead to larger uncertainties. It is therefore critical to apply high quality out-of-scope detection to keep the algorithm operating within its designed specifications. An advanced method for out-of-scope detection that uses a density map of the simulated data space was presented in section 2.5. We also acknowledge that the presented inversion process is an ill-posed problem which, as any remote sensing inversion, may lead to equivocalities or ambiguities. Work-arounds exist but are beyond practicability (Ardizzone et al., 2019).

4.4. Validation of uncertainty

Not only the remote sensing products but also the associated uncertainty products require validation, which is not a trivial task and requires careful consideration of the suitability of the in-situ observations. The in-situ data will have different sources of uncertainty and a direct comparison with the remotely sensed uncertainties would be in most cases meaningless. The authors of the IOCCG Report 18 on Uncertainty Estimates also concluded on this subject: "… *comparison need to be conducted between equivalent quantities (i.e., considering the same sources of uncertainty in their budget), which is not straightforward as the different approaches are based on different hypotheses and mathematical frameworks.*"

A possible approach for uncertainty verification has been outlined by Merchant et al. (2017) based on the analysis of the frequency distribution of the following ratio:

$$\frac{x_{sat} - x_{ref}}{\sqrt{\sigma_{sat}^2 + \sigma_{ref}^2 + \sigma_{geo}^2}} \tag{30}$$

where x_{sat} is the product value derived from satellite data and x_{ref} the reference (in-situ) data, divided by the square root of the sum of

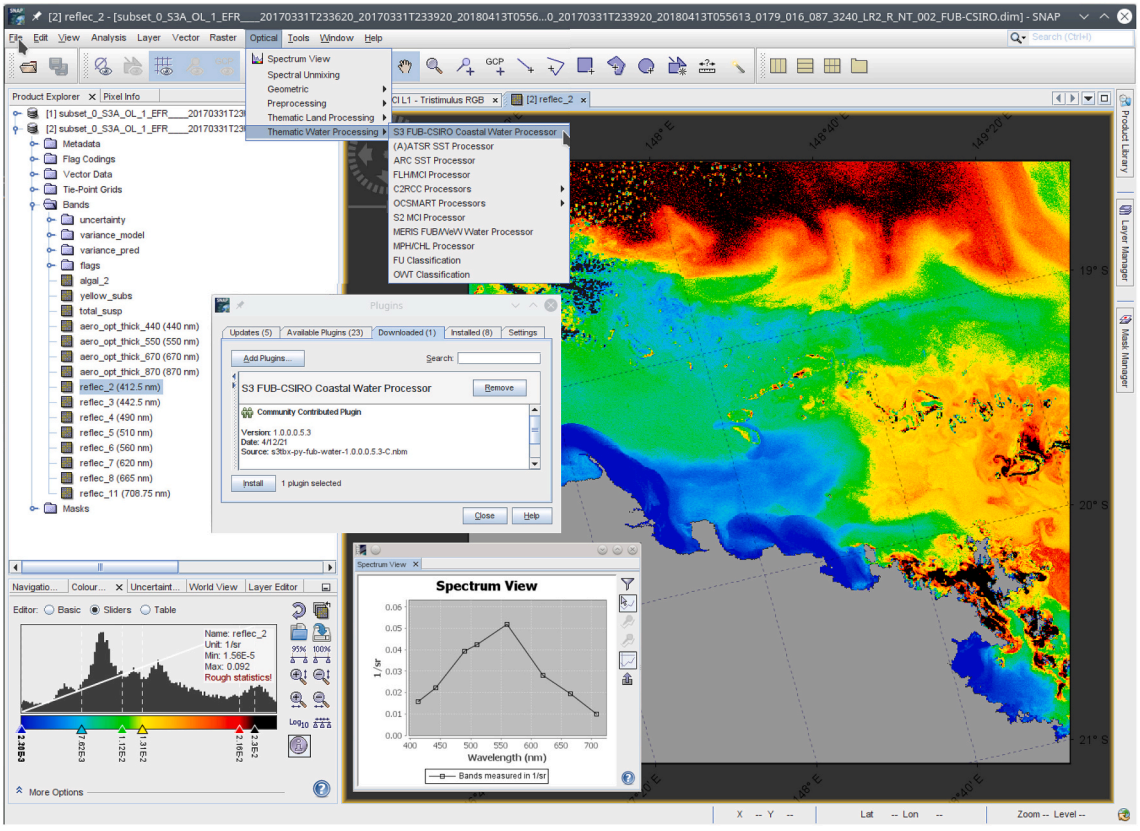

Fig. 14. Ensemble ANN atmospheric correction algorithm running under SNAP (version 8) showing remote sensing reflectance at 412.5 nm for selected test image (see Fig. 9).

variances of the satellite product σ^2_{sat}, the in-situ observations σ^2_{ref} and the variances of the geophysical variability σ^2_{geo} associated with the temporal, spatial and definitional mis-match of the satellite and in-situ data. If the uncertainties and the geophysical variability are correctly defined the ratio will be normally-distributed with a standard deviation equal to unity (Merchant et al., 2017). However, this approach requires that the uncertainties of the in-situ data and geophysical variability are fully known. In our case σ^2_{ref} and σ^2_{geo} are currently not fully determined so that this approach could not be applied. Nevertheless it provides a promising tool for future validation of the satellite derived reflectance uncertainty products.

4.5. A plugin for SNAP

The described ensemble ANN algorithm including the advanced out-of-scope detection and uncertainty framework has been integrated into the S3 FUB-CSIRO Coastal Water Processor version 1.0.0.0.5.3, which is available as a Python/C plugin for the Sentinel Application Platform (SNAP). The plugin allows both batch processing of Sentinel-3 OLCI full or reduced resolution data through the SNAP command-line Graph Processing Tool (GPT) as well as single-image processing using the Graphical User Interface (GUI). The GUI also allows non-remote sensing experts to analyze and visualize Sentinel-3 data. A screenshot of the ANN processor in operation is shown in Fig. 14. The algorithm can be downloaded at: https://github.com/s3tbx-fub-csiro/s3tbx-fub-csiro.git.

5. Summary and conclusions

This manuscript presented the development of a new ensemble neural network atmospheric correction approach for Sentinel-3 OLCI that provides pixel-based estimation of the inherent model inversion uncertainty and sensor noise propagation. Using concurrent in-situ radiometric observations the algorithm was validated across a wide range of optical water types and compared with the operational processor (IPF-OL-2 version 7.01) currently deployed by EUMETSAT in the marine ground segment. In this comparison the ensemble ANN method showed advanced performance in comparison to the operational processor with a band-averaged (412–708 nm) mean absolute percentage error of 16% compared to 37% and a four-times lower band-averaged bias of -0.00045 sr^{-1}. We provided an analytic framework to propagate OLCI's SNR and estimated its contribution to the total reflectance uncertainty. The uncertainties associated with the sensor noise variance were found to be two orders of magnitude lower than the uncertainty components of the prediction and total model variances at 412.5 nm. As the sensor noise accounts for the smallest contribution to the total uncertainty its propagation can be neglected if time-efficient processing is of importance, which will reduce the computation time by a factor of ~20. The overall largest uncertainty component was found to be the total model uncertainty from averaging the responses of the five slightly different adapted networks of the ensemble. A comparison with the operational Level 2 OLCI uncertainty products will be conducted once these are approved for use. The derived uncertainty products provide a useful tool to guide future algorithm improvements, and allow further uncertainty propagation through in-water and bio-optical algorithms as well as data assimilation into BGCMs. The algorithm could be extended to other past, present and future ocean colour missions, which would require sensor-specific radiative transfer simulations and subsequent training and validation of the inverse models. We will consider exploiting the full spectral information of Sentinel-3 OLCI including the UV and SWIR bands in a future update of the algorithm including a validation of the uncertainty products.

Glossary

Abbreviations and symbols

AC	Atmospheric Correction
AERONET	Aerosol Robotic Network
AIMS	Australian Institute of Marine Science
ANN	Artificial Neural Network
BGCM	Biogeochemical Models
BOA	Bottom of the Atmosphere
CCD	Charge-Couple Device
CDOM	Coloured Dissolved Organic Matter
CODARep	Copernicus Online Data Access Reprocessing Repository
C-OPS	Compact-Optical Profiling System
CSIRO	Commonwealth Scientific and Industrial Research Organization
CT	Convexity Test
DALEC	Dynamic Above-water Radiance and Irradiance Collector
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
eVT	extended Validation Test
FUB	Freie Universität Berlin
ESA	European Space Agency
FR	Full Resolution
GPS	Global Positioning System
GPT	Graph Processing Tool
GUI	Graphical User Interface
IMOS	Integrated Marine Observing System
IOCCG	International Ocean Colour Coordinating Group
L-BFGS	Limited Memory Broyden-Fletcher-Goldfarb-Shanno
LUT	Look-Up Table
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
MERIS	Medium Resolution Imaging Spectrometer
MOMO	Matrix-Operator-MOdel
NCRIS	National Collaborative Research Infrastructure Strategy
NIR	Near-infrared
ODA	Online Data Archive
OLCI	Ocean and Land Colour Instrument
OP-PROC	Operational Processor
RMSE	Root Mean Squared Error
RR	Reduced Resolution
RT	Radiative Transfer
RV	Research Vessel
SNAP	Sentinel Application Platform
SNR	Signal to Noise Ratio
SWIR	Shortwave infra-red
TOA	Top of Atmosphere
TSS	Total Suspended Solids
UV	Ultra-violet
VT	Validation Test
AOT	Aerosol Optical Thickness
E	Irradiance
N	Size of population (=24,640,000 simulated spectra)
N_V	Size of validation sample (=100,000)
N_T	Size of training sample (=100,000)
N_a	Number of ensemble ANNs (=5)
L	Radiance
$\underline{\underline{J}}$	Jacobian matrix
p	Barometric pressure
R_{rs}	Remote sensing reflectance
S	Activation function
$\boldsymbol{W}$	Weight matrix
ws	Wind speed
$\underline{X}$	Input vector
$\underline{Y}$	Output vector
ϵ	Residual sunglint and skyglint term
λ	Wavelength
ρ	Reflectance
t	Diffuse transmittance
θ_0	Sun zenith
σ^2_{ANN}	Model variance
$\sigma^2_{t,ANN}$	Total model variance
$\overline{\sigma^2_{ANN}}$	Model variance of the mean
$\overline{\sigma^2_P}$	Prediction variance of the mean
$\overline{\sigma^2_N}$	Instrument noise variance of the mean
$\sigma_{t,ANN}$	Total model uncertainty
$\overline{\sigma_P}$	Mean prediction uncertainty

(continued on next page)

(*continued*)

$\overline{\sigma_N}$	Mean sensor noise uncertainty
σ_t or $\sqrt{\overline{\sigma^2_t}}$	Total uncertainty
ξ	Air-sea interface reflection coefficient
General notation	
$\square$	The box indicates a vector component or general scalar
$\overline{\square}$	Overline denotes an average
$\underline{\square}$	Underline denotes a vector
$\underline{\overline{\square}}$	Simultaneous over- and underline denotes an average vector
$\square_i$	Vector component related to the input
$\square_o$	Vector component related to the output
$\square_T$	Vector component related to the training sample
$\square_V$	Vector component related to the validation sample
$\square_P$	Vector component of the prediction variance
$\square_N$	Vector component of the noise variance
$\square_a$	Vector component belonging to ANN #a (= 1,…,N_a)
$\square^{(n)}$	Vector #n from a sample or population

Author credit statement

TS and MS conceptualized the study, developed the methodology including the software and prepared an advanced draft of the manuscript. MS implemented the algorithm as a plugin for SNAP. JL assisted with the DALEC processing and analysis including match-up extractions. DBP provided data processing and satellite archive management support. DBP and JL reviewed the advanced draft of the manuscript. All authors have read and agreed to the published version of the manuscript.

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.

Acknowledgments

Data was sourced from Australia's Integrated Marine Observing System (IMOS) – IMOS is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS). NCRIS (IMOS) and CSIRO are acknowledged for funding the Lucinda Jetty Coastal Observatory. Funding was partly provided by eReefs, which is a public-private collaborative project between CSIRO, the Australian Institute of Marine Science, the Bureau of Meteorology, the Great Barrier Reef Foundation and the Queensland Department of Environment and Science. We are grateful to the Australian Institute of Marine Science (AIMS) for supporting the DALEC deployments on the RV Solander. We acknowledge EUMETSAT for the provision of Sentinel-3 data, specifically David Dessailly and Ewa Kwiatkowska for the provision of the latest reprocessing. This study was supported by the EUMETSAT Copernicus Collaborative Exchange Program in form of a travel grant to Michael Schaale. We also wish to thank the ESA Sentinel Application Platform and Sentinel tool-box development teams for making the SNAP software freely available. Three anonymous reviewers are thanked for their valuable comments on the manuscript.

References

ACRI-ST, 2013. Sentinel-3 Optical Products and Algorithm Definition: Pixel-by-pixel uncertainty propagation in OLCI clear water branch. In: Tech. Report, Ref. S3-L2-SD-01-C01-ACR-TN, Version 2.0.

Aires, F., 2004a. Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 1. Network weights. J. Geophys. Res. 109, D10303. https://doi.org/10.1029/2003JD004173.

Aires, F., 2004b. Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 2. Output errors. J. Geophys. Res. 109, D10304. https://doi.org/10.1029/2003JD004174.

Aires, F., 2004c. Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 3. Network Jacobians. J. Geophys. Res. 109, D10305. https://doi.org/10.1029/2003JD004175.

Alikas, K., Ansko, I., Vabson, V., Ansper, A., Kangro, K., Uudeberg, K., Ligi, M., 2020. Consistency of radiometric satellite data over lakes and coastalwaters with local field measurements. Remote Sens. 12 https://doi.org/10.3390/rs12040616.

Antoine, D., Morel, A., 1999. A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour (MERIS instrument): principle and implementation for atmospheres carrying various aerosols including absorbing ones. Int. J. Remote Sens. 20 https://doi.org/10.1080/014311699212533.

Antoine, D., Slivkoff, M., Klonowski, W., Kovach, C., Ondrusek, M., 2021. Uncertainty assessment of unattended above-water radiometric data collection from research vessels with the Dynamic Above-water Radiance (L) and Irradiance (E) Collector (DALEC). Opt. Express. https://doi.org/10.1364/oe.412022.

Ardizzone, L., Kruse, J., Wirkert, S., Rahner, D., Pellegrini, E.W., Klessen, R.S., Maier-Hein, L., Rother, C., Köthe, U., 2019. Analyzing inverse problems with invertible neural networks. In: 7th International Conference on Learning Representations, ICLR 2019.

Arras, K., 1998. An Introduction to Error Propagation: Derivation, Meaning, and Examples of Equation $C_y = F_x C_x F_x^T$. Lausanne Swiss Fed. Inst. Technol, Lausanne.

Babin, M., 2000. Coastal surveillance through observation of ocean colour (COASTLOOC). Final Report, Project ENV4-CT96-0310. In: Laboratoire de Physique et Chimie Marines. Villefranche-sur-mer, France.

Bailey, S.W., Franz, B.A., Werdell, P.J., 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. Opt. Express 18, 7521. https://doi.org/10.1364/oe.18.007521.

Baird, M.E., Cherukuru, N., Jones, E., Margvelashvili, N., Mongin, M., Oubelkheir, K., Ralph, P.J., Rizwi, F., Robson, B.J., Schroeder, T., Skerratt, J., Steven, A.D.L., Wild-Allen, K.A., 2016. Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: comparison with satellite data. Environ. Model. Softw. https://doi.org/10.1016/j.envsoft.2015.11.025.

Bennartz, R., Fischer, J., 2000. A modified k-distribution approach applied to narrow band water vapour and oxygen absorption estimates in the near infrared. J. Quant. Spectrosc. Radiat. Transf. 539–553.

Bi, S., Li, Y., Wang, Q., Lyu, H., Liu, G., Zheng, Z., Du, C., Mu, M., Xu, J., Lei, S., Miao, S., 2018. Inland water Atmospheric Correction based on Turbidity Classification using OLCI and SLSTR synergistic observations. Remote Sens. 10 https://doi.org/10.3390/rs10071002.

Białek, A., Douglas, S., Kuusk, J., Ansko, I., Vabson, V., Vendt, R., Casal, T., 2020. Example of Monte Carlo method uncertainty evaluation for above-water ocean colour radiometry. Remote Sens. 12 https://doi.org/10.3390/rs12050780.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Inc, USA.

Brajard, J., Jamet, C., Moulin, C., Thiria, S., 2006. Use of a neuro-variational inversion for retrieving oceanic and atmospheric constituents from satellite ocean colour sensor: Application to absorbing aerosols. Neural Netw. 19 https://doi.org/10.1016/j.neunet.2006.01.015.

Brando, V., Lovell, J., King, E., Boadle, D., Scott, R., Schroeder, T., 2016. The potential of autonomous ship-borne hyperspectral radiometers for the validation of ocean color radiometry data. Remote Sens. 8, 150. https://doi.org/10.3390/rs8020150.

Brandt, S., 2014. Data analysis: Statistical and computational methods for scientists and engineers. https://doi.org/10.1007/978-3-319-03762-2.

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: European Space Agency. (Special Publication) ESA SP.

Doerffer, R., 2008. Algorithm Theoretical Basis Document (ATBD) MERIS regional coastal and lake case 2 water project atmospheric correction ATBD. Water Version 1, 1–42.

Draper, N.R., Smith, H., 1998. Applied regression analysis. John Wiley & Sons.

ESA, 2013. Reference Model for MERIS Level 2 Processing Third MERIS reprocessing: Ocean Branch. In: Doc. No. PO-TN-MEL-GS-0026, Issue 5, rev 4.

EUMETSAT, 2018. Sentinel-3 OLCI Marine User Handbook. D-64295 Darmstadt, Germany. EUMETSAT. https://www.eumetsat.int/media/45743.

EUMETSAT, 2019. Sentinel-3 Product Notice – OLCI Level-2 Ocean Colour, Product Notice ID EUM/OPS-SEN3/TEN/19/1068317. In: S3.PN-OLCI-L2M.001 Version 1. EUMETSAT, p. 10. https://www.eumetsat.int/media/43298.

EUMETSAT, 2021. Sentinel-3 OLCI L2 report for baseline collection OL_L2M_003 - EUM/RSP/REP/21/1211386. EUMETSAT. https://www.eumetsat.int/media/47794.

Fan, Y., Li, W., Gatebe, C.K., Jamet, C., Zibordi, G., Schroeder, T., Stamnes, K., 2017. Atmospheric correction over coastal waters using multilayer neural networks. Remote Sens. Environ. 199 https://doi.org/10.1016/j.rse.2017.07.016.

Fell, F., Fischer, J., 2001. Numerical simulation of the light field in the atmosphere-ocean system using the matrix-operator method. J. Quant. Spectrosc. Radiat. Transf. 3, 351–388.

Fischer, J., Grassl, H., 1984. Radiative transfer in an atmosphere–ocean system: an azimuthally dependent matrix-operator approach. Appl. Opt. https://doi.org/10.1364/ao.23.001032.

Frouin, R., Pelletier, B., 2015. Bayesian methodology for inverting satellite ocean-color data. Remote Sens. Environ. 159 https://doi.org/10.1016/j.rse.2014.12.001.

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. https://doi.org/10.3389/feart.2019.00145.

Gao, B.C., Montes, M.J., Li, R.R., Dierssen, H.M., Davis, C.O., 2007. An atmospheric correction algorithm for remote sensing of bright coastal waters using MODIS land and ocean channels in the solar spectral region. IEEE Trans. Geosci. Remote Sens. 45 https://doi.org/10.1109/TGRS.2007.895949.

GCOS, 2016. The Global Observing System For Climate Implementation Needs. World Meteorol. Organ, p. 200.

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

Gillis, D.B., Bowles, J.H., Montes, M.J., Moses, W.J., 2018. Propagation of sensor noise in oceanic hyperspectral remote sensing. Opt. Express. https://doi.org/10.1364/oe.26.00a818.

Gordon, H.R., 1997. Atmospheric correction of ocean color imagery in the Earth Observing System era. J. Geophys. Res. 102, 17081–17106.

Gordon, H.R., Wang, M.H., 1994. Retrieval of water-leaving radiance and aerosol optical-thickness over the oceans with {SEAWIFS} - a preliminary algorithm. Appl. Opt. 33, 443–452.

Gordon, H.R., Clark, D.K., Brown, J.W., Brown, O.B., Evans, R.H., Broenkow, W.W., 1983. Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison of ship determinations and CZCS estimates. Appl. Opt. https://doi.org/10.1364/ao.22.000020.

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, 1–24. https://doi.org/10.3390/rs11030220.

He, X., Bai, Y., Pan, D., Tang, J., Wang, D., 2012. Atmospheric correction of satellite ocean color imagery using the ultraviolet wavelength for highly turbid waters. Opt. Express. https://doi.org/10.1364/oe.20.020754.

Hedderich, J., Sachs, L., 2016. Angewandte Statistik Methodensammlung mit R, New York.

Hollstein, A., Fischer, J., 2012. Radiative transfer solutions for coupled atmosphere ocean systems using the matrix operator technique. J. Quant. Spectrosc. Radiat. Transf. 113 https://doi.org/10.1016/j.jqsrt.2012.01.010.

Hooker, S.B., McClain, C.R., 2000. The calibration and validation of SeaWiFS data. Prog. Oceanogr. https://doi.org/10.1016/S0079-6611(00)00012-4.

Hooker, S.B., Esaias, W.E., Feldman, G.C., Gregg, W.W., Mcclain, C.R., 1992. An overview of SeaWiFS and ocean color. In: SeaWiFS Tech. Rep. Ser. NASA Tech. Memo. 104566, Vol. 1, p. 34.

Ibrahim, A., Franz, B.A., Ahmad, Z., Bailey, S.W., 2019. Multiband atmospheric correction algorithm for ocean color retrievals. Front. Earth Sci. 7 https://doi.org/10.3389/feart.2019.00116.

IOCCG, 2010. Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. In: Wang, M. (Ed.), Reports of the International Ocean-Colour Coordinating Group, No. 10. IOCCG, Dartmouth, Canada.

IOCCG, 2019. Uncertainties in ocean colour remote sensing. In: Mélin, F. (Ed.), IOCCG Report Series, No. 18, International Ocean Colour Coordinating Group, Dartmouth, Canada.

Jamet, C., Thiria, S., Moulin, C., Crepon, M., 2005. Use of a neurovariational inversion for retrieving oceanic and atmospheric constituents from ocean colar imagery: A feasibility study. J. Atmos. Ocean. Technol. 22 https://doi.org/10.1175/JTECH1688.1.

JCGM (Joint Committee For Guides In Measurements), 2008. Evaluation of measurement data — Guide to the expression of uncertainty in measurement. Int. Organ. Stand, Geneva. https://doi.org/10.1373/clinchem.2003.030528. ISBN 50.

JCGM (Joint Committee For Guides In Measurements) S1, 2008. Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method. Evaluation JCGM 101, 2.

Jones, E.M., Baird, M.E., Mongin, M., Parslow, J., Skerratt, J., Lovell, J., Margvelashvili, N., Matear, R.J., Wild-Allen, K., Robson, B., Rizwi, F., Oke, P., King, E., Schroeder, T., Steven, A., Taylor, J., 2016. Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the Great Barrier Reef. Biogeosciences. https://doi.org/10.5194/bg-13-6441-2016.

Karpouzli, E., Malthus, T., 2003. The empirical line method for the atmospheric correction of IKONOS imagery. Int. J. Remote Sens. 24 https://doi.org/10.1080/0143116021000026779.

Kuchinke, C.P., Gordon, H.R., Franz, B.A., 2009. Spectral optimization for constituent retrieval in Case 2 waters I: Implementation and performance. Remote Sens. Environ. 113 https://doi.org/10.1016/j.rse.2008.11.001.

Lavender, S.J., Pinkerton, M.H., Moore, G.F., Aiken, J., Blondeau-Patissier, D., 2005. Modification to the atmospheric correction of {SeaWiFS} ocean colour images over turbid waters. Cont. Shelf Res. 25, 539–555. https://doi.org/10.1016/j.csr.2004.10.007.

Lee, Z.P., Arnone, R., Hu, C., Werdell, P.J., Lubac, B., 2010. Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm. Appl. Opt. 49 https://doi.org/10.1364/AO.49.000369.

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, 1997–2010. https://doi.org/10.3390/rs11202400.

Liu, D.C., Nocedal, J., 1989. On the limited memory {BFGS} method for large scale optimization. Math. Program. 45, 503–528. https://doi.org/10.1007/bf01589116.

Maritorena, S., d'Andon, O.H.F., Mangin, A., Siegel, D.A., 2010. Merged satellite ocean color data products using a bio-optical model: characteristics, benefits and issues. Remote Sens. Environ. 114 https://doi.org/10.1016/j.rse.2010.04.002.

Martinetz, T., Schulten, K., 1991. A "Neural-Gas" network learns topologies. Artificial Neural Networks 1, 397–402. https://doi.org/10.1016/C2009-0-12894-5. http://webmail.inb.uni-luebeck.de/inb-publications/pdfs/MaSc91.pdf.

McKinna, L.I.W., Cetinić, I., Chase, A.P., Werdell, P.J., 2019. Approach for propagating radiometric data uncertainties through NASA ocean color algorithms. Front. Earth Sci. 7, 1–17. https://doi.org/10.3389/feart.2019.00176.

McKinna, L.I.W., Cetinić, I., Werdell, P.J., 2021. Development and validation of an empirical ocean color algorithm with uncertainties: a case study with the particulate backscattering coefficient. J. Geophys. Res. Ocean. 126 https://doi.org/10.1029/2021JC017231.

Merchant, C.J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., De Leeuw, G., Mittaz, J., Poulsen, C., Povey, A.C., Reuter, M., Sathyendranath, S., Sandven, S., Sofieva, V.F., Wagner, W., 2017. Uncertainty information in climate data records from Earth observation. Earth Syst. Sci. Data 9. https://doi.org/10.5194/essd-9-511-2017.

Metrology, J.C.F.G.I., 2017. International vocabulary of metrology – Basic and general concepts and associated terms (VIM), 3rd. Tetrahedron Lett, p. 58.

Milutinović, S., Bertino, L., 2011. Assessment and propagation of uncertainties in input terms through an ocean-color-based model of primary productivity. Remote Sens. Environ. 115 https://doi.org/10.1016/j.rse.2011.03.013.

Mobley, C.D., 2015. Polarized reflectance and transmittance properties of windblown sea surfaces. Appl. Opt. https://doi.org/10.1364/ao.54.004828.

Moore, G., Lavender, S., 2010. OLCI Bright Waters Atmospheric Correction (mesotrophic to high turbidity). In: Algorithm Theoretical Basis Document, S3-L2-SD-03-C08-ARG-ATBD_BWAC.

Morel, A., Bélanger, S., 2006. Improved detection of turbid waters from ocean color sensors information. Remote Sens. Environ. 102, 237–249. https://doi.org/10.1016/j.rse.2006.01.022.

Nazeer, M., Nichol, J.E., Yung, Y.K., 2014. Evaluation of atmospheric correction models and Landsat surface reflectance product in an urban coastal environment. Int. J. Remote Sens. 35 https://doi.org/10.1080/01431161.2014.951742.

Parker, W.S., 2013. Ensemble modeling, uncertainty and robust predictions. Wiley Interdiscip. Rev. Clim. Chang. 4 https://doi.org/10.1002/wcc.220.

Penny, W.D., Roberts, S.J., 1998. Error bars for linear and nonlinear neural network regression models.

Plass, G.N., Kattawar, G.W., Catchings, F.E., 1973. Matrix operator theory of radiative transfer. 1: Rayleigh scattering. Appl. Opt. 12, 314–329.

Ruddick, K.G., Ovidio, F., Rijkeboer, M., 2000. Atmospheric correction of {SeaWiFS} imagery for turbid coastal and inland waters. Appl. Opt. 39, 897. https://doi.org/10.1364/ao.39.000897.

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.

Santer, R., Zagolski, F., Ramon, D., Fischer, J., Dubuisson, P., 2005. Uncertainties in radiative transfer computations: consequences on the MERIS products over land. Int. J. Remote Sens. 26 https://doi.org/10.1080/01431160500166508.

Schaale, M., Schroeder, T., 2013. An extended validation test for data input into parameterized retrieval algorithms. AIP Conf. Proc. 1531, 951–954. https://doi.org/10.1063/1.4804929.

Schroeder, T., 2005. Fernerkundung von Wasserinhaltsstoffen in Küstengewässern mit MERIS unter Anwendung expliziter und impliziter Atmosphärenkorrekturverfahren (Remote sensing of coastal waters with MERIS on basis of explicit and implicit atmospheric correction algorithms). Dissertation. Freie Universität Berlin, Dep. Earth Sci. https://doi.org/10.17169/refubium-11210.

Schroeder, T., Behnert, I., Schaale, M., Fischer, J., Doerffer, R., 2007a. Atmospheric correction algorithm for MERIS above case-2 waters. Int. J. Remote Sens. https://doi.org/10.1080/01431160600962574.

Schroeder, T., Schaale, M., Fischer, J., 2007b. Retrieval of atmospheric and oceanic properties from MERIS measurements: A new Case-2 water processor for BEAM. Int. J. Remote Sens. https://doi.org/10.1080/01431160701601774.

Sentinel-3 OLCI Processing Baseline [WWW Document], n.d. URL https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-olci/processing-baseline (accessed 4.6.21).

Shen, M., Duan, H., Cao, Z., Xue, K., Qi, T., Ma, J., Liu, D., Song, K., Huang, C., Song, X., 2020. Sentinel-3 OLCI observations of water clarity in large lakes in eastern China: implications for SDG 6.3.2 evaluation. Remote Sens. Environ. 247 https://doi.org/10.1016/j.rse.2020.111950.

Siegel, D.A., Maritorena, S., Robinson, W., 2000. Atmospheric correction of satellite ocean colour imagery: the black pixel assumption. Appl. Opt. 39, 3582–3591.

Singh, R.K., Shanmugam, P., He, X., Schroeder, T., 2019. UV-NIR approach with non-zero water-leaving radiance approximation for atmospheric correction of satellite imagery in inland and coastal zones. Opt. Express. https://doi.org/10.1364/oe.27.0a1118.

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

Steven, A.D.L., Baird, M.E., Brinkman, R., Car, N.J., Cox, S.J., Herzfeld, M., Hodge, J., Jones, E., King, E., Margvelashvili, N., Robillot, C., Robson, B., Schroeder, T., Skerratt, J., Tickell, S., Tuteja, N., Wild-Allen, K., Yu, J., 2019. eReefs: an operational information system for managing the Great Barrier Reef. J. Oper. Oceanogr. 1–17 https://doi.org/10.1080/1755876x.2019.1650589.

Taylor, B.N., Kuyatt, C.E., 1994. NIST Technical Note 1297 1994 Edition. In: Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results. Natl. Inst. Stand. Technol.

Vanhellemont, Q., 2020. Sensitivity analysis of the dark spectrum fitting atmospheric correction for metre- and decametre-scale satellite imagery using autonomous hyperspectral radiometry. Opt. Express 28. https://doi.org/10.1364/oe.397456.
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 https://doi.org/10.1016/j.rse.2021.112284.
Wang, M., 2007. Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: Simulations. Appl. Opt. https://doi.org/10.1364/AO.46.001535.
Wang, M., 2016. Rayleigh radiance computations for satellite remote sensing: accounting for the effect of sensor spectral response function. Opt. Express. https://doi.org/10.1364/oe.24.012414.
Wang, M., Shi, W., 2007. The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing. Opt. Express 15, 15722. https://doi.org/10.1364/oe.15.015722.