ESA Climate Change Initiative Aerosol_cci+

End-to-end ECV uncertainty budget (E3UB) Version 2.1 (D2.2b)

Document reference:
Aerosol_cci+_E3UB_v2.1.doc

DOCUMENT STATUS SHEET

	FUNCTION	NAME	DATE	SIGNATURE
LEAD AUTHOR	Editor, author	Thomas Popp	02.03.2020 23.04.2020	
CONTRIBUTING AUTHORS	Co-authors	Peter North Yves Govaerts Marta Luffarelli	10.03.2020 24.04.2020 13.11.2020	
REVIEWED BY	Science leader review	Thomas Popp	10.03.2020 27.04.2020 17.01.2021	
APPROVED BY	Technical officer (ESA)	Simon Pinnock		
ISSUED BY	Project manager	Thomas Popp		

EXECUTIVE SUMMARY

This document summarizes the status of uncertainty characterization close to the end of the second year of the Aerosol_cci+ project. Uncertainty characterization means the identification of different sources of error, assessment of their behaviour, and calculation of the sensitivity of retrieval algorithms to each source.

Algorithm theoretical baseline documents (ATBD) have been prepared for the two dual view algorithms involved in the Aerosol_cci+ project, which detail (amongst other things) the propagation and treatment of uncertainties. This second version of this report summarises the major principles and knowledge contained within the documents of both algorithms to compare the techniques used and contextualise them within the field of metrology. These uncertainty estimates are based on the developer's understanding of their retrievals and their sensitivities to the environment, with significant consideration given to the pre-launch calibration of each sensor.

Just as it is important to validate a data product to demonstrate that it is fit-for-purpose, it is necessary to validate the estimates of uncertainty to demonstrate that they represent the distribution of error. This is done as outlined in the Product Validation Plan (PVP) and the results are summarized in the Product Validation and Intercomparison Report (PVIR).

This document consists of 6 sections. After an introduction, it summarises the terminology of error characterisation (which is harmonized between all CCI projects). The sources of error are then generalised, followed by an outline of the uncertainty estimation techniques used in both dual view algorithms of the project. Some advice from data producers on appropriate use of aerosol data is provided, leading to final conclusions.

Issue	Date	Modified Items / Reason for Change
0.9	02.03.2020	Document structure and first draft
1.0	10.03.2020	Update by responsible for mature algorithm Review by science leader
1.1	27.04.2020	Revision according to RIDs raised by ESA on 17/04/2020
2.0	13.11.2020	Update by responsibles for second algorithm
2.1	17.01.2021	Science leader review, minor formal updates

TABLE OF CONTENTS

1 INTRODUCTION

This report aims to review the results of uncertainty characterization provided by the dual view algorithms within the Aerosol_cci+ project. It summarises the known sources of error in the aerosol products, how each of them can be characterized and how they combine together. It also identifies remaining gaps in those elements of the comprehensive uncertainty characterisation.

The scope of this report is to provide an overview of the management of uncertainty, not to replicate the ATBDs for the algorithms (in which the methods to estimate uncertainty are detailed). The report has been and will be updated regularly to represent the progress made on uncertainty propagation and sensitivity studies produced during the project. Results of the validation of uncertainty values against reference data are provided in the Product Validation and Intercomparison Report (PVIR).

In Aerosol_cci+ only test datasets (up to one full year global for each sensor) of the dual view sensors SLSTR (onboard Sentinel-3A and Sentinel-3B), AATSR (ENVISAT) and ATSR-2 (ERS-2) are processed to allow sufficient statistics for validation of the AOD products and their uncertainties. Once benchmarked, full mission reprocessing with new algorithm versions is then subject of the operational Copernicus Climate Change Service.

1.1 References

1.1.1 Applicable Documents

[AD1] The Statement of Work, reference ESA-CCI-EOPS-PRGM-SOW-18-018, issue 1, revision 6, dated May 31st, 2018, and its specific annex C.

[AD2] The Contractor's Proposal reference 3022091 revision 1.1, dated 10 December 2018

1.1.2 Reference Documents

[RD1] ATBD for SU's AATSR algorithm, v4.3, dated 15.05.2017.

[RD2] Product Validation Plan, version 1.2, dated 08.10.2019.

[RD3] Aerosol_cci2 Comprehensive Error Characterisation Report (CECR), version 3.2, dated 17.08.2017.

[RD4] ATBD V1.0 for CISAR SLSTR algorithm V2.0, dated November 2020.

1.1.3 Academic References

Beers, Y.: Introduction to the theory of error, Addison-Wesley, Massachusetts, 1957.

Bevan, S., North, P., Los, S. & Grey, W. (2012). A global dataset of atmospheric aerosol optical depth and surface reflectance from AATSR. Remote Sensing of Environment 116, 199-210.

BIPM: Guide to the Expression of Uncertainty in Measurement (GUM), Bureau International des Poids et Mesures, http://www.bipm.org/en/publications/guides/gum.html, 2008.

Cox, C. and Munk, W. (1954). Statistics of the sea surface derived from sun glitter, Journal of Marine Research, 13, 198-227.

Hughes, Ifan and Hase, Thomas P. A. (2010) *Measurements and their uncertainties : a practical guide to modern error analysis.* Oxford : New York, NY: Oxford University Press. ISBN 9780199566327.

Koepke, P. (1984). Effective Reflectance of Oceanic Whitecaps, *Applied Optics*, 23(11), 1816-1824.

Kotchenova, S.Y. and Vermote, E.F. (2007). Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part II. Homogeneous Lambertian and anisotropic surfaces. *Applied Optics*, Vol. 46, Issue 20, pp. 4455-4464.

Monahan, E.C. and O'Muircheartaigh, I. (1980). Optimal power-law description of oceanic whitecap dependence on wind speed, *Journal of Physical Oceanography*, **10**(12), 2094–2099, 1980.

Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters), *Journal of Geophysical Research*, 93(C9), 10749-10768, 1988.

North, P.R.J. (1996), Three-dimensional forest light interaction model using a Monte Carlo method, *IEEE Transactions on Geoscience and Remote Sensing*, 34(5), 946-956.

North, P., Briggs, S., Plummer, S. & Settle, J. (1999). Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multiangle imagery. IEEE Transactions on Geoscience and Remote Sensing 37(1), 526-537.

Popp, Thomas, Report on AVHRR aerosol demonstration dataset (D5_8), revision 1.1, H2020 project FIDUCEO, dated 11.10.2019

Povey, A.C. and R.G. Grainger: Known and unknown unknowns: uncertainty estimation in satellite remote sensing, Atmos. Meas. Tech., 8, 4699–4718, 2015.

Sayer, A. M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., Patadia, F., Popp, T., Povey, A. C., Stebel, K., and Witek, M. L.: A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing, Atmospheric Measurements and Techniques, 13, 373–404, https://doi.org/10.5194/amt-13-373-2020, 2020.

List of acronyms

(A)ATSR	(advanced) along track scanning radiometer
AERONET	aerosol robotic network
AOD	aerosol optical depth
ATBD	algorithm theoretical basis document
BRDF	bidirectional reflectance function
CCI	climate change initiative
ECMWF	European centre for mid-range weather forecasting
ECV	environmental climate variable
LUT	look up table
pdf	probability density function
rmse	root mean square error
SNR	signal to noise ratio
SLSTR	Sea and Land Surface Temperature Radiometer
SSA	single scattering albedo
SU	Swansea University
TOA	top of atmosphere
UV/VIS	ultraviolet/visible (parts of the spectrum)

2 DEFINITION OF TERMS

This section is an (adapted) copy of Section 6 of the CCI Project Guidelines that were an output of discussions of the initial ten ECV teams (including Aerosol_cci) and the Climate Model User Group during the very first CCI colocation meeting in 2010. As a result of those discussions the CCI program agreed on these common definitions, from which we copy the parts relevant for Aerosol_cci.

2.1 Describing error and uncertainty

A **measurement** is a set of operations intended to determine the value of a quantity. Following BIPM (2008), it is useful to define the term **measurand** as the particular quantity subject to measurement such that the phrases "true value of a quantity" and "value of the measurand" are synonymous.

Very few instruments directly measure the measurand. An instrument generally reports a quantity from which the magnitude of the measurand is estimated (e.g. an instrument sensitive to infrared light might be used to measure the temperature of an object). The process of measurement is intrinsically inexact. The difference between a measured value and the value of the measurand is called the error. Traditionally (e.g. Beers, 1975), the word "error" has also meant a numerical value that estimates the variability of the error if a measurement is repeated (i.e. a width of the distribution of possible errors). To avoid this ambiguity, the CCI program has adopted the BIPM (2008) definitions

- **error (of measurement)**: the result of a measurement minus a true value of the measurand;
- **uncertainty (of measurement)**: a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

The "true" value of the error is rarely known such that its magnitude is hypothetical. Error is frequently viewed as having a random and a systematic component, defined by BIPM (2008) as

- **random error**: result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatable conditions;
- **systematic error**: the mean that would result from an infinite number of measurements of the same measurand carried out under repeatable conditions minus the true value of the measurand.

A more detailed discussion of these concepts applied to satellite remote sensing can be found in Povey and Grainger (2015).

Two qualitative terms not defined in BIPM (2008) but commonly used to describe a measurement (e.g. Beers, 1957, Hughes and Hase, 2010) are precision and accuracy, defined here as

- **precision**: qualitative measure of the (relative) magnitude of the random uncertainty;
- **accuracy**: qualitative measure of the (relative) magnitude of the systematic uncertainty.

Although it is not possible to compensate for random error, the resulting uncertainty in our estimate of the measurand can usually be reduced by averaging repeated, independent observations. The statistical distribution of random error can be described by a probability density function (pdf) for which the **expected value** (i.e. the average over the pdf) is zero. As the random error often arises from the addition of many effects, the central limit theorem suggests that a Gaussian distribution should be a good representation of the pdf. Hence, the uncertainty resulting from random error is commonly represented by the one-sigma standard deviation that would be obtained from repeated measurements of the same quantity under the same conditions. If N uncorrelated observations are available, the random component of uncertainty is their one-sigma standard deviation multiplied by a factor of $1/\sqrt{N}$. The smallest possible change in value that can be observed can be taken as half the uncertainty (which can also be used as the detection limit of the instrument).

In some circumstance, a correction can be applied to compensate for systematic errors. Afterwards, the expected value of the error is assumed to be zero (i.e. the correction leaves only random errors), but there will be random and systematic errors in the correction itself.

An **error budget** is a summary of the known sources of error in a measurement with estimates of their resulting uncertainty (and, preferably, information on how those uncertainties combine). Standard methods of error propagation (e.g. Hughes and Hase, 2010) can be used to transform uncertainties into measurement units. The total uncertainty is the combined total accounting for any correlation between component errors.

When multiple measurands are estimated simultaneously, their uncertainty may not be independent. Their mutual uncertainty is represented with a covariance matrix $\mathbf{S}_{ij} = \langle\sigma_i\sigma_j\rangle$, where each term is the expected value of the product of the uncertainties σ_i of the ith and jth measurands. If the measurands are independent then the off-diagonal terms are zero and the uncertainty on each measurand is given by the square-root of the corresponding diagonal element.

2.2 Validation of Measurements

Validation is the assessment of a measurement and its uncertainty. This is principally achieved by **external validation**, the comparison of a measurement to an independent measurement (from a different instrument). The independent estimate of the measurand is termed the **validation value**. The **discrepancy** is defined as the difference between the measurement and the validation value. A small average discrepancy (e.g. the root-mean-square) between the measurement and validation value is indicative of an accurate measurement but could also result from a fortuitous cancellation of error terms. The uncertainty is assessed by its ability to characterise the observed distribution of discrepancies.

It is only practical to report individual discrepancies for small data sets. For the large number of measurements typical in satellite remote sensing, validation involves statistically characterising the discrepancies. The behaviour of the instrument (or algorithm) is often expected to be a function of the conditions observed, so it is typical to characterize discrepancies separately over a number of “regimes” (e.g. land and sea). The choice of regimes could come from a cluster analysis of discrepancy (if the difference in regimes causes differences in systematic error) but more commonly comes from knowledge of the measurement process.

Consider a set of n measurements $\{x_1 \pm \delta x_1, x_2 \pm \delta x_2, x_3 \pm \delta x_3, \ldots x_n \pm \delta x_n\}$ together with a set of validation values $\{v_1 \pm \delta v_1, v_2 \pm \delta v_2, v_3 \pm \delta v_3, \ldots v_n \pm \delta v_n\}$. The statistical characterization of the discrepancies within a regime is made through three **quality parameters**.

- **Bias,** b, the mean value of the discrepancy,

$$b = [\ \Sigma_{i=1}^{n} x_i - v_i\ ] / n\ .$$

The expectation value of the bias is the combination of systematic errors in the measurement and in the validation value. The bias can only be attributed to the measurement if the systematic error in the validation value is known. Ideally, the bias would be zero.

- **Stability**, s, the change in bias with time,

$$s = [\ b(t + \Delta t) - b(t)\ ] / \Delta t\ .$$

Ideally, the stability would be zero over any timescale. In remote sensing the stability can display periodicity related to factors, such as instrument drift or solar illumination of the satellite (both over an orbit and seasonally). It is suggested that the stability is estimated at the same temporal scale that any trends in the data are calculated.

In some case **internal validation** can be used to check reported uncertainty. Consider the situation where an instrument measures the same quantity under conditions where the reported uncertainty does not vary. Then the variability of the measurements should agree with the reported random uncertainty.

2.3 Comparing Measurements with a Model

Further understanding can be achieved through comparison of measurements with model output. A model field is sampled as if viewed by a satellite and the same quality parameters are calculated. However,

- the uncertainty in the model may not be reported and so has to be assumed, and
- the bias cannot be attributed to the model or measurements without reference to additional information.

If the model evaluates substantially different scales to the satellite an estimate of uncertainty due to interpolation must be included. If the model is at a coarser resolution than the measurements an approach could be to compare the model value with a (weighted) average of the measurements. In that case, correlations in the systematic uncertainty need to be considered. The statistical comparison of model and measurement data must account for the influence of sampling. For example, the comparison of monthly time series from model output and averaged measurements may show discrepancies due to a lack of observations in certain regions, such as those with persistent cloud coverage.

3 SOURCES OF UNCERTAINTIES

3.1 Classification

Despite their extensive use, the classification of uncertainties as random or systematic is limited. A random uncertainty can appear to introduce a systematic bias after propagation through a non-linear equation due to its asymmetric distribution, and the distribution of a systematic error has finite width. The use of these terms is better understood as synonyms for the non-technical meanings of noise and bias, respectively.

Regardless, users have an interest in the causes of uncertainty in a measurement. The source of an error affects how it is realised and its relative importance. Povey and Grainger (2015) proposed five classifications of error by their source.

- **Measurement errors**: These result from statistical variation in the measurand or random fluctuations in the detector and electronics. To assess these accurately requires the comparison of the instrument to a thoroughly characterised reference. The response may evolve over time, necessitating the periodic repeat of calibration procedures.
- **Parameter errors**: Retrievals use auxiliary information to constrain features of the environment not easily determined from the measurements. Parameters will be produced by an independent retrieval and have associated uncertainties that propagate into the results.
- **Resolution errors (also named representativity errors)**: Aerosol is a continuous field in space and time, but satellite observations only sample it and aren't necessarily representative. Filtering procedures (for quality control) can further limit the sample. As aerosol retrievals are only performed in cloud-free conditions, the concept is also known as "fair-weather bias". Filtering can also remove exceptional events, as high aerosol optical depth (AOD) plumes often fail cloud clearing, producing a low bias in averages.
- **Approximation errors (also named forward model errors)**: It is not always practical to evaluate the most precise formulation of a forward model. For example, the atmosphere may be approximated as plane parallel or look-up tables (LUTs) may be used rather than solving the equations of radiative transfer. Such approximations will introduce error, which can be assessed by comparing the performance of the rigorous and simplified forward models through simulated data.
- **System errors**: It is also not always possible to constrain every aspect of the environment with the available information. In this project, the type of aerosol observed is an important example. These properties are assumed and errors result from their inaccuracy. The errors are a non-linear function of the observed state and are known to be a significant source of uncertainty, but that uncertainty cannot be readily quantified.

Measurement and parameter errors are both intrinsic sources of uncertainty. Measurement errors affect the quantities measured and analysed by the retrieval. Parameter errors are propagated from auxiliary inputs, such as meteorological data or empirical constants. Resolution errors result from finite sampling of a constantly varying system. These can be important as satellites do not sample randomly but with a systematic bias due to the satellite's orbit and quality control or filtering. Approximation errors represent aspects of the analysis that could have been done more precisely but do not affect the fundamental measurand. System errors express choices in the analysis that alter the measurand. The system error results from the difference between the assumed system and reality.

3.2 Qualitative description of major sources of uncertainties

The following table provides an overview of the common sources of error across the two algorithms in the project, based on the description of the algorithms in their ATBDs [RD1-4].

Table 3-1: *Qualitative error budget for aerosol retrievals.*

Source of uncertainty	Description	Qualitative estimate of contribution
Cloud screening and safety zone	Capabilities depend on available spectral range (e.g. thermal bands are important); safety zone also masks elevated AOD around clouds	High for UV/VIS sensors, medium for stratospheric algorithms
Overpass time	Polar orbiting sensors provide typically one or two sun-synchronous overpass times per day	High when comparing to different sensors or against models
Land surface reflectance (BRDF)	Can be estimated from vegetation index and/or mid-infrared bands, drawn from a climatology or ECV, or retrieved alongside AOD from multi-view data	High for nadir-only sensors, with larger uncertainty at higher reflectances
Ocean surface reflectance	Estimated using white caps parameterisation and possibly a climatology of ocean colour	Medium
Calibration	Absolute radiance calibration is critical with spectral calibration being less critical due to the broad-band features considered	Medium
Aerosol optical properties	This includes spectral extinction, absorption, phase function and shape (degree of sphericity)	Medium to high for sensors with low information content, low for AOD < 0.15
Vertical aerosol profile	Different assumptions are made for different aerosol types but sensitivity at TOA is small for VIS/IR sensors, increasing in the TIR	Medium for UV observations and absorbing aerosol, low otherwise
Directional reflectance ratio	A surface model is used to describe angular and spectral variation of surface reflectance by minimizing the discrepancy between the parameterised surface model and the measured directional reflectances – any uncertainties of this surface model remain as uncertainties in the aerosol retrieval.	Medium for multi-view sensors
Pixel size	Ranges from 1x1 km^2 for radiometers to 16x7 km^2 for polarization instruments to approximately 0.25˚x0.5˚ for spectrometers	Medium when pixels dimension approach 50 km (approximate scale of aerosol variation)
Trace gas concentration profiles	Critical absorption bands are usually avoided	Low
Radiative transfer forward model	Typical accuracy < 1%	Low
Look-up table discretization	Uncertainty often a function of the number of discretization points	Low
Wind speed	Used to estimate ocean reflectance	Low
Sampling	Practically all sensors under-sample the aerosol fields in time; different samplings lead to bias between different products	Depends strongly on the repeat cycle of the sensor and its swath width
Aggregation to 10x10 km^2	Aims to improve the signal-to-noise ratio and exclude outliers	Reduces random error (but not systematic) and may decrease representivity of data

4 METHODS OF DETERMINING UNCERTAINTIES

This section summarizes principles used for characterising uncertainties first for individual measured pixels (level 2 products in orbit projection), and secondly for spatial and temporal averages (level3 gridded products). A good overview of different methods to estimate level2 uncertainties is given in section 2 of Sayer, et al., 2020.

In response to user requirements, all Aerosol_cci+ products include a pixel-level estimate of the uncertainty. There are several approaches to estimating such values.

- Prognostic uncertainty estimates can be calculated with analytical calculations, such as the traditional propagation of errors via Jacobians, assume that the errors are normally distributed, that the uncertainties on the retrieval inputs are accurately quantified, and that the formulation of the algorithm is physically valid (or, at least, that the distribution of error resulting from the formulation is known).
- Diagnostic uncertainty estimates come from validation activities, such as comparison against the AERONET sun photometer network, which can approximate the error in a measurement by the discrepancy between the retrieved and the validation values. This can be a useful investigation of uncertainty after the retrieval processing has been completed where little is known about the distribution of error. However, this requires the availability of reference measurements under all retrieval conditions (e.g. for situations near clouds or in coastal waters this is usually not the case) and that the uncertainties of the reference data are significantly smaller than the retrieval errors and therefore can be neglected (for AERONET AOD with uncertainties of 0.01 – 0.02 this is the case except for near-zero values of AOD).
- Theoretical information content analysis can be used when neither of the previous options is possible, such as before the launch of the satellite. Significant differences between pre-launch expectations and post-launch results could indicate inaccurate assumptions in the algorithm.
- Ensembles of different retrieval algorithms or different parameters may be useful in characterising errors that cannot be quantified directly. Aerosol type can be assessed in this manner, by attempting retrieval on one measurement with several types.
- Quality flags provide the user with qualitative uncertainty information, indicating where the formulation of the algorithm may not be valid (such that analytical calculations are not necessarily meaningful). This can be used to exclude pixels with unknown or incomplete uncertainty estimates from any analysis. However, at best these give only a qualitative indication of the expected uncertainty and a user must test the impact of applying any quality flag on the obtainable sampling and results in the intended application.

The use of uncertainty propagation techniques implies that all systematic biases can be removed. This is one of the main goals of the cyclic algorithm development and evaluation approach implemented in Aerosol_cci: to understand causes for biases and subsequently remove them. By comparing each new algorithm version with independent ground-based reference data, biases (in different regions or seasons) are identified and the next algorithm development step measures are tested to reduce or even eliminate those biases. This allows then to apply uncertainty propagation with unbiased standard uncertainties. For the latest versions of the Swansea algorithm evaluated in Aerosol_cci2 with AATSR, overall biases were within AERONET uncertainties: ±0.01 for high AOD (>0.2) and ±0.02 for low AOD (<0.2) over both land and ocean. For SLSTR first analysis show somewhat larger biases (~0.05) which need to be further analysed.

4.1 AOD from dual view algorithms

Two algorithms are used within the Aerosol_cci+ project to produce the AOD ECV. Their techniques for determining the pixel-level uncertainty differ and are briefly summarised here. (For a more exhaustive description, see their respective ATBDs). In this second version of this document the uncertainty propagation of both algorithms by Swansea university and by Rayference are described here.

4.1.1 Error propagation approach in dual view SU

While the original CCI baseline dataset (Bevan et al., 2012) used an empirical error approximation, within Aerosol_cci this has been developed to give full analytic propagation of uncertainty for each retrieval. Over both land and ocean, the retrieval uses non-linear optimisation of an error function, of the form

$$\mathrm{X}^2 = \sum_{\lambda=1}^{\lambda=n} \sum_{\Omega=0}^{\Omega=55} \frac{\left(M(\lambda,\Omega) - O(\lambda,\Omega)\right)}{\sigma_{M(\lambda,\Omega)}^2 + \sigma_{O(\lambda,\Omega)}^2}$$

where $\sigma_{M(\lambda,\Omega)}^2$ and $\sigma_{O(\lambda,\Omega)}^2$ denote estimates of 1 s.d. uncertainty in model and observation of surface reflectance at waveband / and view direction Ω (nadir is here denoted by Ω = 0°, forward/oblique view by Ω = 55°), with number of wavebands *n* = 4 for (A)ATSR and *n* = 5 for SLSTR. It is possible to also include the full covariance matrix into the X^2formulation, but currently error in model and observations are approximated as uncorrelated between channels. For correctly normalised value of chi sq, the estimate of 1 s.d. error in t_{550} is derived from the second derivative (curvature) of the error surface near the optimal value:

$$s_{t_{550}} = \left(\partial^2 \mathrm{C}^2 \big/ \partial^2 t_{550} \right)^{-0.5}$$

The curvature term is estimated by a parabolic fit of the error function for surrounding values of t_{550}.

4.1.1.1 Surface Reflectance Model uncertainty

Over land, model uncertainty was evaluated by inversion (against the test dataset) of surface BRDF values computed from a 3D Monte Carlo model. The 3D model and test dataset are described in North (1996) and North et al. (1999). Uncertainties in the modelled surface reflectance per spectral channel of the satellite instrument were estimated from this and subsequent optimization.

Over vegetated surface (NDVI>0.7)

$$\sigma_{M\_land}=\{0.01, 0.01, 0.06, 0.02, 0.02\}$$

Over bright surface (NDVI <0.1):

$$\sigma_{M\_land}=\{0.01, 0.01, 0.02, 0.15, 0.08\}$$

corresponding to the (A)ATSR / SLSTR channels {555nm, 670m, 870nm, 1.6 µm, 2.2µm}.

For the range 0.1 <=NDVI<=0.7, we use a linear combination of the uncertainties. No significant angular dependence of model error was found.

Over ocean the surface model is based on the models of Cox and Munk (1954) for glint, Monahan & O'Muircheartaigh (1980) and Koepke (1984) for foam fraction and spectral reflectance, and Morel's case I water reflectance model (1988). Uncertainty in ocean reflectance model $M_{ocean}(\lambda,\Omega,W,C)$ as function of wavelength λ, viewing direction Ω, wind speed W (ms^{-1}) and pigment concentration C (mg m^{-3}) is given by

$$s^2_{M\_ocean} = s^2_{M\_ocean\_W} + s^2_{M\_ocean\_C},$$

where

$$\sigma_{M\_ocean\_W} = |M_{ocean}(\lambda,\Omega,W+\sigma_W,C) - M_{ocean}(\lambda,\Omega,W,C)|$$

$$\sigma_{M\_ocean\_C} = |M_{ocean}(\lambda,\Omega,W,C+\sigma_C) - M_{ocean}(\lambda,\Omega,W,C)|.$$

Uncertainties in wind speed and pigment concentration are assigned values of 3 ms^{-1} and 0.1 mg m^{-3}; these values are assigned arbitrarily pending a review of realistic uncertainties, but lead to realistic uncertainty estimates in ocean AOD.

4.1.1.2 Observation errors

The per-channel observation error gives an estimate of the 1 s.d. uncertainty in derived land surface reflectance, and includes errors due to instrument calibration, radiative transfer model and LUT, and uncertainty in aerosol absorption parameterization.

$$s^2_O = s^2_{RT} + s^2_{inst} + s^2_{AerMod}.$$

Approximations for these are given by:

$$s^2_{inst} = T_S(l,q)(a_l + b_l R_{TOA}(l,q)),$$

where for ATSR-2 and AATSR we use:

$$a = \{0.0005, 0.0003, 0.0003, 0.0003\}$$

$$b = \{0.024, 0.032, 0.02, 0.033\}$$

Currently the *a* term is neglected. Values used here were reported during CCI project by D. Smith, RAL. For SLSTR on Sentinel-3A and 3B we use the same values, with a value of 0.033 for relative uncertainty in 2.2µm. Sensitivity of retrieval in the processing will be updated to use further uncertainty evaluation of the channels as available.

The term Ts gives scaling from TOA to surface at

$$T_S(\lambda,\Omega)=\frac{\delta R_{SURF}(\lambda,\Omega)}{\delta R_{TOA}(\lambda,\Omega)},$$

and is derived from the LUT coefficients at time of inversion
Based on Kotchenova and Vermote (2007), error due to RT at all channels is approximated as

$$s_{RT}^2=0.006.$$

The error due to uncertainty in aerosol absorption is approximated by

$$\sigma_{AerMod}=0.05P_R(\lambda,\Omega).$$

where P_R denotes atmospheric path radiance, estimated from LUT values, based on variability within the typical range 0.9-1.0 for aerosol SSA.

4.1.1.3 Status of treated / ignored uncertainty components in the SU algorithm

Based on the qualitative analysis of sources of uncertainty (table 3-1) the following table 4-1 shows which of these sources of error are treated in the Swansea algorithm, which are ignored, and a justification why and error source is negligible, or a plan to develop a method to estimate them.

Table 4-1: *Qualitative error budget for aerosol retrievals.*

Source of uncertainty	**Treatment in Swansea algorithm (treated / ignored)**	**Explanation (justification or plan to overcome this simplification)**
Cloud screening and safety zone	Ignored	New approach of combined AOD / COD retrieval will be tested
Overpass time	Solar angle implicitly impacts error estimate	This is an issue of documentation to the users, not of uncertainty propagation
Land surface reflectance (BRDF)	Treated with sophisticated surface reflectance model	-
Ocean surface reflectance	Treated with sophisticated surface reflectance model	-
Calibration	Treated through propagation of reflectance uncertainties; the quality of level1b calibration uncertainties which are needed as input is a concern for SLSTR	-
Aerosol optical properties	Treated through fitting of several bands / viewing directions	-
Vertical aerosol profile	neglected	Above 500nm the reflectance signal has only weak dependence on the vertical aerosol profile
Directional reflectance ratio	Treated with sophisticated surface reflectance model	-
Pixel size	neglected	Uncritical as compared to larger spectrometer pixels (500 m allow stringent cloud identification)
Trace gas concentration profiles	neglected	The algorithm uses climatological gas concentration values while the absolute impact of uncertainties / variations in the channel absorptions on the measured reflectances is very small through use of window channels with few absorption lines
Radiative transfer forward model	neglected	Those uncertainties are implicitly covered within the calibration coefficient uncertainties
Look-up table discretization	neglected	Has impact on minimizing the cost function in a small fraction of retrievals – this is under revision
Wind speed	neglected	Wind speed values come from numerical weather models, but

		arbitrary uncertainty value in wind speed currently used
Sampling	Implicitly treated	A trade-off between quality (filtering weak retrievals out) and coverage (increase sampling) has been worked out through experience over many years
Aggregation to 10x10 km^2	Not treated in uncertainty propagation	Aims to improve the signal-to-noise ratio and exclude outliers

4.1.2 Error propagation approach in CISAR

4.1.2.1 Error propagation scheme approach in CISAR

The CISAR (Combined Inversion of Surface and AeRosol) algorithm infers aerosol and cloud optical thicknesses and surface reflectance over land and sea surfaces from observations acquired by the SLSTR radiometer in the S1, S2, S3, S5, S6μm bands.

FASTRE, the forward model included in the CISAR algorithm (Govaerts and Luffarelli 2018), is explicitly solved during the inversion process, *i.e.*, it does not rely on pre-computed solutions stored in look-up tables, allowing a continuous variation of the state variables in the solution space. The retrieval scheme is based on an Optimal Estimation (OE) approach where the cost function accounts for the differences between the observations and the forward radiative transfer model, the retrieved state variables and their prior information and finally smoothness constraints on temporal and spectral variations of the atmospheric properties.

This retrieval algorithm implements a *prognostic* uncertainty inversely proportional to the magnitude of the Hessian matrix (second partial derivative) assuming that the errors are normally distributed, that the uncertainties on the retrieval inputs are accurately quantified, and that the inverted forward model is capable of accurately (without bias) simulate the observations.

4.1.2.2 Type of uncertainties

The CISAR algorithm relies on the following input information:

- SLSTR L1b observations in bands S1 to S6 converted into bidirectional reflectance factors and the pixel navigation accuracy.
- Model parameters including the total column water vapour and ozone, the surface wind speed, the aerosol layer height, the surface pressure, a cloud mask, a land-sea mask and the pixel elevation;
- The forward model uncertainty;
- Prior information on the observed system including the aerosol fine and coarse mode ratio, the surface reflectance and the aerosol and cloud optical thickness.

Most of this input information is known with a given uncertainty that are combined assuming that they are not correlated. The total measurement system uncertainty is composed of:

- The SLSTR radiometric uncertainty $\mathbf{S}_N$ is composed of the radiometric noise assumed to be equal to 2% and the pixel geolocation accuracy. This latter accuracy is in principle needed as observations are accumulated in time over a period of 16 days assuming that exactly the same pixel area is observed during that period. However, as native Level 1B are averaged over a 5 x 5 km area, this uncertainty can be neglected. The resulting error matrix $\mathbf{S}_N$ is diagonal.
- The equivalent model parameter radiometric uncertainties $\mathbf{S}_B$. Only the uncertainty on water vapour and ozone are considered with

  $$\sigma_B^2(\mathbf{b};\tilde{\lambda},\Omega_0,\Omega_v)=\left(\frac{\partial y(\mathbf{x},U_{oz};\Omega,\tilde{\lambda})}{\partial U_{oz}}\sigma_{U_{oz}}\right)^2+\left(\frac{\partial y(\mathbf{x},U_{wv};\Omega,\tilde{\lambda})}{\partial U_{wv}}\sigma_{U_{wv}}\right)^2$$

  where $\sigma_{U_{oz}}$ and $\sigma_{U_{wv}}$ are the uncertainties of the water vapour and ozone fields respectively. The terms of the $\mathbf{S}_B$ matrix writes $\mathbf{S}_B\,(i,j)=\delta(i,j)\sigma_B^2\,(\lambda,\,\Omega)$.
- Forward model uncertainty $\mathbf{S}_F$. This noise results from the fast forward model assumptions and approximations. Let $y_R(\mathbf{x},\mathbf{b})$ be an accurate forward model with an explicit representation of the atmospheric vertical profile. A global estimation of this error is performed that does not depend on the actual value of **x**. The error covariance is

  $$\sigma_F^2(\tilde{\lambda},\Omega_0,\Omega_v,\phi)=\left(\mathbf{y}(\Omega,\tilde{\lambda})\right)^2\frac{1}{N}\sum_{\{\mathbf{x},\mathbf{b}\}}\left(\frac{y_m(\mathbf{x}_\Delta,\mathbf{b}_\Delta,\Omega_\Delta)-y_R(\mathbf{x},\mathbf{b},\Omega)}{y_R(\mathbf{x},\mathbf{b},\Omega)}\right)^2$$

  where $\mathbf{x}_\Delta$ represents the discrete value of **x** and {**x**,**b**} is the domain of variation of **x** and **b**. The terms of the matrix $\mathbf{S}_F$ writes $\mathbf{S}_F\,(i,j)=\sigma_F^2\,(\tilde{\lambda},\Omega_0,\Omega_v,\varphi)$.

These three uncertainty diagonal matrices are combined into the matrix $\mathbf{S}_y=\mathbf{S}_N+\mathbf{S}_B+\mathbf{S}_F$.

- Prior information uncertainty. Prior information $\mathbf{x}_\mathbf{b}$ is provided to the CISAR algorithm on:
  - the surface parameter magnitude;
  - the aerosol optical thickness of each vertices bounding the solution space;
  - cloud optical thickness;
  - aerosol optical thickness temporal variability;
  - aerosol optical thickness spatial variability;
  - on AOT/COT spectral variability.

  An uncertainty $\sigma_{\mathbf{xb}}$ is assigned to each of these terms as described in 0.

4.1.2.3 AOT aerosol uncertainty estimation

The retrieval uncertainty is based on the OE theory, assuming linear behaviour of $y_m(\mathrm{x}, \mathrm{b}; \mathrm{m})$ in the vicinity of the solution $\hat{x}$. Under this condition, the retrieval uncertainty $\sigma_{\hat{x}}$ is determined by the shape of the cost function $\mathrm{J}(\mathrm{x})$ at $\hat{x}$

$$\sigma_x^2 = \left(\frac{\partial^2 J(x)}{\partial x^2}\right)^{-1}$$

The CISAR algorithm 0 retrieves the optical thickness τ_j for each vertex j that bounds the solution space. The total aerosol optical thickness $\tau_a = \sum_j \tau_j$. The associated uncertainty is expressed as:

$$\sigma_{\tau_a}^2 = \sum_i^n \sigma_{\tau_i}^2 + \sum_i^n \sum_{j(j\neq i)}^n \sigma_{\tau_i}\sigma_{\tau_j}$$

4.1.2.4 Status of treated / ignored uncertainty components in CISAR

Based on the qualitative analysis of sources of uncertainty (table 3-1) the following table 4-2 shows which of these sources of error are treated in the CISAR algorithm, which are ignored, and a justification why and error source is negligible, or a plan to develop a method to estimate them.

Table 4-2: *Qualitative error budget for the CISAR aerosol retrievals.*

Source of uncertainty	**Treatment in CISAR algorithm (treated / ignored)**	**Explanation (justification or plan to overcome this simplification)**
Cloud screening and safety zone	Treated	AOD and COD are retrieved simultaneously
Overpass time	Solar angle implicitly impacts error estimate	The forward model uncertainty accounts for the SZA value.
Land surface reflectance (BRDF)	Treated and fully coupled with aerosol retrieval	Aerosol retrieval is fully coupled with the retrieval of surface reflectance, including for the uncertainty estimation.
Ocean surface reflectance	Surface reflectance radiatively coupled with aerosol retrieval	The uncertainty on the surface wind speed and pressure is not propagated as their contribution is small as compared to the water vapour and ozone absorption.
Calibration	Radiometric noise is treated but possible systematic calibration error are not considered	The current version of CISAR does not propagate systematic uncertainties.
Aerosol optical properties	Treated through combining vertices of aerosol single scattering properties	It is fully propagated.
Vertical aerosol profile	Aerosol layer height is taken from a climatology data set but no uncertainty is propagated.	This uncertainty is not propagated in the current version of CISAR.
Directional reflectance ratio	Not applicable	Not applicable
Pixel size	Partially accounted through the navigation accuracy	Uncritical when pixels are aggregated at 5 km resolution.
Trace gas concentration profiles	neglected	The algorithm uses climatological gas concentration values while the absolute impact of uncertainties / variations in the channel absorptions

		on the measured reflectances is very small through use of window channels with few absorption lines
Radiative transfer forward model	Treated	Those uncertainties are implicitly covered
Look-up table discretization	Not needed, the radiative transfer equation is solved online.	Not applicable for the CISAR algorithm.

4.2 AOD propagation to gridded products

Propagation to averaged gridded products requires knowledge of the spatial and temporal correlation structures of different components of AOD uncertainties (propagated from input reflectances and from assumptions / simplifications in the retrieval algorithm). As one major component, uncertainties of the measured reflectances need to be split in components with different correlation structures and their respective correlation lengths need to be provided. This information is not (yet) available for the dual view instruments, but over the coming years space agencies plan to work it out. For other uncertainty components from the retrieval process correlation information needs to be estimated based on underlying physics. A study within the Horizon-2020 project FIDUCEO demonstrated for AVHRR AOD this propagation of components with different correlation structures (Popp, 2019). In this case the dominant uncertainty component was common, i. e. had global correlation. As a consequence, AOD uncertainties were reduced to some extent by averaging into grid cells, but remained closer to the hypothetical case of fully correlated uncertainties than to the other extreme of fully random uncertainties.

In Aerosol_cci2 a case study was conducted which evaluated AATSR AOD uncertainties for several simple formula to estimate gridded daily AOD uncertainties by combining AOD variability and propagated level2 uncertainties against differences to AERONET [RD3]. None of the tested formula performed generally well, while the most reasonable value was achieved by the sum of the propagated random uncertainty and the AOD variability.

So far, Aerosol_cci level3 products contain the worst case scenario for propagated uncertainties, i.e. the simple spatial average of all contributing pixel level uncertainties, which equals mathematically the case of spatially fully correlated uncertainties. Temporal correlations of uncertainties are neglected (and when averaging monthly means uncertainties are considered as fully independent). This is regarded as well justified since uncertainties of level1 measurements from different orbits are largely independent and also most atmospheric retrieval conditions are either changing very fast (e.g. clouds) or very slowly (e.g. surface) as compared to aerosols. The only element which may violate this assumption lies in aerosol type / optical properties, where uncertainties within the duration of an aerosol plume may have some temporal correlation.

5 GUIDELINES FOR USING THE PRODUCTS

Irrespective of the data set used, users should:

- Read the summary (~1 page) in the user guide. It summarises the key features and limitations of the data to provide a guide to avoid common confusions or mistakes.
- Be aware that filtering data by pixel level uncertainty as low quality indicates that there is little confidence that the algorithm is appropriate for that data. Depending on how conservative the data producer has chosen to be, that judgement may remove good data and/or exclude poor data and may introduce spatio-temporal artefacts.
- Use the Level 3 products (daily, monthly averaged gridded products) prepared alongside the Level 2 data. The retrieval experts have carefully considered sampling issues that may not be easily managed in post-processing without additional information (that is not necessarily available).
- Uncertainties in level3 products currently represent an upper limit of propagated uncertainties, although to our knowledge, this seems to be a reasonable conservative estimate of Level-3 uncertainty.
- Please provide feedback to the Aerosol_cci project and data producers on the quality and utility of the products. Without input, it is impossible to produce more useful products. With their feedback or questions users can contact the Aerosol_cci+ science leader, Thomas Popp (e-mail: thomas-dot-popp-at-dlr-dot-de) or the responsible scientist for the dual view retrieval algorithm at Swansea university, Peter North (e-mail: p-dot-r-dot-j-dot-north-at-swansea-dot-ac-dot-uk) and at Rayference, Yves Govaerts (yves-dot-govaerts-at-rayference-dot-eu).

6 CONCLUSIONS

An estimate of the uncertainty on a measurement is necessary to make appropriate use of the information conveyed by a measurement. Awareness of the importance of pixel-level uncertainty estimates is beginning to pervade the aerosol remote sensing community through efforts such as the yearly AeroSAT workshop. A standard for evaluating pixel level uncertainties has been worked out (Sayer, et al., 2020). By comparing the performance of the algorithms evaluated within the Aerosol_cci+ project to each other and validation data, the quality of their uncertainty estimates can be evaluated and potential areas for improvement can be illuminated.

This End-to-End ECV Uncertainty Budget is the second issue where the current status of uncertainty propagation in the second test Climate Research Data Package from two algorithms is reflected, also including datasets from the second algorithm (CISAR, optimal interpolation technique). One more version of the document is envisaged which is associated with the final dataset package revision including datasets from both dual view algorithms. Through the intercomparison of the results of uncertainty estimates in two different algorithms we intend to learn more about their strengths and weaknesses so that we can improve those uncertainty estimates. Our strategy for improving the uncertainty calculations relies on following elements:

- Validation of the pixel-level uncertainties by statistical comparison to the best estimate of the true error represented by the difference to AERONET measurements.
- Comparison of the uncertainties obtained with the two mathematically different uncertainty calculation methods in the two algorithms of Aerosol_cci+
- Analysis of progress on uncertainty propagation in other projects (e.g. other CCI projects such as SST_cci, Cloud_cci, CMUG, Horizon2020 project FIDUCEO, ESA Sentinel-3 LAW project) to assess how far this can be used in one of the dual view algorithms of Aerosol_cci+ This refers particularly to those elements of uncertainties where we are aware of deficits, such as cloud-mask induced uncertainties or propagation of uncertainties to gridded products
- For the cloud-mask induced uncertainty the approach demonstrated in FIDUCEO (Popp, 2019) requires a probabilistic cloud mask, so that AOD retrieval differences between a conservative and a more relaxed cloud mask can be assessed. This is not applicable in the two Aerosol_cci+ algorithms. However, the CISAR algorithm applies a new approach where a cloud mask is avoided but aerosol and cloud optical depth are retrieved together. This new approach holds significant potential to reduce the cloud-mask induced errors and to provide new insight into the associated uncertainties.

One important element which is to our current knowledge out of scope for Aerosol_cci+ is the propagation to gridded (level3) products, since this requires as key input uncertainties of level1b reflectance datasets which contain uncertainties broken into different components with different correlation structures (e.g. in FIDUCEO: common = globally correlated component, independent = randomly correlated component, structured component with a correlation function of the decreasing correlation with increasing distance between pixels).
Another limitation where we expect little progress is for conditions with no reference data, such as coastal water (AERONET stations are on land near the coastline, so that for the 10km super

pixels of Aerosol_cci+ and the typical 50km validation window they usually mix results of retrievals over land with results over water (coastal and open water).

End of the document