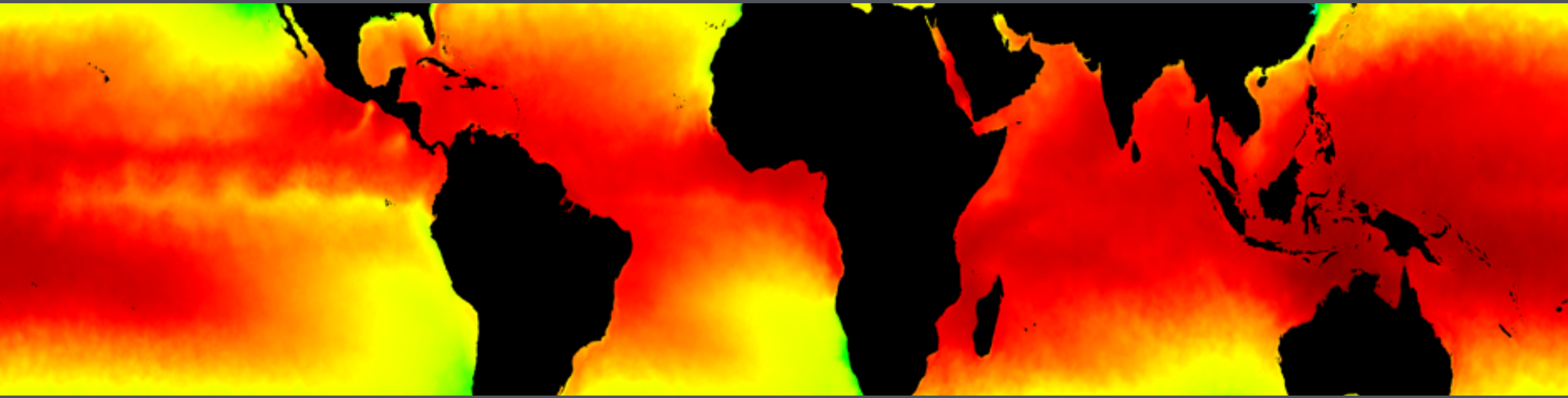
2nd LOTUS Workshop, WMO
17-19 September 2018

University of Reading

METROLOGICAL REQUIREMENTS FOR QA OF OZONE DATASETS AT THE CLIMATE DATA STORE



Michael Taylor (michael.taylor@reading.ac.uk), Jonathan Mittaz, Chris Old, Catherine Dieval, Christopher Merchant (University of Reading)

Sarah Douglas, James Ryder, Joanne Nightingale (NPL)

Christelle Barbey, Céline Fouron, Guillaume Duveau (Telespazio France)

Copernicus Europe's eyes on Earth | Climate Change Service | NPL National Physical Laboratory | FIduceo | TELESPAZIO a LEONARDO and THALES company

OVERVIEW

- The C3S EQCO process for quality assurance
  - The aims of EQCO
  - User requirements for QA of climate data at the CDS
  - QA portal and tab at the CDS
  - Gap analysis of ECVs and QA recommendations

- The FIDUCEO approach to uncertainty calculation
  - Measurement equations, measurement trees and the GUM
  - Effects tables and correlation structures
  - Calculation of pixel-level correlation matrices
  - Examples and applications

Climate Change

TELESPAZIO a LEONARDO and THALES company

NPL National Physical Laboratory

University of Reading

EQCO = Evaluation and Quality Control for Observations

(C3S 51 Lot 2)

Climate Change

University of Reading

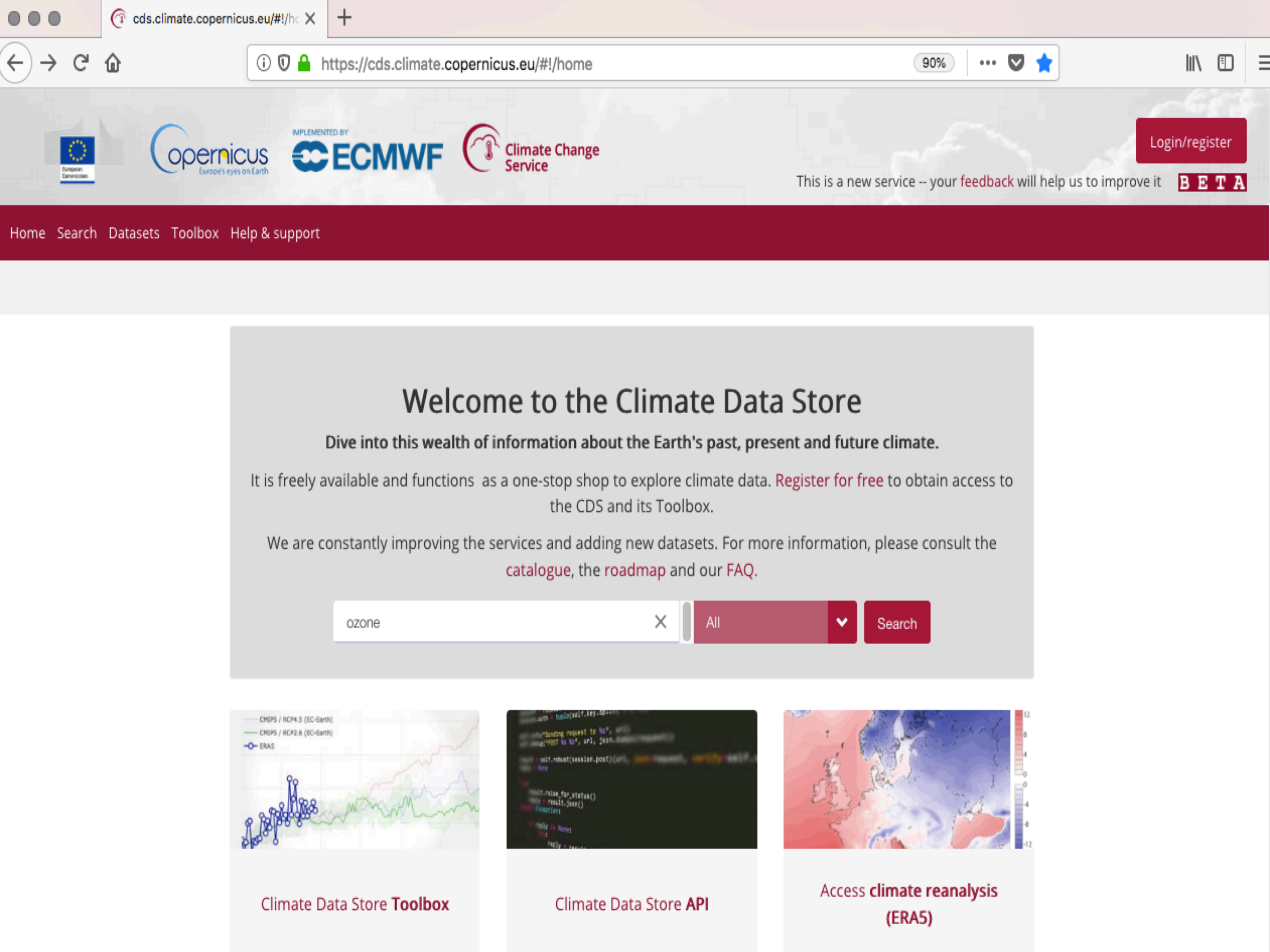
THE AIMS OF EQCO

- Identify what information users need to understand the data (i.e. QA)
- Design a prototype EQC function for the CDS (i.e. to display via a portal and grade QA completeness)
- Design a QA evaluation template and summary
- Provide scientific gaps and recommendations

USER REQUIREMENTS SURVEY

- “There are lots of data sets”
- “Not easy to decide which to use for my application”
- “Documentation often lacks detail or is dispersed"
- ”How do I know which dataset is of good quality?”
- “Validation is key”
- ”I’d like information on product traceability, algorithms and uncertainty”

Essential climate variables, ECVs (as defined by GCOS)	**Total number of products found**
Precipitation (in situ)	53
Surface Air Temp (in situ)	70
LAI	33
fAPAR	30
Sea Surface Temperature	50
Soil Moisture	62
Ozone and Aerosols	**180**
Ocean Colour	37


cds.climate.copernicus.eu/#!/hc
https://cds.climate.copernicus.eu/#!/home
90%
Copernicus
Europe's eyes on Earth
IMPLEMENTED BY
ECMWF
Climate Change Service
Login/register
This is a new service -- your feedback will help us to improve it
BETA
Home
Search
Datasets
Toolbox
Help & support
Welcome to the Climate Data Store
Dive into this wealth of information about the Earth's past, present and future climate.
It is freely available and functions as a one-stop shop to explore climate data. Register for free to obtain access to the CDS and its Toolbox.
We are constantly improving the services and adding new datasets. For more information, please consult the catalogue, the roadmap and our FAQ.
ozone
All
Search
Climate Data Store Toolbox
Climate Data Store API
Access climate reanalysis (ERA5)

Login/register

This is a new service -- your feedback will help us to improve it BETA

Search results

ozone

Sort by

Relevancy

Title

Type

Product type

Variable domain

Spatial coverage

Temporal coverage

All | Datasets

Showing 1-4 of 4 results for **ozone** ×

Ozone monthly gridded data from 1970 to present

This dataset provides estimates of the montly mean values of the **ozone** concentration, mixing ration

Climate Data Store Roadmap

): Other long-lived greenhouse gases Atmosphere (composition): **Ozone** Ocean (physics): Sea surface

ERA5 hourly data on single levels from 2000 to present

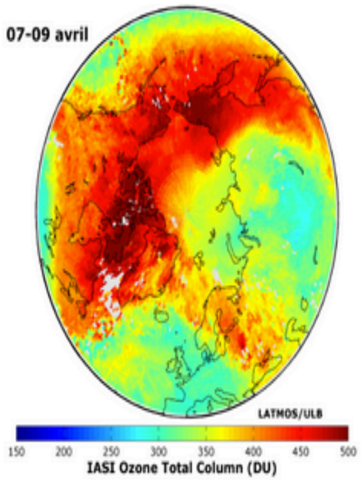
undefined

ERA5 hourly data on pressure levels from 2000 to present

pressure levels". Variables in the dataset are: Divergence, Fraction of cloud cover, Geopotential, **Ozone**

Ozone monthly gridded data from 1970 to present

This dataset provides estimates of the montly mean values of the ozone **concentration**, **mixing ration** and **content** over the globe from a large set of satellite sensors. Most of the ozone data products in this dataset have been developed as part of the ESA Ozone Climate Change Initiative project. They represent the current state-of-the-art in Europe for satellite-based ozone climate data record production, in line with the "Systematic observation requirements for satellite-based products for climate" as defined by GCOS (Global Climate Observing System).



The dataset is organised around the vertical aggregation of the ozone data in four main products:

- Ozone total column retrieval from UV-nadir sensors;
- Ozone total and tropospheric column retrieval from IASI sensors;
- Ozone profile retrieval from UV-nadir sensors;
- Ozone profile retrieval from limb and occultation sensors.

When dealing with satellite data it is common to encounter references to processing levels which describes the amount of processing applied to the raw data, in this case, Level-3 and Level-4. Level-3 means that data are on a regular latitude/longitude expectabily with gaps in space and time. Level-4 data was futher reprocessed in order to fill any eventual gaps in the dataset.

Another common reference is to Climate Data Recors (CDR) and interim-CDR (ICDR). For this datset, both the ICDR and CDR parts of each product were generated using the same software and algorithms. The CDR is intended to have sufficient length, consistency, and continuity to detect climate variability and change. This is the case for instance with the ozone vertically integrated values computed from the passive remote-sensing UV spectrometry onboard of nadir sensors such as SBUV, TOMS, GOME, SCIAMACHY or OMI. The ICDR provides a short-delay access to current data where consistency with the CDR baseline is expected but was not extensively checked. The user is invited to read the documentation in order to determine for each product which are the time spans of the CDR and ICDR parts.

Contact

copernicus-support@ecmwf.int

License

Licence to Use Copernicus Products

Related data

Carbon dioxide data from 2002 to present derived from satellite sensors

Methane data from 2002 to present derived from satellite sensors

Climate Change

University of Reading

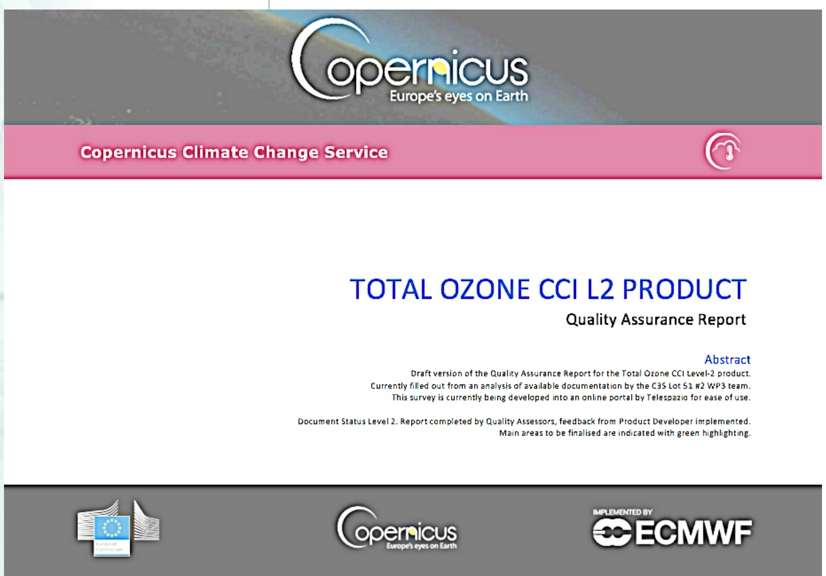
WHAT USERS SAY THEY NEED FOR QA

PRODUCT DETAILS	UNCERTAINTY CHARACTERISATION
PRODUCT GENERATION	VALIDATION
QUALITY FLAGS	INTER-COMPARISON

Climate Change

University of Reading

Quality Assurance Reports


Copernicus
Europe's eyes on Earth
Copernicus Climate Change Service

TOTAL OZONE CCI L2 PRODUCT

Quality Assurance Report

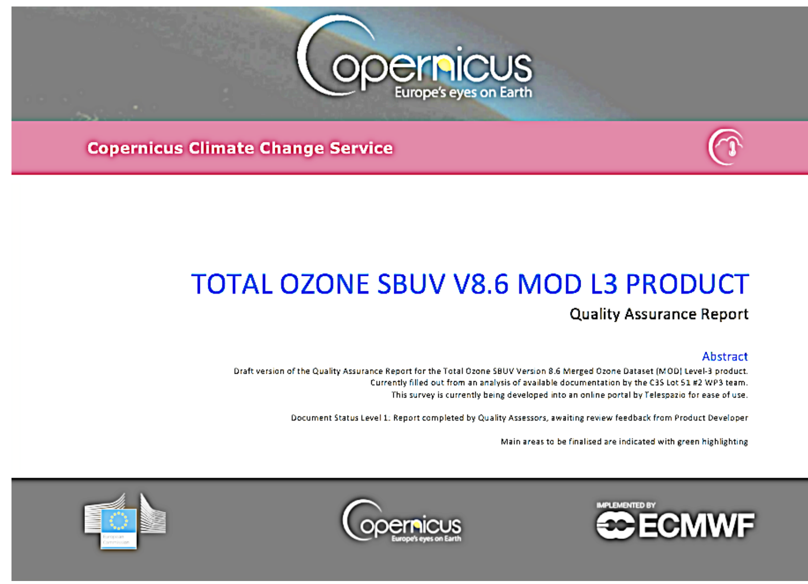
Abstract

Draft version of the Quality Assurance Report for the Total Ozone CCI Level-2 product.
Currently filled out from an analysis of available documentation by the C3S Lot 51 #2 WP3 team.
This survey is currently being developed into an online portal by Telespazio for ease of use.

Document Status Level 2: Report completed by Quality Assessors, feedback from Product Developer implemented.
Main areas to be finalised are indicated with green highlighting.

Copernicus Europe's eyes on Earth

IMPLEMENTED BY ECMWF


Copernicus
Europe's eyes on Earth
Copernicus Climate Change Service

TOTAL OZONE SBUV V8.6 MOD L3 PRODUCT

Quality Assurance Report

Abstract

Draft version of the Quality Assurance Report for the Total Ozone SBUV Version 8.6 Merged Ozone Dataset (MOD) Level-3 product.
Currently filled out from an analysis of available documentation by the C3S Lot 51 #2 WP3 team.
This survey is currently being developed into an online portal by Telespazio for ease of use.

Document Status Level 1: Report completed by Quality Assessors, awaiting review feedback from Product Developer

Main areas to be finalised are indicated with green highlighting

Copernicus Europe's eyes on Earth

IMPLEMENTED BY ECMWF

Climate Change Service

beta6 This is a new service -- your feedback will help us to improve it

Home Search Datasets Toolbox Help & support

Quality report example full fields

- Product details
- Product generation
- Quality flags
- Uncertainty characterisation (content in construction)
- Product validation
- Product inter-comparison

Product information

ECV: Earth radiation budget

Product ID 123

Version number: v1.0

Organisation: My organisation

Point of contact

Name	Email
John Doe	john@dd.com

Processing Level of product: Level 2

Physical Quantity Name: Active Fire Maps

DOI https://climate.copernicus.eu/

Summary description of the product

Timeliness

Primi igitur omnium statuuntur Epigonus et Eusebius ob nominum gentilitatem oppressi. praediximus enim Montium sub ipso vivendi termino his vocabulis appellatos fabricarum culpasse tribunos ut adminicula futurae molitioni pollicitos.

Primi igitur omnium statuuntur Epigonus et Eusebius ob nominum gentilitatem oppressi. praediximus enim Montium sub ipso vivendi termino his vocabulis appellatos fabricarum culpasse tribunos ut adminicula futurae molitioni pollicitos.

Physical Quantity Definition

Primi igitur omnium statuuntur Epigonus et Eusebius o

termino his vocabulis appellatos fabricarum culpasse t

Physical Quantity Units: m

Product status: Completed

Date l

Fri, 05

Are further versions expected?

Yes

- ➢ Standardised form
- ➢ Will be on the C3S Climate Data Store
- ➢ Able to compare quickly against other data products

Coverage and Resolution

Temporal

Record start date: Sun, 05/20/2018 - 12:00

Record end date: Thu, 05/31/2018 - 12:00

Temporal resolution

T	Unit
12	Month(s)
1	Year(s)

Climate Change

University of Reading

GRADING OF QA COMPLETENESS

	WHITE (not provided)	GREY (basic)	BLUE (intermediate)	GREEN (excellent)
DETAILS	No ATBD	Basic ATBD or journal articles	ATBD details input data and all process steps	ATBD conforms to QA4ECV and is up to date
GENERATION	no TC	Incomplete TC available	Complete and detailed TC available	TC with uncertainty sources and propagation completed
QUALITY FLAGS	no QF's	Basic QF's with limited details on derivation and usage	Detailed QF's with good description of each flag and usage	Comprehensive QF's with full descriptions that allow a detailed understanding of data quality
UNCERTAINTY	no uncertainty	uncertainty characterised statistically	uncertainty characterised by values propagated through some parts of the processing chain	full metrological uncertainty characterisation starting from L0 data
VALIDATION	no validation	Validation available, but only basic information provided	Must include either: •Total product bias •Total product standard deviation	Must include at least one total product and one regional bias and standard deviation And full temporal coverage should also be validated
INTER-COMPARISON	No inter-comparison	Inter-comparison available with partial coverage and final output only	Inter-comparison with representative coverage and some internal process comparisons (e.g. cloud screening, RTM choice, etc.)	Full inter-comparison including internal processes and good temporal and spatial coverage, carried out routinely

Methane data from 2002 to present derived from satellite sensors

Overview | Download data | Documentation | Quality Assurance

Quality Evaluation

PRODUCT DETAILS	PRODUCT GENERATION	QUALITY FLAGS	UNCERTAINTY CHARACTERISATION	PRODUCT VALIDATION	PRODUCT INTER-COMPARISON
Product Information	Input data and uncertainties	Quality flags	Uncertainty characterisation method	Reference data representativeness	Scale of inter-comparison
Product Description	Sensor calibration		Percentage of uncertainty sources included	Reference data and uncertainty inclusion	Inter-comparison method
Coverage and Resolution	Algorithm method		Uncertainty values provided	Validation method	Inclusion of product uncertainties
Data gaps	Algorithm tuning		Temporal stability	Validation results	Discrepancy between products identified and resolved
Data set limitations and target	Sensitivity analysis		Geolocation uncertainty		
Documentation	Internal processes				
	Traceability				

KEY
Not provided
Basic
Intermediate
Excellent

Full Quality Assurance Report | PDF

Contact

copernicus-support@ecmwf.int

License

GHG-CCI Licence

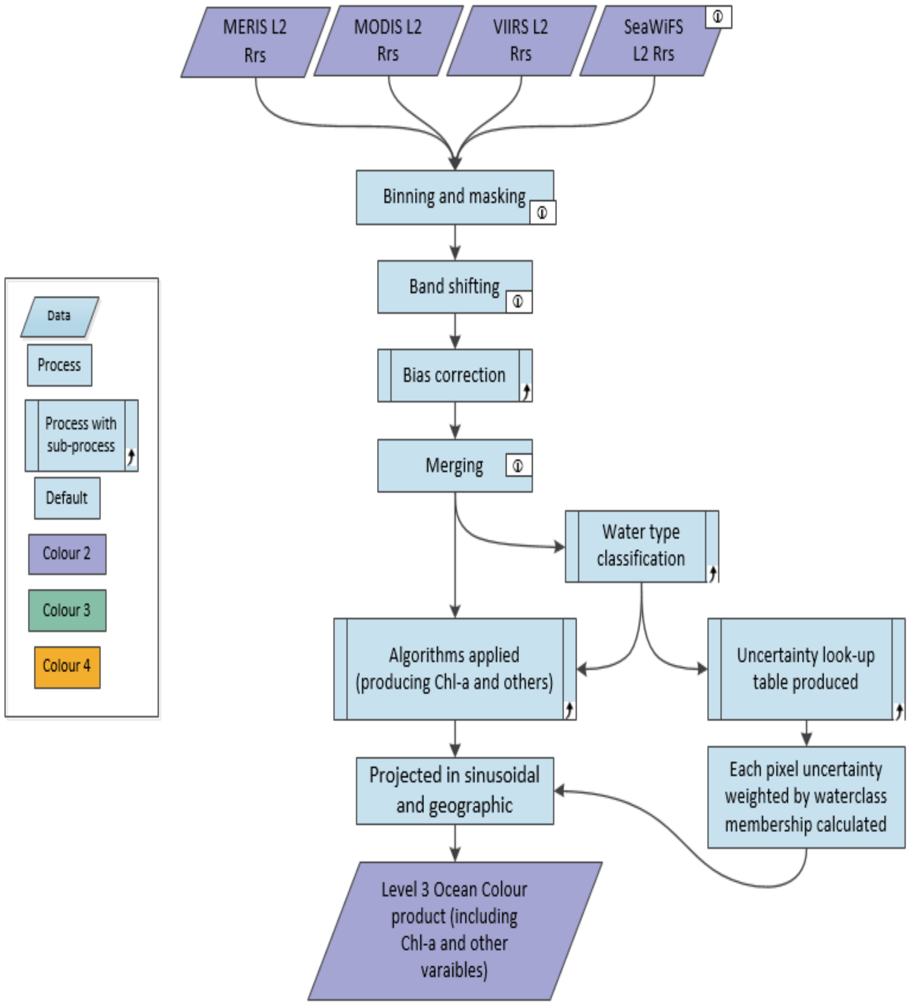
Related data

Carbon dioxide data from 2002 to present derived from satellite sensors

Quality Information

Generation

This traceability chain shows the processes applied to the input data to produce this data set.
Please click on the image for an interactive version.



[Full Quality Assurance Report]

[Generation description]

[Algorithm theoretical basis document (ATBD)]

Quality Flags provided:

1. MERIS_nobs_sum, MODISA_nobs_sum, SeaWiFS_nobs_sum, VIIRS_nobs_sum, total_nobs_sum
2. Chlor_a_log10_bias
3. Chlor_a_log10_rmsd

Uncertainty Characterisation method: Based on validation data

Validation activities:

The Ocean Colour CCI project was validated against several in situ data networks which are commonly used for calibration and validation activities for satellites (MOBY, BOUSSOLE, AERONET-OC, SeaBASS, NOMAD, MERMAID, AMT, ICES, HOT, GeP&CO; Valente et al., 2016). 14582 satellite-in-situ match-ups were analysed finding a small bias and a strong correlation coefficient overall globally. Validation activities also included Chl-a comparisons of global trends with L3 single-sensor satellite data from MERIS, MODIS and SeaWiFS; a long-term analyses in Longhurst biogeographical regions – these are regions of the ocean in which similar characteristics are generally found – including analysis of the time series, phenology and the correlation with climate index (el nino and la nina); and finally a general phenology assessment.

Inter-comparison activities:

No activities currently listed.

Climate Change

University of Reading

SCIENTIFIC GAP ANALYSIS

We carried out an in-depth scientific gap analysis of Level-2 and Level-3 products in different ECV classes focusing on:

- Generation
- Quality flags
- Uncertainty characterisation
- Validation
- Inter-comparison
- Documentation
- Terminology

Essential Climate Variables (defined by GCOS)
Precipitation (in situ)
Surface Air Temp (in situ)
LAI
fAPAR
Sea Surface Temperature
Soil Moisture
Ozone and Aerosols
Ocean Colour

Climate Change

MAIN SCIENTIFIC GAPS FOUND

Uncertainties

- Uncertainties often derived from validation reference data (difference between satellite and *in situ* measurements)

Validation

- Few datasets used uncertainties associated with validation reference data (and often uncertainties on the reference data were also not available)
- Different methodologies were sometimes used for validation activities
- Often only small areas of the total product were represented in the validation activities

Terminology

- Range of uses of quality metric terms, e.g. error and uncertainty
- Often validation reference data is called the "true value"

Climate Change

University of Reading

RECOMMENDATIONS

Uncertainties

- All Level-1 data needs to have associated uncertainties.
- Important retrieval processes need associated data/model uncertainties.
- Validation reference data should provide uncertainties and these should be used in the comparison of the data.

Good practice validation guides for all ECVs

- This exists or is being developed for a few ECVs (e.g. CEOS WGCV LPV is leading on this). Needs to be more widespread to ensure reference data is representative and that the match-up and reference uncertainties are both considered.
- This guide should also clarify the terminology using metrological definitions.

THE FIDUCEO APPROACH

- Metrological starting point = measurement equation
- From this we construct a measurement tree → for uncertainty propagation using the GUM
- Effects tables → characterise physical sources of uncertainty & their correlation properties
- Modeling and bias correction → to remove errors due to each effect
- Allan deviation → to trace noise uncertainty & behaviour
- Adaptive filters and robust statistics → to remove outliers
- Sensitivity coefficients → to propagate Level-0 uncertainties to Level-1
- SNOs and EIV optimisation to harmonise inter-satellite data to produce long-term climate data records of retrieved ECVs

EXAMPLE: post-calibration of TOA radiances

- HgCdTe detectors for the 11μm and 12 μm IR channels of the AVHRR have a quadratic calibration equation ('measurement equation'):

ICT Radiance, L_{ICT}

$$L = a_0 + \frac{(\epsilon + a_1)L_{ICT} - a_2(\bar{C}_S - \bar{C}_{ICT})^2}{\bar{C}_S - \bar{C}_{ICT}}(\bar{C}_S - C_E) + a_3(\bar{C}_S - C_E)^2 + O$$

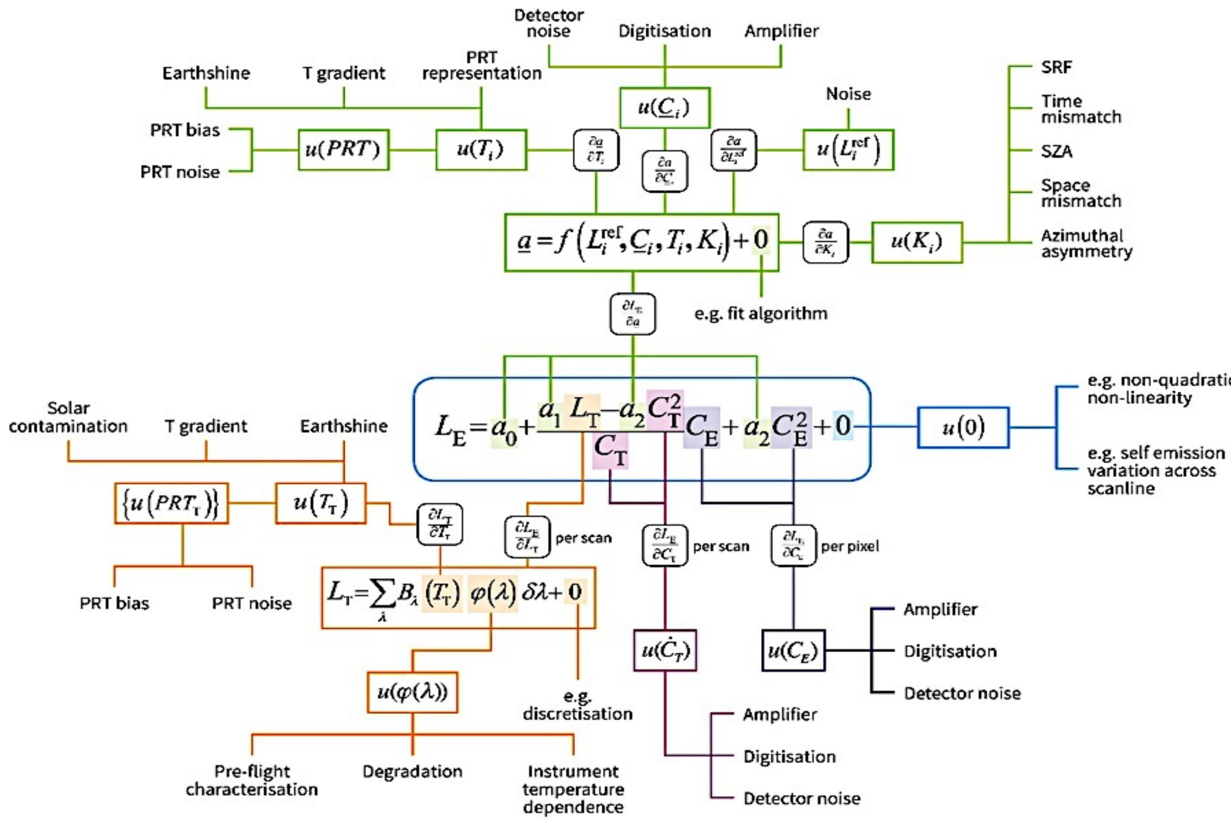
Averaged Space Counts C_S

Averaged ICT Counts C_{ICT}

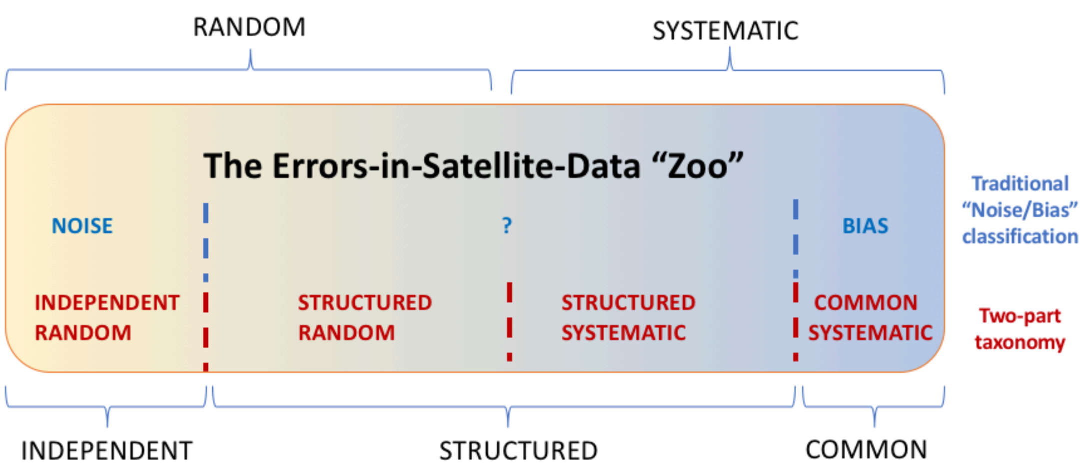
Earth Counts, C_E

- The $+O$ term is important as it captures the residual of the model assumption

We expand the measurement equation to construct a measurement tree → to identify effects and account for uncertainties & sensitivities



We account for independent, structured and common uncertainties



3 uncertainty magnitudes (per-pixel):

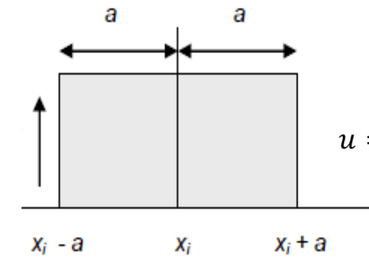
1. **Independent** - due to errors that are random with no spatio-temporal correlation between pixels
2. **Structured** - due to errors from random or systematic processes that have spatio-temporal correlation between pixels
3. **Common** - due to calibration of harmonised radiances

The total uncertainty (independent, structured and common per channel per pixel) is typically **non-Gaussian**

We construct an ‘effects table’ for uncertainty sources

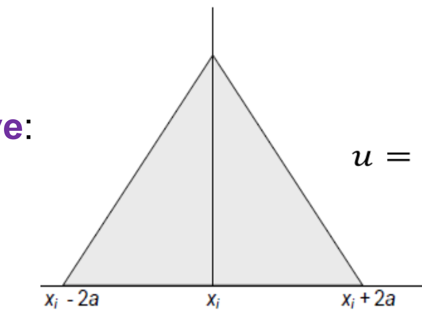
Name of effect		Earth Count Noise	Averaged Space Count Noise	Averaged IWCT Count Noise
Affected term in measurement function		C_E	C_S	C_t
Instruments in the series affected		All	All	All
Correlation type and form	Pixel-to-pixel [pixels]	Random*	Rectangular Absolute	Rectangular Absolute
	from scanline to scanline [scanlines]	Random*	Triangular	Triangular
	between images [images]	N/A	N/A	N/A
	Between orbits [orbit]	Random	Random	Random
	Over time [time]	Random	Random	Random
Correlation scale	Pixel-to-pixel [pixels]	[0]		
	from scanline to scanline [scanlines]	[0]	n = 51	n = 51
	between images [images]	N/A	N/A	N/A
	Between orbits [orbit]	[0]	[0]	[0]
	Over time [time]	[0]	[0]	[0]
Channels/ bands	List of channels / bands affected	All	All	All
	Correlation coefficient matrix	Identity matrix (1s down diagonal only)*	Identity matrix (1s down diagonal only)*	Identity matrix (1s down diagonal only)*
Uncertainty	PDF shape	Digitised Gaussian	Digitised Gaussian	Digitised Gaussian
	units	Counts	Counts	Counts
	magnitude	Provided per pixel	Provided per scanline	Provided per scanline
Sensitivity coefficient		, Eq 4-1	, Eq 4-2	, Eq 4-3

Rectangular Absolute:
Fully systematic or systematic within a calibration period



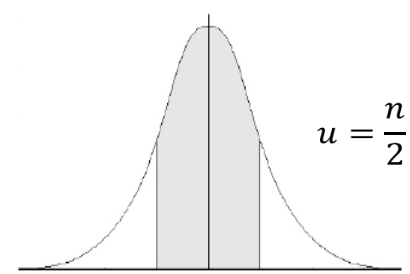
$$u = a/\sqrt{3}$$

Triangular Relative:
Rolling Averages



$$u = a/\sqrt{6}$$

Bell-shape Relative:
Weighted rolling averages, splines, smoothing, other



$$u = \frac{n}{2} - 1/\sqrt{3}$$

We correct for effects and propagate uncertainties using the Guide to the expression of Uncertainty in Measurement (GUM)

$$u_c^2(y) = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}u(x_i, x_j)$$

Adding in quadrature

Sensitivity coefficient times uncertainty

Correlation term

Sensitivity coefficients times covariance

2 because symmetrical

We also calculate correlation matrices and length-scales from codified effects tables and the data

- **mean correlation length-scales** (cross-track, cross-line) for combined structured effects
- **cross-channel error correlation matrices *R*** for independent and structured effects:

Average <> over adequate sample of pixels and invert from covariance to correlation units

$$\boldsymbol{R}_{c,i} = \boldsymbol{U}_{c,i}^{-1} \left\langle \sum_{j} \sum_{i|j} \boldsymbol{C}_c^{p,j} \boldsymbol{U}_c^{p,i} \boldsymbol{R}_c^{p,i} {\boldsymbol{U}_c^{p,i}}^{\mathrm{T}} {\boldsymbol{C}_c^{p,j}}^{\mathrm{T}} \right\rangle_p {\boldsymbol{U}_{c,i}^{-1}}^{\mathrm{T}}$$

Cross-channel error correlation matrix from independent effects

Combined (per-pixel) over all effects in terms of the measurement function

Sensitivity in this channel for term *j*

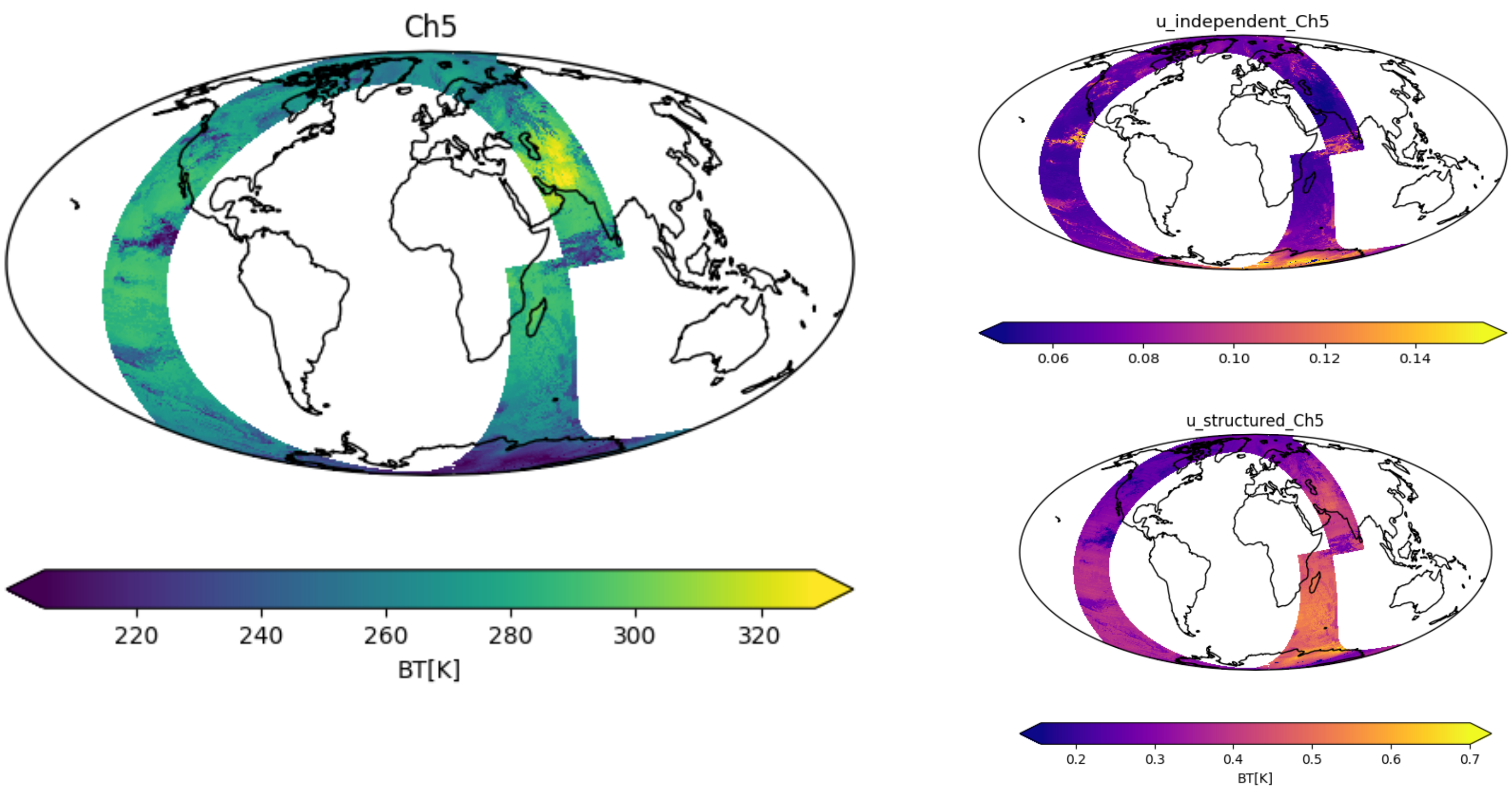
Uncertainty at the pixel from the effect in a given channel

Cross-channel error correlation matrix of a particular independent effect (in codified effects table), *i*, relevant to measurement function term *j*

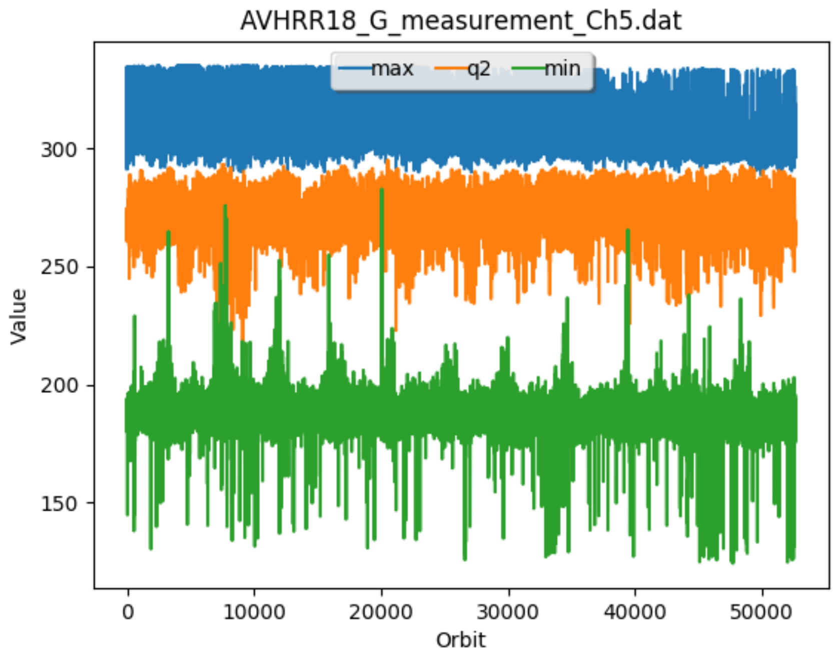
FCDR = **codified effects table** + **satellite data** + **harmonisation data**

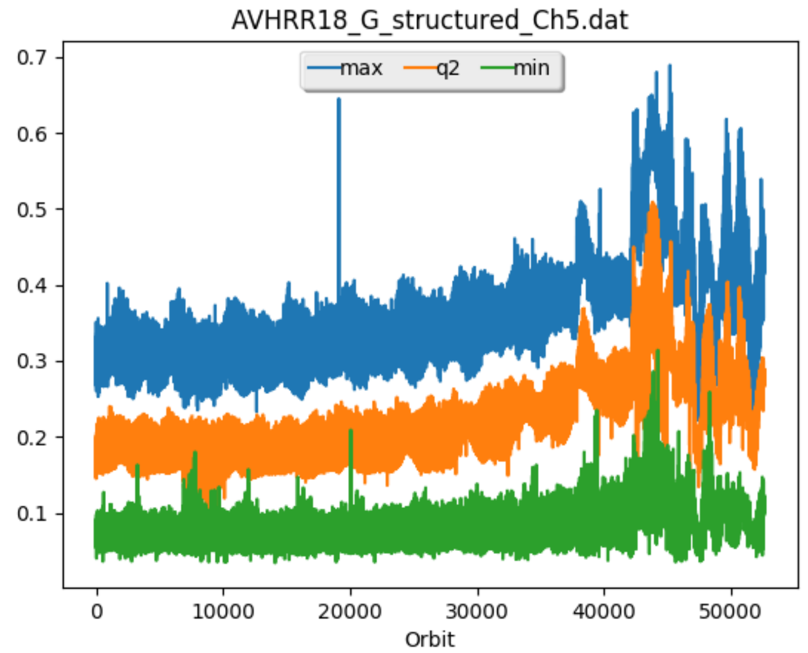
EXAMPLE: orbital uncertainties

AVHRR Ch5 (12µm) on NOAA-11 1992/07/04 10:32-12:14 UCT



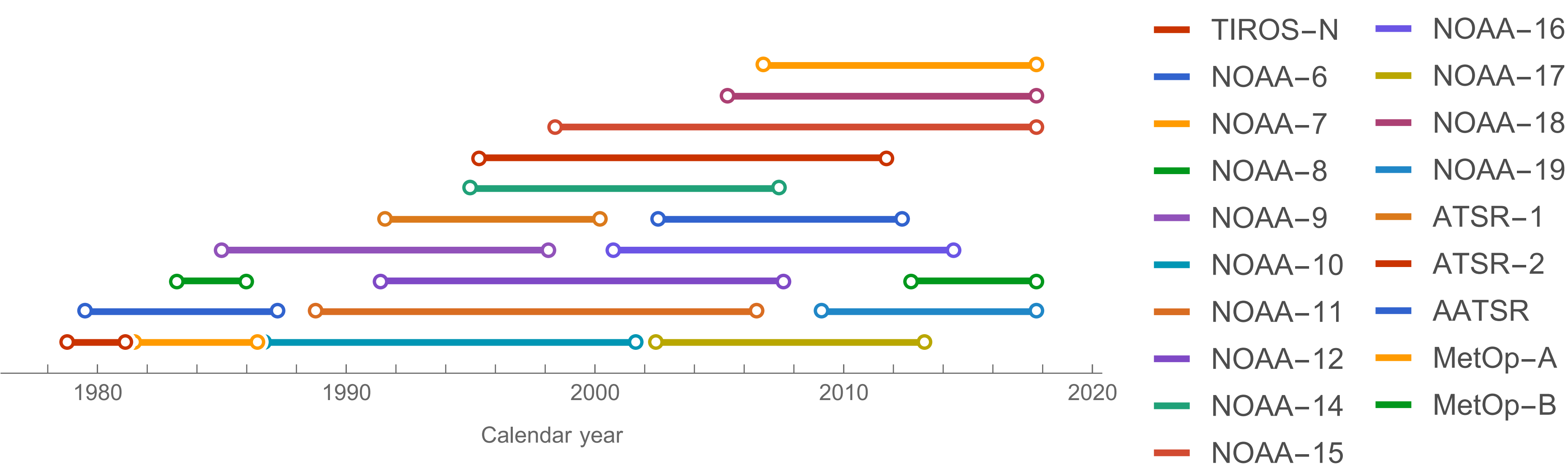
EXAMPLE: drift of systematic uncertainty over time





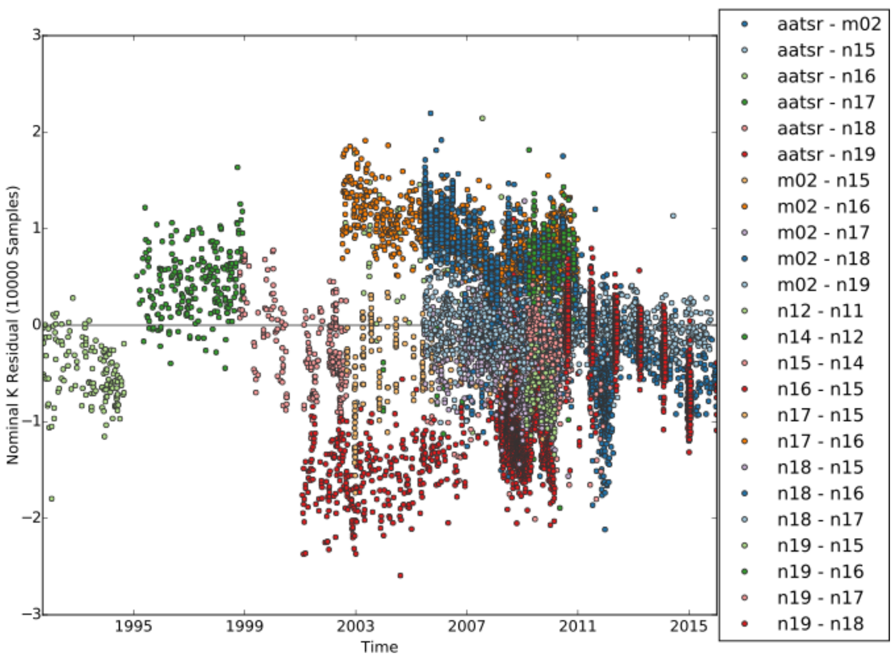
EXAMPLE: harmonisation of inter-satellite data using SNOs & EIV

HARMONISATION: What are the calibration coefficients a_n that minimise the differences between actual and expected inter-sensor differences?

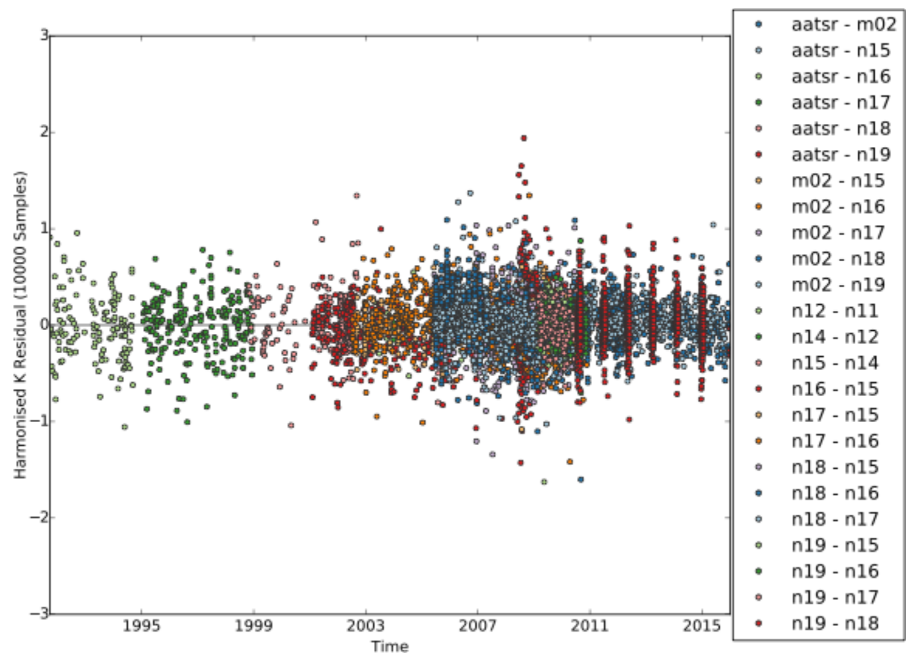


EXAMPLE: harmonisation of inter-satellite data using SNOs & EIV

Nominal Residuals



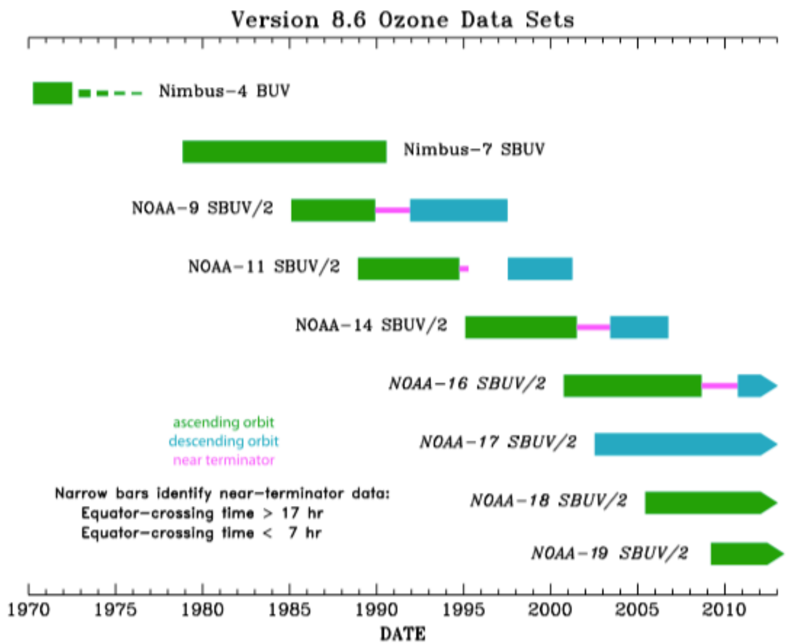
Harmonised Residuals



PROPOSAL: post-calibration and harmonisation of SBUV sensors

- To we use the FIDUCEO approach to post-calibrate and harmonise SBUV/2 radiances and establish a baseline to retrieve metrologically-traceable ozone products? → form a measurement tree for the SBUV
- What resources are needed? → e.g. 20-50 Tb storage (for Level-1c) + parallel computing for orbit processing
- How long might it take? → e.g. ~2 years + 2-3 software engineers
- What are the possible impacts on trend detection? → e.g. due to systematic uncertainty propagation to Level-3 MZM in SBUV v8.6 MOD

SBUV sensor series and Dobson validation



McPeters et al., JGR 118, 8032-9, 2013

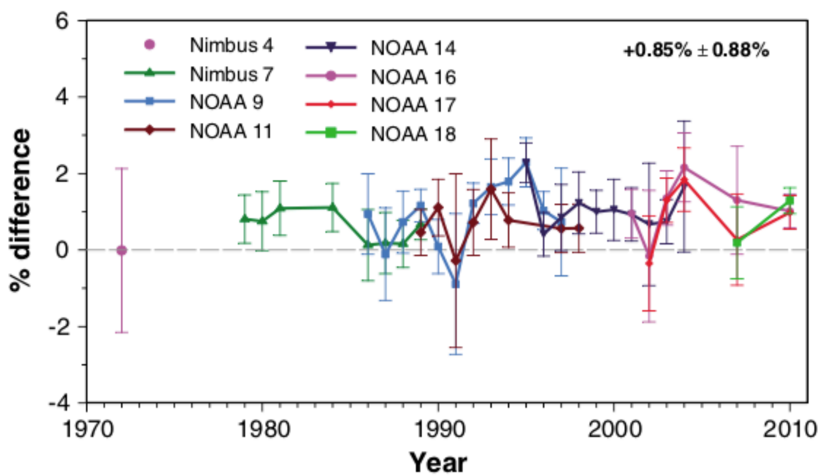
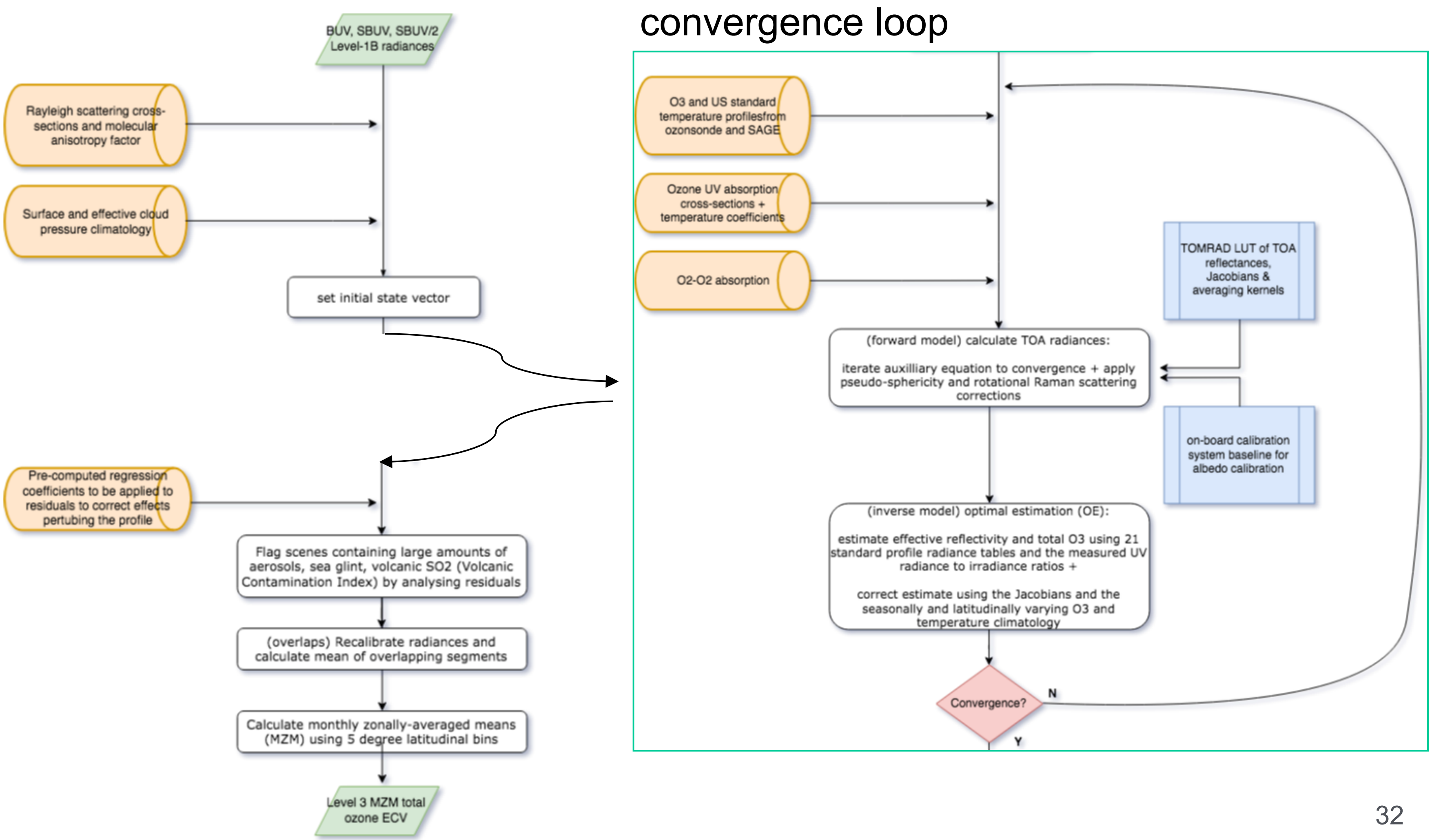


Figure 4. Total ozone from a series of SBUV instruments compared with ozone measured by the World Standard Dobson instrument I83 at Mauna Loa Observatory each year. 2σ standard errors are plotted for each year's comparison.

Top-level process diagram for SBUV v8.6 MOD



Level-2 OE RETRIEVAL

a posteriori retrieval equation

NWP error covariance of the prior value

$$\hat{\mathbf{z}} = \mathbf{z}_a + \mathbf{S}_a\mathbf{K}^T(\mathbf{K}\mathbf{S}_a\mathbf{K}^T + \mathbf{S}_\epsilon)^{-1}(\mathbf{y} - \mathbf{F}(\mathbf{z}_a)) + \mathbf{0}$$

$\partial \mathbf{F}/\partial \mathbf{y}$

Radiances

Assumptions & approximation

$$\mathbf{S}_\epsilon = \mathbf{S}_c + \mathbf{S}_F$$

Forward model (RTTOV)

Prior value (linearisation point)

cross-channel radiance error covariance and forward model error covariance

$$\mathbf{U}_{c,i}\mathbf{R}_{c,i}\mathbf{U}_{c,i}{}^{\mathrm{T}} + \mathbf{U}_{c,s}\mathbf{R}_{c,s}\mathbf{U}_{c,s}{}^{\mathrm{T}} + \mathbf{U}_{c,h}\mathbf{U}_{c,h}{}^{\mathrm{T}}$$

single-pixel uncertainty info from the new SBUV FCDR

University of Reading

CONCLUDING THOUGHTS

- C3S is live: https://climate.copernicus.eu
- C3S EQCO is defining essential info for QA of observational ECVs
- C3S EQCO has designed an evaluation matrix to help users understand product QA
- C3S EQCO raise awareness of metrology in QA to help plug scientific gaps
- FIDUCEO has demonstrated how to generate metrologically-traceable pixel-level uncertainties with correlation info at Level-1 → Level-2+
- FIDUCEO helps improve data quality by applying a metrologically-consistent recalibration and harmonisation prior to OE retrieval
- FIDUCEO has generated FCDRs for VIS, IR and MW sensors and a logical next step would be to apply the methodology to UV radiances e.g. from the SBUV.

Many thanks for listening