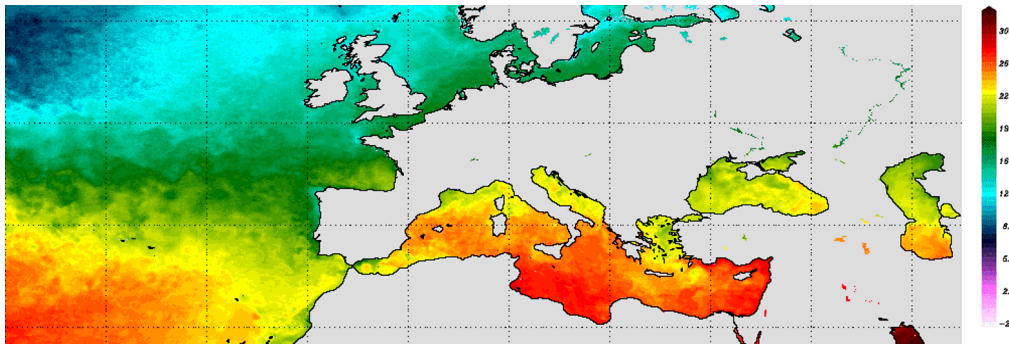# ESA CCI Phase 3 Sea Surface Temperature (SST)

Algorithm Theoretical Basis Document D2.1 v3

Issue Date: **20 January 2023**

Document Ref: SST_CCI_D2.1_ATBD_v3.1

Contract: ESA/AO/1-9322/18/I-NB

SIGNATURES AND COPYRIGHT

Title **:** **ESA CCI Phase 3 Sea Surface Temperature (SST)**

Volume **:** **Algorithm Theoretical Basis Document D2.1 v3**

Issued **:** **20 January 2023**

Authored :

[signature]

_____

Dr Owen Embury (UoR)

Authored : Dr Simon Good (UKMO)

Dr Jacob Høyer (DMI)

Address **:** University of Reading,
Whiteknights,
Reading,
Berkshire,
RG6 6AH,
United Kingdom

TABLE OF CONTENTS

1.   INTRODUCTION

1.1   Purpose and Scope

This is the Algorithm Theoretical Basis Document (ATBD) for the ESA Sea Surface Temperature (SST) Climate Change Initiative (CCI) project. This version covers the algorithms used to generate the version 3 Climate Data Record (CDR) from input satellite radiances to output single-sensor SST and merged multi-sensor SST products.

1.2   Executive summary

This document summarises the algorithms and auxiliary data used in the production of the Level 2, Level 3 and Level 4 Sea Surface Temperature CDR products. The version 3 SST-CCI CDR provides a baseline record from 1980 through end-2021.

The version 3 SST CDR is produced using data from multiple satellite sensor systems. Input sensors are the Along Track Scanning Radiometer (ATSR), the Advanced Very-High Resolution Radiometer (AVHRR), the Sea and Land Surface Temperature Radiometer (SLSTR), and Advanced Microwave Scanning Radiometers (AMSR) instruments. Input satellite brightness temperatures (BTs) from infrared sensors are first cloud screened using the physically-based Bayesian method based on Merchant et al. (2005) and SSTs are retrieved using an Optimal Estimation (AVHRR; Merchant et al. 2008) or coefficient (ATSR, SLSTR; Embury and Merchant 2012) technique. Microwave sensors are processed using the two-step multiple linear regression from Alerskans et al. (2020). As the satellite SSTs are a measure of the temperature of the skin (infrared sensor) or sub-skin (microwave sensor) of the water at the time it was observed an adjustment to a daily mean temperature at 20cm depth is calculated (allowing comparison with the historical in situ record). Quality level and retrieval uncertainty estimates are calculated, and the SSTs are remapped to a regular latitude-longitude grid. These are then used to produce global daily products. This ATBD describes the algorithms used to generate the SST products, including the (1) identification of clear-sky pixels for valid retrieval, (2) the SST retrieval procedure itself, (3) estimating the daily average SST from the instantaneous skin observation, (4) assigning a pixel quality level, (5) remapping the data to a regular global grid, and (6) infilling to remove data gaps.

1.3   Changes from version 2.1

Compared to the previous v2.1 CDR the major changes are:
- Longer time series: 1980 to 2021 (previous CDR was Sept 1981 to 2016)
- Updated radiative transfer model: now using RTTOV 12.3 with "lblrtm_v12.8-aer_v_3.6" coefficients released in October 2020. This version supports variable CO_2 and the CAMS aerosol species used for tropospheric aerosol allowing correction for desert-dust aerosol.
- Improved prior information: Use of CAMS reanalysis and climatology for tropospheric dust aerosol. Use of ERA5 for processing of recent sensors (MetOp AVHRR and SLSTR). All sensors processed using previous CDRv2.1 as the prior SST.

- New bias aware optimal estimation retrieval for single-view AVHRR sensors reduces systematic biases.
- Addition of early AVHRR/1 data in 1980s, and improved AVHRR processing to reduce data gaps in 1980s.
- Addition of dual-view SLSTR data from 2016 onwards.
- Use of full-resolution MetOp AVHRR data (previously used 'global area coverage' Level 1 data).
- Inclusion of L2P passive microwave AMSR data.
- Early Level 4 data (pre-1997) are adjusted relative to HadSST4 to reduce impact of intermittent periods of anomalous satellite calibration.

1.4   Acronyms

Acronym	Definition
AMSR	Advanced Microwave Scanning Radiometer
ARC	ATSR Reprocessing for Climate
ASDI	ATSR Saharan Dust Index
ATBD	Algorithm Theoretical Basis Document
ATSR	Along-Track Scanning Radiometer
AVHRR	Advanced Very-High Resolution Radiometer
BT	Brightness Temperature
C3S	Copernicus Climate Change Service
CAMS	Copernicus Atmosphere Monitoring Service
CCI	Climate Change Initiative
CDR	Climate Data Record
CLASS	Comprehensive Large Array-data Stewardship System
DV	Diurnal Variability
ECMWF	European Centre for Medium-Range Weather Forecasts
EPS	EUMETSAT Polar System
ERA	European Reanalysis
ERS	European Remote Sensing
ESA	European Space Agency
FRAC	Full Resolution Area Coverage
GAC	Global Area Coverage
GDS	GHRSST Data Specification
GHRSST	Group for High Resolution SST
ICDR	Interim Climate Data Record
IR	Infrared
L2	Level-2 data product
L2P	Level-2 Pre-processed data product
L3	Level-3 data product
L3C	Level-3 Collated data product

Algorithm Theoretical Basis Document D2.1 v3

L3U	Level-3 Uncollated data product
L4	Level-4 data product
LECT	Local Equator Crossing Time
NOAA	National Oceanographic and Atmospheric Administration
NWP	Numerical Weather Prediction
OE	Optimal Estimation
OSI SAF	Ocean and Sea Ice Satellite Application Facility
PDF	Probability Density Function
POES	Polar Operational Environmental Satellites
RDAC	Regional Data Assembly Centre (GHRSST)
RFI	Radio Frequency Interference
RTM	Radiative Transfer Model
S3-MPC	Sentinel 3 Mission Performance Centre
SAFE	Standard Archive For Europe
SLSTR	Sea and Land Surface Temperature Radiometer
SST	Sea Surface Temperature
TCWV	Total Column Water Vapour
TOA	Top Of Atmosphere
WS	Wind Speed

2. SATELLITE INSTRUMENTS

2.1 AVHRR

The series of Advanced Very High Resolution Radiometers (AVHRRs) are a series of multipurpose visible and infrared (IR) imaging instruments carried onboard the National Oceanic and Atmospheric Administration (NOAA) Polar Operational Environmental Satellites (POES) and EUMETSAT Polar System (EPS) MetOp satellites. The first AVHRR instrument was carried onboard the TIROS-N satellite launched in October 1978, and the final AVHRR launched onboard MetOp-C in November 2018.

The AVHRR is an across-track scanning radiometer using six spectral channels (early versions of the instrument had four or five channels), with a spatial resolution of approximately 1.1 km at nadir. There are 2048 pixels in each scan for a swath width of about 2800 km. However, due to hardware limitations when the instruments were originally designed it was not possible to record a complete orbit of full resolution data for transmission to the ground station. Therefore, the onboard processor samples the real-time data to produce reduced resolution Global Area Coverage (GAC) data with a nominal resolution of ~4 km. This is achieved by averaging four pixels along the first scanline and then skipping a pixel before averaging the next four pixels. The next two scan lines are discarded before resuming the sampling on the fourth scanline. Each four-pixel average is then considered to be representative of a 15-pixel cell (5 pixels across track by 3 pixels along track) as shown in Figure 1. The more recent MetOp satellites do not have this limitation and record full orbit data at native resolution. Note that the NOAA distributes the full resolution MetOp data as Full Resolution Area Coverage (FRAC).
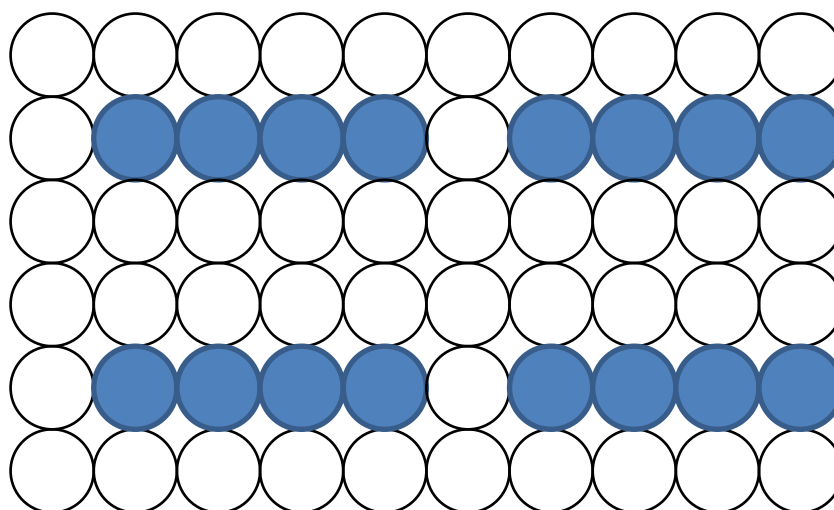

Figure 1: Schematic of GAC pixels (blue) which are the average of four full resolution pixels (circles). White circles indicate full resolutions pixels which are not included in the averaging, so data is not available.

Algorithm Theoretical Basis Document D2.1 v3

Satellite	Sensor	Overpass Time	Operations start	Operations end
TIROS-N	AVHRR/1		Nov 1978	Jan 1980
NOAA-6	AVHRR/1	AM	Jul 1979	Mar 1982
NOAA-7	AVHRR/2	PM	Sep 1981	Feb 1985
NOAA-8	AVHRR/1	AM	May 1983	Oct 1985
NOAA-9	AVHRR/2	PM	Feb 1985	Nov 1988
NOAA-10	AVHRR/1	AM	Nov 1986	Sep 1991
NOAA-11	AVHRR/2	PM	Nov 1988	Dec 1994
NOAA-12	AVHRR/2	AM	Sep 1991	Dec 1998
NOAA-14	AVHRR/2	PM	Jan 1995	Oct 2002
NOAA-15	AVHRR/3	AM	Oct 1998	Dec 2010
NOAA-16	AVHRR/3	PM	Jan 2001	Dec 2010
NOAA-17	AVHRR/3	~10:00 AM	Jun 2002	Dec 2010
NOAA-18	AVHRR/3	PM	May 2005	Ongoing
MetOp-A	AVHRR/3	9:30 AM	Oct 2006	Nov 2021
NOAA-19	AVHRR/3	PM	Feb 2009	Ongoing
MetOp-B	AVHRR/3	9:30 AM	Jan 2013	Ongoing
MetOp-C	AVHRR/3	9:30 AM	Nov 2018	Ongoing

Table 1: List of platforms carrying AVHRR sensors. NOAA platforms are in drifting orbits with initial daytime overpass defined as afternoon or morning. MetOp platforms are in maintained orbits with fixed 9:30 AM equator crossing (while onboard fuel remains).

Channel	Central Wavelength (μm)	AVHRR/1	AVHRR/2	AVHRR/3
1	0.63	X	X	X
2	0.87	X	X	X
3A	1.61			X
3B	3.74	X	X	X
4	10.8	X	X	X
5	12.0		X	X

Table 2: List of AVHRR channels by instrument type. AVHRR can only transmit one of 3A or 3B – typically 3A is used during day and 3B at night; however, some NOAA satellites transmit 3B at all times.

2.2 ATSR

The Along Track Scanning Radiometer (ATSR) instruments are well calibrated, dual-view radiometers intended to produce long-term, consistent SST observations. Three ATSR instruments have flown on board ESA's two European Remote Sensing (ERS) satellites and Envisat satellite. All three satellites were in stable sun-synchronous orbits with near-constant Local Equator Crossing Times (LECTs) – the ERS-1 and ERS-2 platforms had a LECT of 10:30 and Envisat had a crossing time of 10:00 all of which were maintained within a few minutes.

The ATSR instruments had four key design features making them better suited to climate applications than the AVHRR instruments:

- The instrument spectral response functions were accurately measured during pre-flight calibration and characterisation.
- The ATSRs are exceptionally well calibrated with two accurate onboard calibration targets at temperatures of ~260 K and 300 K which greatly reduces non-linearity errors for ocean observations (for comparison AVHRR instruments have a single on-board target at ~290 K and rely on a space view (2.7 K) to provide the second point.
- The infrared detectors are actively cooled to ~82 K to reduce instrument noise and avoid temperature dependent effects on calibration. In addition, the instrument fore-optics are cooled below ambient temperature to reduce self-emission issues, and the instrument is enclosed to prevent stray-light affecting the detectors.
- The dual-view capability using a single telescope with a conical scanning pattern provides both a nadir-view and an inclined forward view (~55°). Having two views of the Earth's surface allows the instrument to gather more information and more effectively separate surface and atmospheric effects; i.e. the SST retrieval can be made more robust to atmospheric conditions, including water vapour and stratospheric aerosol.

The first ATSR carried onboard ERS-1 was a four-channel radiometer with channels are 1.6, 3.7, 11, and 12 μm. However, the ATSR-1 instrument would only transmit one of the 1.6 or 3.7 μm channels with the selection based on the 1.6 μm reflectance intended to separate day and night-conditions. There were two major issues affecting the ATSR-1 instrument:

1. The 3.7 μm channel failed in May 1992, less than a year after the satellite was launched
2. In order to preserve mission lifetime the temperature of the actively cooled detectors were allowed to rise, by-end of mission the detectors were operating at over 110 K.

The second ATSR carried onboard the ERS-2 added three reflectance channels at 0.55, 0.67, and 0.87 μm mainly for vegetation monitoring. The channels had only limited availability over ocean due to telemetry bandwidth limitations; depending on instrument operating mode the visible channels may be transmitted for a reduced narrow swath, reduced 8-bit digitization, or interlaced (record every-other) pixels. The final, Advanced ATSR, instrument onboard Envisat was functionally the same as ATSR-2 but without the bandwidth limitations, so all seven channels are always available in full resolution.

Algorithm Theoretical Basis Document D2.1 v3

Channel	Central Wavelength (µm)	ATSR1	ATSR2	AATSR
1	0.55		X^2	X
2	0.67		X^2	X
3	0.87		X^2	X
4	1.6	X^1	X	X
5	3.7	X^1	X	X
6	10.8	X	X	X
7	12.0	X	X	X

Table 3: List of ATSR channels for each instrument. (1) ATSR1 would only transmit one of 1.6 and 3.7 µm depending on the 1.6 µm intended to separate day and night conditions. However, after the failure of the 3.7 µm channel only 1.6 µm data is transmitted. (2) ATSR2 visible channels have limited availability over ocean resulting in a reduced width swath and/or reduced 8-bit digitisation.

2.3 SLSTR

The Sea and Land Surface Temperature Radiometer (SLSTR) instrument is a well calibrated, dual-view radiometer intended to produce long-term, consistent SST observations. The design of the SLSTR instrument builds on the heritage of the earlier (A)ATSR instruments adding more spectral bands and a wider swath. The first SLSTR instrument is carried onboard the Sentinel-3A satellite launched in February 2016, with the second Sentinel-3B was launched in April 2018.

Channel	Central Wavelength (µm)	Band width (µm)	Spatial resolution (m)
S1	0.555	0.02	500
S2	0.659	0.02	500
S3	0.865	0.02	500
S4	1.375	0.15	500
S5	1.61	0.06	500
S6	2.25	0.05	500
S7	3.74	0.38	1000
S8	10.85	0.9	1000
S9	12.0	1.0	1000

Table 4: List of SLSTR channels

2.4 AMSR

Advanced Microwave Scanning Radiometers (AMSR) instruments have been used on three satellites: ADEOS II, EOS Aqua, and GCOM-W1. ADEOS II only operated for 10 months before suffering from solar panel failure so it not suitable for use in the CDR. The EOS-Aqua AMSR

(known as AMSR-E) was launched in May 2002 and operated for over nine years before failing in October 2011. The GCOM-W1 satellite launched in May 2012 carrying the AMSR2 instrument. Both the AMSR-E and AMSR2 sensors are in early afternoon orbits (1:30 LECT).

The AMSR is a passive microwave radiometer measuring both horizontal and vertical polarisations at six frequencies from 6.9 to 89 GHz to give 12 channels (AMSR2 measures an additional frequency at 7.3 GHz to give 14 channels). The spatial resolution depends on the frequency, with the footprint varying from approximately 5 km at 89 GHz to 50 km at 6.9 GHz as shown in Table 5. The measurements are sampled every 10 km (5 km for 89 GHz) during the instrument scan which means there is significant overlap between adjacent footprints at lower frequencies. The antenna scans in a conical pattern with a satellite zenith (or Earth incident) angle of 55° as shown in Figure 2.

Channel (GHz)	AMSR-E footprint (km)	AMSR2 footprint (km)
6.93	43 x 75	35 x 62
7.3	–	35 x 62
10.65	29 x 51	24 x 42
18.7	16 x 27	14 x 22
23.8	18 x 32	15 x 26
36.5	8 x 14	7 x 12
89.0	4 x 6	3 x 5

Table 5: Approximate spatial resolution (footprint or field of view size) of AMSR channels.
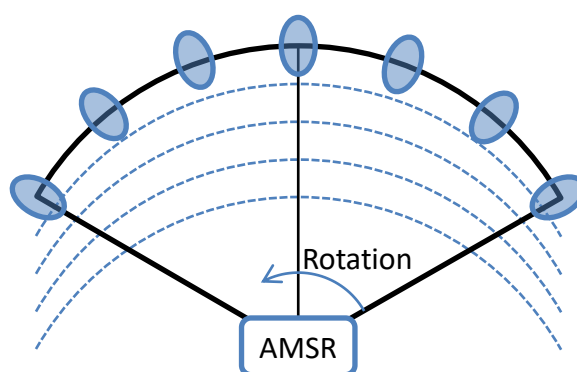
Figure 2: AMSR scanning geometry. Satellite zenith angle of 55° results in an elliptical footprint aligned along the direction of view as the antenna rotates.

Passive microwave SSTs are a lower resolution (approximately 50 km) and lower accuracy than infrared SST measurements. Furthermore, they are affected by Radio Frequency Interference (RFI) and cannot be used within ~100 km of land or sea-ice due to side-lobe contamination (Alerskans et al., 2020). However, they also have advantages:

- Microwave radiometers can retrieve SST through clouds, allowing SST to be measured in all conditions except rain.
- Microwave retrievals of SST are not affected by aerosols and the vertical distribution of water vapour in the same way as infrared retrievals are.

3.   AUXILIARY DATA

3.1   NWP data

Numerical Weather Prediction (NWP) data is used as input to the radiative transfer model (RTM) needed for both cloud detection and SST retrieval algorithms. It is also used by the diurnal variability (DV) model used to adjust the SST to standard time and depth. The required parameters for the various algorithms are shown in Table 6 below.

The ERA5 reanalysis (Hersbach et al., 2020) is the main NWP input used for MetOp and SLSTR processing. While, the older ERA-Interim reanalysis (Dee et al., 2011) is used for the ATSR and NOAA AVHRR processing as the earlier ERA5 data were not available at the time the algorithm development occurred for these sensors.

Parameter	Type	Algorithms
Atmospheric temperature	Analysis, profile	RTM
Atmospheric water vapour	Analysis, profile	RTM
Surface pressure	Analysis, surface	RTM
Mean sea level pressure	Analysis, surface	RTM
10m wind U-component	Analysis, surface	RTM, DV, MW
10m wind V-component	Analysis, surface	RTM, DV, MW
2m air temperature	Analysis, surface	RTM
2m dew point temperature	Analysis, surface	RTM
Sea surface temperature[1]	Analysis, surface	RTM, DV
Skin temperature[1]	Analysis, surface	RTM
Total cloud cover	Analysis, surface	Bayes
Total Column Water Vapour	Analysis, surface	Bayes, ARC, OE
Sea ice fraction	Analysis, surface	DV
Surface Sensible Heat Flux	Forecast, surface	DV
Surface Latent Heat Flux	Forecast, surface	DV
Surface Solar Radiation	Forecast, surface	DV
Surface Thermal Radiation	Forecast, surface	DV
East-West Surface Stress	Forecast, surface	DV
North-South Surface Stress	Forecast, surface	DV

Table 6: NWP inputs to algorithms: RTM – Radiative Transfer Model; Bayes – Bayesian Cloud Detection; OE – Optimal Estimation SST retrieval; ARC – Coefficient based SST retrieval; MW – Microwave SST retrieval; DV – Diurnal Variability model. Note: Both Bayes and OE algorithms use the output from the RTM. (1) Also see section 3.4.

3.2   Tropospheric aerosol

In addition to the NWP data described in previous section3.1, the RTM can also make use of aerosol data. These are taken from the Copernicus Atmosphere Monitoring Service (CAMS)

reanalysis (Inness et al., 2019) and the CAMS aerosol climatology (Bozzo et al., 2020) as the reanalysis is only available from 2003 onwards.

The CAMS aerosol climatology provides monthly layer-integrated mass concentrations (kg/m^2) for 11 aerosol components on 60 model levels and 3° × 3° horizontal resolution. The components are: Sea Salt bin1 (0.03-0.5 µm); Sea Salt bin2 (0.5-5.0 µm); Sea Salt bin3 (5.0-20.0 µm); Mineral Dust bin1 (0.03-0.55 µm); Mineral Dust bin2 (0.55-0.9 µm); Mineral Dust bin3 (0.9-20.0 µm); Organic Matter (hydrophilic); Organic Matter (hydrophobic); Black Carbon (hydrophilic); Black Carbon (hydrophobic); (tropospheric) Sulfates. Mineral dust is the most significant component for SST retrieval as it can cause BT impacts of more than 1 K, while the other components typically cause impacts of ~0.1 K.

The CAMS aerosol climatology used as the primary tropospheric aerosol input, with all aerosol component profiles passed to the RTM. When CAMS reanalysis data are available (i.e. 2003 to 2021) the climatological mineral dust profiles are scaled to match the daily total column dust quantities from the reanalysis.

3.3   Stratospheric aerosol

There were three major volcanic eruptions during the CDR period: El Chichón (1982), Mount Pinatubo (1991), and Mount Hudson (1991); which resulted in elevated levels of stratospheric sulfate aerosol with impacts on the infra-red channels used for SST retrieval. Although the dual-view ATSR and SLSTR sensors are capable of an aerosol robust SST retrieval (e.g. Embury and Merchant, 2012), the single-view AVHRR sensor will be biased unless we include appropriate stratospheric aerosol loadings in the RTM calculations. An auxiliary dataset of stratospheric aerosol was derived from the High-Resolution Infrared Radiation Sounders (HIRS) for the ESA SST CCI CDR version 2 1.   Dee, D. P. *et al.* The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.* **137**, 553–597 (2011).

2.   Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049 (2020).
3.   Merchant, C. J. *et al.* Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci Data* **6**, 223 (2019).
4.   Merchant, C. J. & Embury, O. Adjusting for Desert-Dust-Related Biases in a Climate Data Record of Sea Surface Temperature. *Remote Sensing* **12**, 2554 (2020).
5.   Inness, A. *et al.* The CAMS reanalysis of atmospheric composition. *Atmos. Chem. Phys.* **19**, 3515–3556 (2019).
6.   Bozzo, A., Benedetti, A., Flemming, J., Kipling, Z. & Rémy, S. An aerosol climatology for global models based on the tropospheric aerosol scheme in the Integrated Forecasting System of ECMWF. *Geosci. Model Dev.* **13**, 1007–1034 (2020).
7.   Matson, M. The 1982 El Chichón Volcano eruptions — A satellite perspective. *Journal of Volcanology and Geothermal Research* **23**, 1–10 (1984).
8.   Lambert, A. *et al.* Measurements of the evolution of the Mt. Pinatubo aerosol cloud by ISAMS. *Geophysical Research Letters* **20**, 1287–1290 (1993).
9.   Merchant, C. J., Harris, A. R., Murray, M. J. & Závody, A. M. Toward the elimination of bias in satellite retrievals of sea surface temperature: 1. Theory, modeling and interalgorithm

comparison. *J. Geophys. Res.* **104**, 23565–23578 (1999).

10.   Baran, A. J. & Foot, J. S. New application of the operational sounder HIRS in determining a climatology of sulphuric acid aerosol from the Pinatubo eruption. *J. Geophys. Res.* **99**, 25673–25679 (1994).

11.   OSI SAF. Global Sea Ice Concentration Climate Data Record v2.0 - Multimission. *EUMETSAT SAF on Ocean and Sea Ice* 9.6 MB (2017) doi:10.15770/EUM_SAF_OSI_0008. (Merchant et al., 2019) using a method based on Baran and Foot (1994).
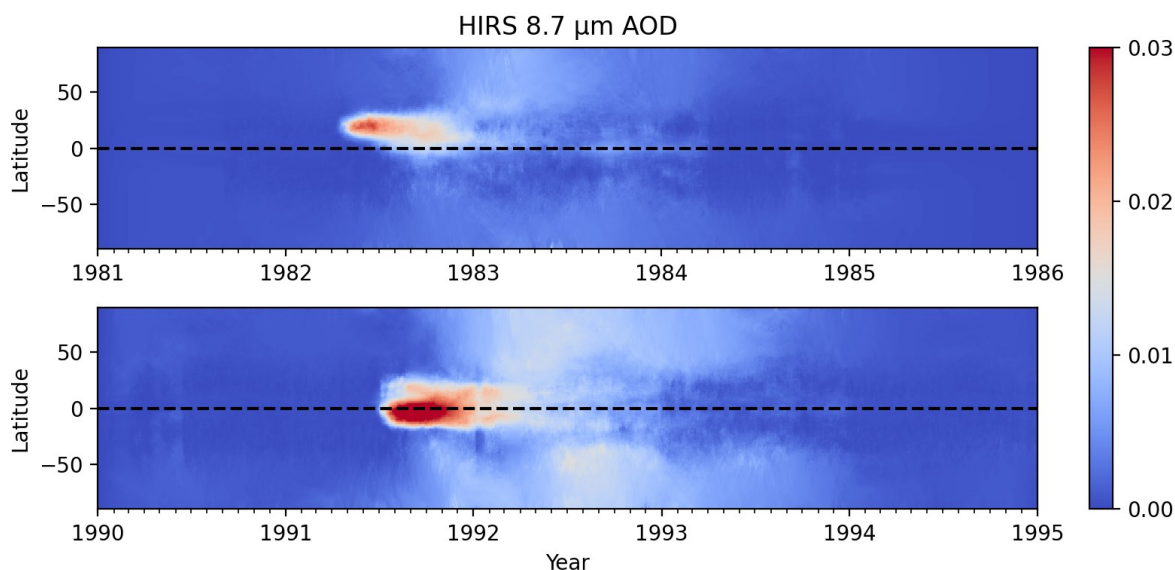

Figure 3: Stratospheric aerosol prior used for CDR processing.

3.4   Prior SST estimate

The prior surface temperature estimate is one of the primary inputs to the RTM. Although SST data is included with the ECMWF NWP (section 3.1), this is actually the prescribed input used as the lower boundary condition for the ECMWF analysis rather than a model output. The source of the SST data included with the NWP depends on the analysis version (ERA-interim or ERA5) and the time-period (ERA-interim used five different sources of SST over the years). In order to maximize consistency in the CDR processing we will instead use the previous version 2.1 of the SST CDR with the desert-dust correction from Merchant and Embury (2020) as the prior when available. However, as the v3.0 CDR targets a longer time-period than version 2.1 it is necessary to fall back to the SST provided with the NWP at the start and end of the processing as shown in Table 7.

Time Period	SST prior	Notes
Jul 1979 – Aug 1981	NWP (ERA-interim: HadISST)	Monthly 1°
Sept 1981 – Dec 2021	SST CDR v2.1 with dust correction	Daily 0.05°
Jan 2022 – onwards	NWP (ERA-5: OSTIA)	Daily 0.25°

Table 7: Prior SST used for CDR processing

3.5   SST climatology

A daily SST climatology is used in the Level 4 processing. This was calculated by averaging the ESA SST CCI version 2.1 analyses for the period (1982 – 2010). A dust correction was applied to limit the impact of dust aerosol on the climatology (Merchant and Embury, 2020).

3.6   SST background error covariance estimates

The background error covariances are specified in the Level 4 analysis system by two components. Associated with each of these are error standard deviations at each grid point and a length scale. One set of standard deviations and its scale are used to represent error correlations at mesoscales and the other at synoptic scales. The relative size of the two error standard deviations at a grid point controls the effective length scale of the combination of the two. These background error covariance estimates are specified seasonally and are described in Roberts-Jones et al. (2016).

3.7   Sea ice concentration data

Sea ice concentration data from the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF) are used by the L4 analysis system. Two datasets are used: OSI-450 (sea ice concentration CDR) to end of 2015, and OSI-430-b (sea ice concentration ICDR) from January 2016 onwards.

The details of each of the OSI SAF products are:

Product	OSI-450
Title	Global Sea Ice Concentration Climate Data Record 1979-2015 (v2.0, 2017) (SMMR/SSMI/SSMIS)
DOI	[10.15770/EUM_SAF_OSI_0008](10.15770/EUM_SAF_OSI_0008)
Product	OSI-430-b
Title	Global Sea Ice Concentration Interim Climate Data Record release 2
DOI	[10.15770/EUM_SAF_OSI_NRT_2008](10.15770/EUM_SAF_OSI_NRT_2008)

3.8   HadSST4

The Met Office Hadley Centre's sea surface temperature data set, HadSST.4.0.1.0 (Kennedy et al., 2019), is used to reduce the impacts of intermittent satellite calibration artefacts in the final Level 4 SST product during the 1980s and early 90s. HadSST4 is a monthly in-situ-based analysis on a 5° latitude-longitude grid.

4.   ALGORITHMS

The major processing steps and data flows from Level 1 to IR SST products are shown in Figure 4, and Figure 5 for microwave products. Microwave sensors are not affected by clouds in the same way as IR so the cloud detection step is not required. Furthermore, Level 3 microwave products are not currently produced due to the resolution of the microwave data. Details of the algorithms for each step are in the following subsections.

Satellite imagery is read from the input Level 1b, and auxiliary data (NWP etc.) is interpolated to the location of the satellite data to run the radiative transfer (4.1) and diurnal variability (4.4) steps. Output from the radiative transfer is used in the IR cloud detection (4.2) and SST retrieval (4.3.2), while microwave SST retrieval (4.3.3) uses the NWP and L1b directly. The retrieved SSTs are combined with the diurnal variability calculations and then quality levels are determined (4.5). This results in the final Level 2P data for microwave sensors which is then passed to the analysis scheme (4.9). For IR sensors and additional atmospheric correction smoothing step (4.6) is applied in the L2P data, and the unsmoothed data are remapped (4.7) to the Level 3 grid to produce Level 3U data which is collated (4.8) and passed to the analysis scheme (4.9) to produce Level 4 data. The Level 4 data are post-processed to reduce the impact of intermittent satellite calibration issues in the 1980s and early 90s (4.10).
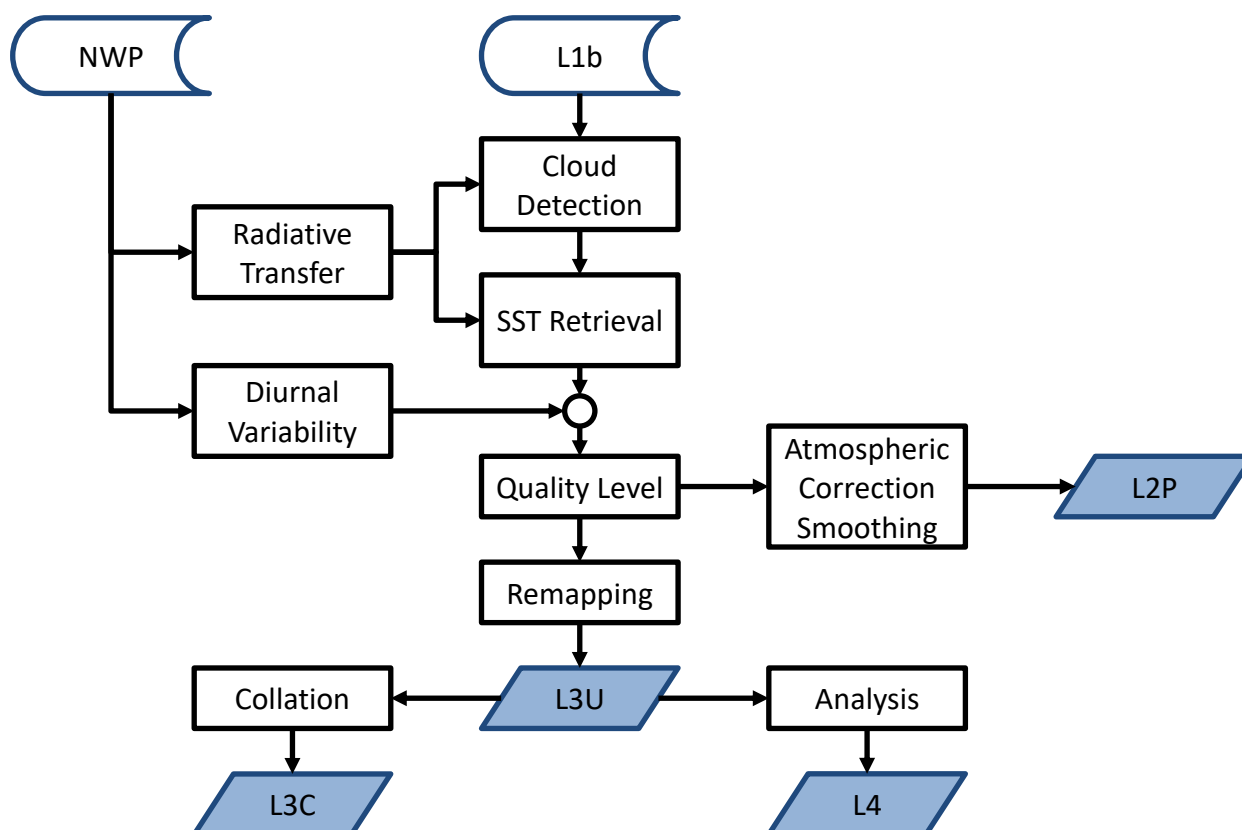

Figure 4: Overview of algorithm steps as applied to different processing levels for IR sensors.
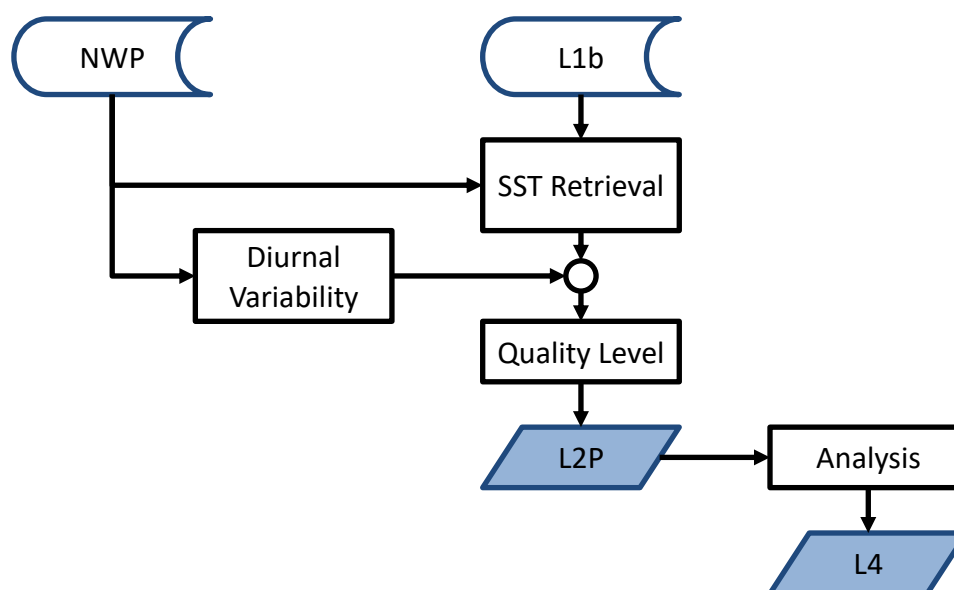
Figure 5: Overview of algorithm steps as applied to microwave sensors.

4.1 Radiative Transfer

Both the Bayesian cloud detection (section 4.2) and optimal estimation retrieval (section 4.3.2) depend on a radiative transfer or forward model to estimate the expected clear-sky BTs based on the available prior information (NWP, aerosol, and prior SST). The CDRv3 processing uses the RTTOV version 12.3 software (Saunders et al., 2018) using the "lblrtm_v12.8/aer_v_3.6" coefficients released in October 2020. This version supports variable CO_2 and the CAMS aerosol species used for tropospheric aerosol (section 3.2).

For the AVHRR instruments an additional forward model bias correction term is applied to ensure consistency between sensors. The bias correction parameters were estimated offline using the bias-aware optimal estimation method of Merchant et al. (2020b) applied to AVHRR as documented in Merchant et al. (2020a) and Appendix A. The MetOp AVHRR sensors were harmonized against ATSR/SLSTR, while the NOAA AVHRRs were harmonized against in situ observations (see Appendix A for details).

4.2 Cloud Detection

Cloud screening is a fundamental pre-processing step for SST retrieval. Traditionally, threshold-based techniques have been used to detect cloud but these often fail under difficult circumstances – for example, in the detection of thin cirrus or low-level fog. The Bayesian cloud detection algorithm presented here was developed originally for ATSR instruments (Merchant et al., 2005) and used in the generation of the ATSR Reprocessing for Climate (ARC) CDR (Merchant et al., 2012) and SST-CCI Phase-I CDR (Merchant et al., 2014). In the SST-CCI Phase-II CDR the Bayesian algorithm is also applied to AVHRR data (Bulgin et al., 2018).

The Bayesian classifier attempts to make optimum use of all available information to calculate the posteriori probability that a given pixel is clear-sky. Available information includes the satellite observations, prior information about the atmospheric and surface condition in the form of NWP analysis/forecast, and the respective errors in these variables. Applying Bayes' theorem to the probability of clear-sky $P(c|\boldsymbol{y}, \boldsymbol{x}_a)$ given both the satellite observations $\boldsymbol{y}$, and the background state $\boldsymbol{x}_a$ yields:

$$P(c|\boldsymbol{y}, \boldsymbol{x}_a) = \frac{P(\boldsymbol{y}|\boldsymbol{x}_a, c)P(\boldsymbol{x}_a|c)P(c)}{P(\boldsymbol{y}|\boldsymbol{x}_a)P(\boldsymbol{x}_a)} \qquad \text{Eq. 1}$$

It is assumed that the background state is independent of the clear-sky probability at the satellite pixel scale so $P(\boldsymbol{x}_a|c) = P(\boldsymbol{x}_a)$. Next the probability of observations given the background state, $P(\boldsymbol{y}|\boldsymbol{x}_a)$, can be expressed as the sum of the probabilities for each possible state. For a binary classifier the possible states are clear (c) and not-clear ($\bar{c}$) giving:

$$P(c|\boldsymbol{y}, \boldsymbol{x}_a) = \frac{P(\boldsymbol{y}|\boldsymbol{x}_a, c)P(c)}{P(\boldsymbol{y}|\boldsymbol{x}_a, c)P(c) + P(\boldsymbol{y}|\boldsymbol{x}_a, \bar{c})P(\bar{c})} \qquad \text{Eq. 2}$$

In order to estimate the clear-sky probability we need to evaluate or estimate the probabilities on the right-hand side of the equation. The probability of the observation given the background state and assuming clear-sky, $P(\boldsymbol{y}|\boldsymbol{x}_a, c)$, can be estimated using a forward model. In principle the probability of an observation for non-clear conditions, $P(\boldsymbol{y}|\boldsymbol{x}_a, \bar{c})$, can also be calculated using a forward model but the distribution is much more complex than the clear-sky case, so a pre-generated lookup table (LUT) is used instead. The probability of clear-sky, $P(c)$, and cloudy-sky, $P(\bar{c})$, can be taken from a climatology or from NWP cloud fraction.

4.2.1 Clear-sky Spectral PDF

Under clear-sky conditions the probability of observations can be calculated using a forward model, $F(\boldsymbol{x}_a)$, and assuming Gaussian background and observation errors following Rodgers (1976):

$$P(\boldsymbol{y}|\boldsymbol{x}_a, c) = \frac{\exp\left(-0.5\big(\boldsymbol{y} - F(\boldsymbol{x}_a)\big)^T (\mathbf{K}\mathbf{S}_a\mathbf{K}^T + \mathbf{S}_\varepsilon)^{-1}\big(\boldsymbol{y} - F(\boldsymbol{x}_a)\big)\right)}{(2\pi)^{\frac{n}{2}}|\mathbf{K}\mathbf{S}_a\mathbf{K}^T + \mathbf{S}_\varepsilon|^{\frac{1}{2}}} \qquad \text{Eq. 3}$$

4.2.2 Cloudy-sky Spectral PDF

For cloudy-sky conditions the probability of observations, $P(\boldsymbol{y}|\boldsymbol{x}_a, \bar{c})$, is taken from a pre-generated lookup tables as described in ATBDv2. Details for the ATSR tables are given in ATBDv2 section 4.2.1.2, and AVHRR tables in section 4.3.2 and Bulgin et al. (2018).

4.3   SST Retrieval

Three different SST retrieval algorithms are used: the ARC algorithm for dual-view sensors (ATSR, SLSTR), Optimal Estimation for single-view sensors (AVHRR), and statistical regression method for microwave sensors (AMSR).

4.3.1 ARC SST retrieval (dual-view sensors)

For dual-view instruments the SST retrieval is based on the techniques developed for the ARC CDR (Embury et al., 2012b; Embury and Merchant, 2012; Embury et al., 2012a). The ARC SST retrieval is a coefficient-based retrieval of the form:

$$\hat{x} = a_0 + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y} \qquad\qquad \text{Eq. 4}$$

Where a_0 is the offset coefficient and $\boldsymbol{a}^{\mathrm{T}} = [a_1, \dots, a_n]$ is a vector of n weighting coefficients for the n BTs in the observation vector $\boldsymbol{y}$.

The coefficients are pre-calculated using least-squares minimization techniques from accurate line-by-line radiative transfer simulations for a range of parameters:
* Satellite zenith angle in the nadir view
* Satellite zenith angle in the forward (or oblique) view
* Prior TCWV
* Instrument detector temperature (ATSR1 only)
* Year

The uncertainty due to radiometric noise (assumed uncorrelated between pixels) is given by:

$$u_{unc} = \sqrt{\boldsymbol{a}^{\mathrm{T}}\mathrm{S}_o\boldsymbol{a}} \qquad\qquad \text{Eq. 5}$$

The uncertainty in the retrieval (assumed correlated on synoptic scales) is given by the fitting error when generating the coefficients. It is tabulated as a function of TCWV and interpolated to the pixel conditions along with the retrieval coefficients (Embury and Merchant, 2012; Bulgin et al., 2016a).

4.3.2 Optimal Estimation (single-view sensors)

For single-view instruments, the SST was retrieved using an optimal estimation (OE) scheme (Merchant et al., 2008; Rodgers, 2000):

$$\hat{x} = x_a + \mathbf{G}\big(\boldsymbol{y} - F(\boldsymbol{x}_a)\big)$$

$$\mathbf{G} = (\mathbf{K}^{\mathrm{T}}\mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S}_a^{-1})^{-1}\mathbf{K}^{\mathrm{T}}\mathbf{S}_\varepsilon^{-1} \qquad\qquad \text{Eq. 6}$$

The retrieved state $\hat{x}$ is the prior state plus an increment of $\mathbf{G}\big(\boldsymbol{y} - F(\boldsymbol{x}_a)\big)$. The matrix $\mathbf{K}$ expresses how the observations change for departures from the prior state x_a, i.e., it is a

matrix where a given row contains the partial derivatives of the BT in a particular channel with respect to each element of the state vector in turn. The partial derivatives are the tangent linear outputs from the forward model $\mathbf{F}$. $\mathbf{S}_\varepsilon$ is the error covariance of the differences between the model and observed BTs. These include the radiometric noise in the observations ($\mathbf{S}_o$) and estimated uncertainty of the forward model ($\mathbf{S}_m$). $\mathbf{S}_a$ is the error covariance matrix for the prior state variables.

The uncertainty due to radiometric noise (assumed uncorrelated between pixels) and due to uncertainty in the retrieval (assumed correlated on synoptic scales) is given by:

$$u_{unc} = \sqrt{\mathbf{G}\mathbf{S}_o\mathbf{G}^{\mathrm{T}}}$$

$$u_{cor} = \sqrt{\mathbf{G}\mathbf{S}_m\mathbf{G}^{\mathrm{T}} + (\mathbf{G}\mathbf{K}-\mathbf{I})\mathbf{S}_a(\mathbf{G}\mathbf{K}-\mathbf{I})^{\mathrm{T}}}$$

Eq. 7

4.3.3 Microwave SST retrieval

The microwave SST retrieval algorithm is the two-step multiple linear regression (Alerskans et al., 2020) with retrieval of wind speed (WS) and then SST. Two stages are required as the relationship between retrieved parameters (WS, SST) and BT is less linear than the IR case. In the first stage a first-guess wind speed is estimated from the observed BTs with the equation:

$$WS_a = a_0 + \sum_{i=1}^{10}(a_{1i}t_i + a_{2i}t_i^2) + a_3(\theta_{sat} - 55)$$

Eq. 8

Where a_0, a_{1i}, a_{2i}, and a_3 are the global wind speed regression coefficients with index i representing the channel number (up to 10 as the 89.0 GHz is not used in the wind speed retrieval). The BTs are in the vector $\boldsymbol{t}$ following $t_i = BT_i - 150$ for all channels except the 23.6 GHz channels where it is $t_i = \ln(290 - BT_i)$. Finally, θ_{sat} is the satellite zenith (or incidence) angle.

For the second stage of the wind speed retrieval, a localised regression is needed. This uses the same form as Eq. 8, except replacing the coefficients a with a different set of regression coefficients which are tabulated as a function of the first guess wind speed. The retrieval is calculated using the two nearest bins to the first-guess, WS_a, and then linearly interpolated.

The retrieved wind speed is then used in the SST retrieval. With the first-guess SST given by:

$$SST_c = c_0 + \sum_{i=1}^{12}(c_{1i}t_i + c_{2i}t_i^2) + c_3(\theta_{sat} - 55) + c_4 WS$$
$$+ \sum_{j=1}^{2}(c_{5j}\cos(j\phi_{rel}) + c_{6j}\sin(j\phi_{rel}))$$

Eq. 9

Where the initial SST retrieval coefficients c are tabulated as a function of latitude and orbit direction (ascending or descending), WS is the retrieved wind speed, and ϕ_{rel} is the relative angle between the satellite azimuth angle and the wind direction (taken from the NWP). The second stage of the SST retrieval uses Eq. 9 again with the coefficients c replaced with a new set of localised coefficients which are tabulated as a function of the retrieved wind speed and first-guess SST. The final retrieved SST is bilinearly interpolated between the closest bins. The binning and interpolation used for the four sets of coefficients is listed in Table 8.

Coefficients	Binning	Interpolation
First-guess WS_a	None (global)	N/A
Retrieved WS_r	WS_a	Linear
First-guess SST_a	Latitude, orbit	None
Retrieved SST_r	SST_a, WS_r	Bi-linear

Table 8: Coefficients used in passive microwave retrievals and binning interpolation used.

The retrieval uncertainties are given by another regression equation:

$$u = e_0 + e_1 SST_r + e_2 SST_r^2 + e_3 WS_r + e_4 WS_r^2 + e_5 \theta_{sat} + e_6 \theta_{sat}^2$$

$$+ \sum_{p=1}^{4} \left(e_{7p} \cos \frac{\varphi_{lat}}{p} + e_{8p} \sin \frac{\varphi_{lat}}{p} \right)$$ Eq. 10

Where φ_{lat} is the latitude and e_i are a set of regression coefficients. There are two sets of regression coefficients: one for the random (uncorrelated) uncertainty to give u_{unc}, and one for the locally correlated component to give u_{cor}.

4.4   Diurnal Variability adjustment

The upper few millimetres of the ocean is referred to as the skin layer. This is cooler than the sub-skin layer immediately below due to latent and sensible heat exchanges with the atmosphere and surface emission of infrared radiation. The sub-skin layer exhibits a diurnal variation in temperature and depth, governed by absorption of solar radiation and wind driven mixing. The deeper mixed layer of the ocean lies below this and is largely unaffected by surface processes. The temperature of this layer is often referred to as the bulk sea surface temperature.

Space-borne infrared instruments measure the temperature of the upper few microns of the sea surface (skin temperature), while microwave instruments are sensitive to the temperature of the upper few millimetres (sub-skin temperature), and in-situ data from buoys or ships are typically a measure of the bulk SST. Furthermore, when combining retrievals from a number of satellite instruments, making observations at different local times (see Figure 6) the diurnal cycle in SST can be aliased into the final product. Depth and time adjustment of the retrieved SST is necessary to meet user requirements for consistency between satellite and in situ records. A standardised depth of 0.2 m is used to match drifting

buoy measurement depth, and standardised time of 10:30 local (am or pm) as this is the approximate mean of the diurnal cycle (so SST at 10:30 or 22:30 can be used to approximate the daily average SST).
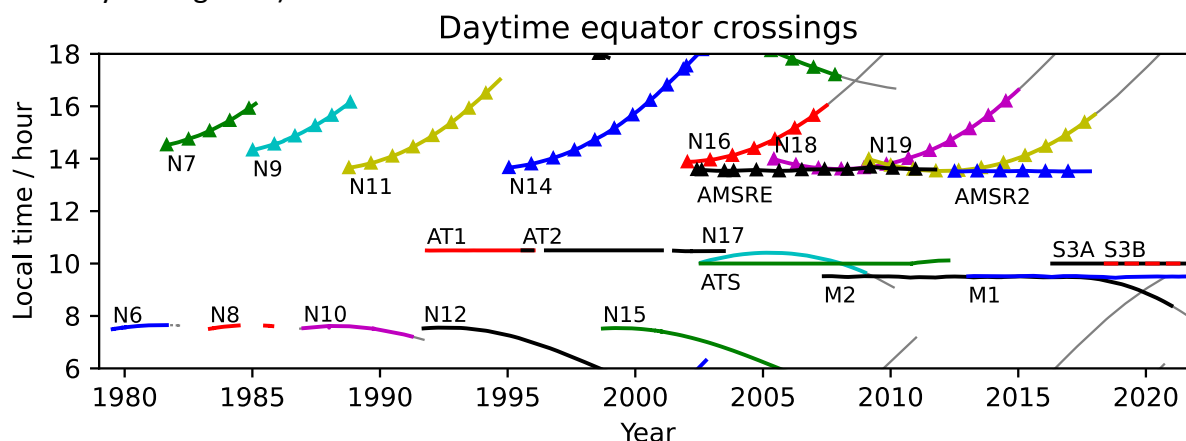

Figure 6: Satellite equator crossing times. Solid lines indicate descending node crossings; lines with triangles indicate ascending node crossings; thin grey lines indicate data were not used in CDR v3.

The DV code was originally developed by Horrocks et al. (2003) and uses the Fairall (1996) model to account for skin effects, ΔT_{skin}, and the Kantha and Clayson (1994) warm-layer model to calculate the depth adjustment, ΔT_{depth}, and diurnal warming throughout the day, ΔT_{time}. All three adjustments are dependent on NWP inputs, $\boldsymbol{x}_a$, listed in Table 6. The time and depth adjusted SST is thus given by:

$$SST_{0.2m}(t_{std}) = SST(t) + \Delta T_{skin}(\boldsymbol{x}_a) + \Delta T_{depth}(\boldsymbol{x}_a) + \Delta T_{time}(\boldsymbol{x}_a, t, t_{std}) \qquad \text{Eq. 11}$$

The additional uncertainty due to the adjustment is:

$$u_{adj} = \sqrt{c(t - t_{std})} \qquad \text{Eq. 12}$$

Where constant c has values listed in Table 9

t	t_{std}	c
04:30 to 10:30	10:30	0.0042
10:30 to 16:30	10:30	0.0054
16:30 to 22:30	22:30	0.0030
22:30 to 04:30	22:30	0.0020

Table 9: DV model uncertainty coefficient c.

Further details can be found in ATBDv2 sections 7 and 8.

4.5   Determining Quality Level

4.5.1 Infrared sensors

Quality level is treated as a concept that is distinct from uncertainty: a highly uncertain SST can have the highest quality level if all the conditions for giving a valid SST and valid SST uncertainty are met: the quality level reflects the degree of confidence in the validity of the uncertainty estimate, not the data uncertainty.

The quality level assigned to a pixel will be the lowest level (row of table) which matches any of the conditions shown in the table below. The assignments are compatible with GHRSST conventions (GHRSST Science Team, 2012): i.e., a particular level is given if none of the conditions higher up any column of the table are met.

Level	Meaning	P(clear)	Sensitivity	χ2	Other
0	No data	<0			No data; land pixel
1	Bad data	<0.5	<0.0	>3	SST < 271.15; ice detected; NWP missing
2	Worst quality	<0.8	<0.10	>2	Limb ($\theta_{sat} > 60$)
3	Low quality	<0.9	<0.20	>1	Twilight (87.5 < θ_{sol} < 92.5)
4	Acceptable quality				ATSR: Aerosol detected: abs(ASDI) > 0.2 AVHRR: solar contamination detected
5	Best quality				

Table 10: Quality level definitions for infrared sensors

For instance, any pixel where P(clear) is unavailable (value is less than zero), required input BTs are unavailable, or which is over land will be assigned quality level of 0. Next, any pixels which have P(clear) < 0.5, calculated SST sensitivity < 0.1 etc. will be assigned quality level of 1 and so on.

Note that:
- Quality level 0 pixels should contain no other data (except land flag in l2p_flags)
- Quality level 2-5 pixels should always contain valid data
- Quality level 1 pixels may contain data in some variables but the data is not suitable for use (bad_data). For instance, the SST retrieval may have been attempted, but rejected as bad_data due to low sensitivity etc. Alternatively, if the retrieved SST is out of range (< 271.15 K) then the value will be missing in the output file.

4.5.2 Microwave sensors

For microwave SSTs the quality level (other than missing and bad data) is dependent on the estimated uncertainty using thresholds are shown in Table 11. The quality checks for removing bad data are fully documented in Alerskans et al. (2020) and briefly summaries below:
- Rain is flagged if the vertical polarized 18 GHz BT is over 240 K.
- RFI is flagged around know sources of RFI (e.g. Ascension Island).

- RFI is also detected by applying secondary SST retrievals which are less accurate, but also less sensitive to RFI as they exclude either the 6.9 or 10 GHz channels. When the SST estimates are significantly different they are flagged as RFI (Alerskans et al., 2020).
- Retrievals within 200 km of sea-ice, or 100 km of land are set to quality level 2.

Level	Meaning	Uncertainty	Other
0	No data		No data; land pixel
1	Bad data		SST < 271.15; SST > 308.15; Rain; RFI
2	Worst quality	$u \geq 1.0$	Proximity to sea ice; proximity to land
3	Low quality	$0.5 < u < 1.0$	
4	Acceptable quality	$0.35 < u < 0.5$	
5	Best quality	$u \leq 0.35$	

Table 11: Quality level definitions for microwave sensors

4.6 Atmospheric Correction Smoothing

The SST retrieval algorithms (section 4.3) used for CDR/ICDR production are optimised to produce a climate-quality SST – i.e. minimum systematic errors, maximum sensitivity to true SST changes and minimum sensitivity to prior SST (i.e. independence from in situ). Meeting these requirements can mean the retrieval has a high sensitivity to radiometric noise resulting in larger uncorrelated uncertainties. For the gridded L3 products uncorrelated errors will be reduced by the averaging process and become negligible over large spatial / temporal scales. However, they can remain significant for full resolution SST products.

Noise in the full resolution SST products can be reduced by noting that the water vapour in the atmosphere varies smoothing on few-km scales. Over these distances the "atmospheric correction" or difference between the SST and top-of-atmosphere BT can be assumed to be constant. For a multi-channel instrument we write a generalised definition of the multi-channel atmospheric correction as:

$$\delta = x - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \qquad \text{Eq. 13}$$

Where the elements of the weighting vector, $\boldsymbol{b}$, satisfy $b_i \geq 0$ for all i and sum to one. This allows us to smooth the atmospheric correction but not the SST itself, exploiting the longer scales of atmospheric variability in order to reduce SST noise.

Atmospheric correction smoothing is applied using a box of 5x5 pixels centred on the target pixel. In order to ensure that process is not contaminated by cloudy or otherwise low quality pixels the smoothing will only utilize pixels with a quality level equal or greater than the centre pixel.

4.6.1 ARC SST (dual-view sensors)

For coefficient-based retrievals used for dual-view sensors we can write the best estimate of the atmospheric correction parameter as:

$$\tilde{\delta} = \langle \hat{x} \rangle - \boldsymbol{b}^{\mathrm{T}} \langle \boldsymbol{y} \rangle = a_0 - (\boldsymbol{a} - \boldsymbol{b})^{\mathrm{T}} \langle \boldsymbol{y} \rangle \qquad \text{Eq. 14}$$

Where the elements of the vector, $\boldsymbol{b}$, are weighted by the radiometric noise in each channel as:

$$b_i = \frac{\sigma_i^{-2}}{\sum_i \sigma_i^{-2}} \qquad \text{Eq. 15}$$

4.6.2 Optimal Estimation (single-view sensors)

For the OE based retrieval we do not need a separate smoothing pass as we can extend the state and observations vectors to directly include the surrounding pixel averages (Merchant et al., 2013). For example, the two-channel retrieval becomes:

$$\boldsymbol{x}^T = [\mathrm{SST} \quad \overline{\mathrm{SST}} \quad \overline{w}]$$

$$\boldsymbol{y}^{\mathrm{T}} = [y_{11} \quad y_{12} \quad \overline{y_{11}} \quad \overline{y_{12}}]$$

$$K = \begin{bmatrix} \dfrac{\partial y_{11}}{\partial \mathrm{SST}} & \dfrac{\partial y_{12}}{\partial \mathrm{SST}} & 0 & 0 \\[2ex] 0 & 0 & \dfrac{\partial \overline{y_{11}}}{\partial \overline{\mathrm{SST}}} & \dfrac{\partial \overline{y_{12}}}{\partial \overline{\mathrm{SST}}} \\[2ex] \dfrac{\partial y_{11}}{\partial \overline{w}} & \dfrac{\partial y_{12}}{\partial \overline{w}} & \dfrac{\partial \overline{y_{11}}}{\partial \overline{w}} & \dfrac{\partial \overline{y_{12}}}{\partial \overline{w}} \end{bmatrix} \qquad \text{Eq. 16}$$

4.7 Remapping (L3U)

The remapping from Level 2 to Level 3 is applied to IR products where the pixel size is similar or smaller than the Level 3 cell size. The lower resolution microwave products are not remapped to Level 3. The remapping follows the GHRRST best practices (GHRSST Science Team, 2012):

- Identify L2 pixels contributing to each L3 cell
- Select highest quality level pixel(s) in the L3 cell
- Calculate the average SST and propagate uncertainties

Figure 7 shows a simple example for a 3x3 grid. The highlighted points (solid fill) are found to lie in the central cell and will be considered when calculating the cell value. In this case the central cell contains a mix of QL1 and QL2 pixels – so only the QL2 pixels will be used to calculate the final cell value.
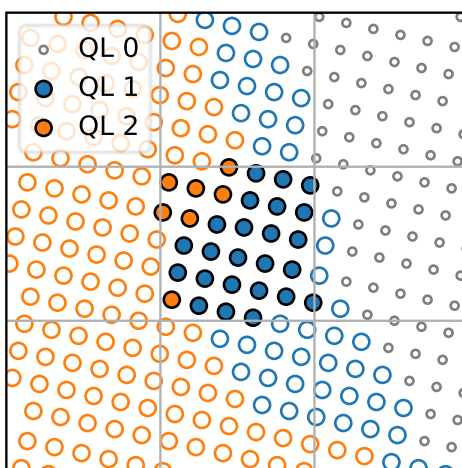
Figure 7: Illustration of remapping from Level 2 to Level 3. Symbols indicate the locations of the L2 data in the Level 3 grid. All highlighted (solid fill) pixels fall within the central cell and will be considered when calculating the value for the central cell.

Propagation of the uncorrelated and correlated uncertainty components is given by:

$$u_{unc} = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n} u_{unc_i}^2}$$

Eq. 17

$$u_{cor} = \frac{1}{n} \sum_{i=1}^{n} u_{cor_i}$$

i.e. the uncorrelated component is reduced by the familiar $1/\sqrt{n}$ while the correlated components are a simple average of the input values.

If the grid cell contains pixels which were not included in the averaging (e.g. due to the presence of cloud etc.), then there is an additional uncertainty due to sampling given by a cubic fit for six bands of SST variability (Bulgin et al., 2016b; SST CCI, 2019):

$$u_{samp} = af^3 + bf^2 + cf + d$$

Eq. 18

Where f is the fraction of pixels used in the cell and coefficients are given in Table 12

SST variability	a	b	c	d
0.0-0.1 K	-0.153	0.322	-0.269	0.10
0.1-0.2 K	-0.154	0.342	-0.352	0.16
0.2-0.3 K	-0.216	0.417	-0.428	0.23
0.3-0.4 K	-0.248	0.449	-0.481	0.28
0.4-0.5 K	-0.231	0.319	-0.369	0.28
0.5-0.6 K	-0.453	0.673	-0.551	0.33

Table 12: Coefficients for sampling uncertainty calculation

4.8   Daily collation (L3C)

The polar orbiting satellites used for CDRv3 (see section 0) typically complete 14-15 orbits each day resulting in the same number of L2P or L3U products. While L3U files are on a global grid, they are very sparse as the sensor will only observe a small fraction of the Earth's surface in each orbit (see Figure 8). For ease of use the SST outputs are collated to produce two files for each 24-hour period – one corresponding to day-time observations and one containing night-time data.
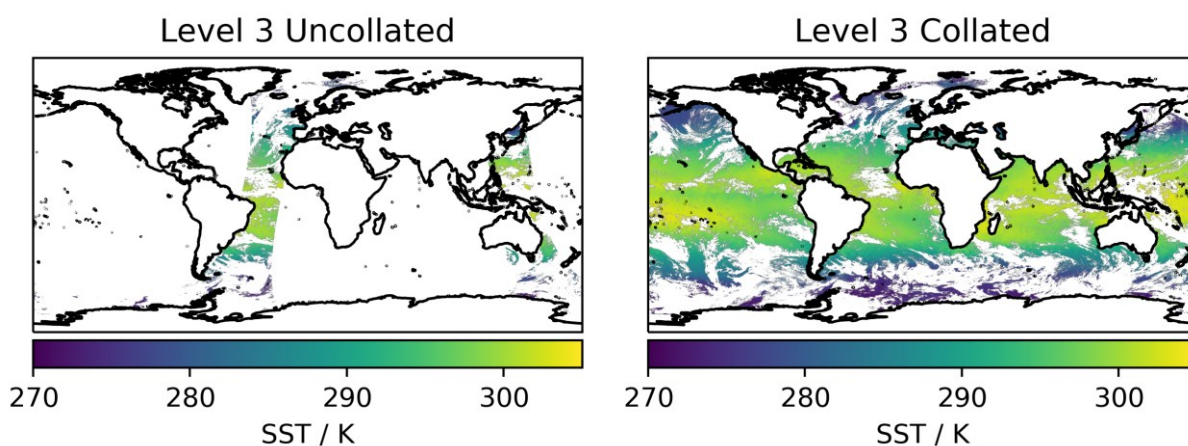

Figure 8: Example plot of AVHRR L3 data. Left – L3U from a single orbit. Right – L3C from 24-hour period.

Following the GHRSST conventions (GHRSST Science Team, 2012), when collating observations from overlapping orbits in the same day the L3C will contain the highest quality observation available in the 24-hour period. The selection of best observation is done as follows:
- Choose input cells with the highest quality_level
- If multiple observations have the same quality_level, then select the one with the lowest estimated uncertainty (variable sst_depth_total_uncertainty is used)

4.9   SST Analysis Scheme (L4)

The SST analyses are produced by minimising a cost function (e.g. Fiedler et al., 2019):

$$J(\delta x) = \frac{1}{2} x^{\mathrm{T}} B^{-1} \delta x + \frac{1}{2} (d - H\delta x)^{\mathrm{T}} R^{-1} (d - H\delta x) \qquad \text{Eq. 19}$$

In this equation, x denotes the ocean state vector (which contains either the SST or sea ice values in the L4 processing) and δx is the departures from the background field. B and R are the background and observation error covariance matrices respectively, d is a vector of differences between observation values and x, and H is the linearised observation operator

which interpolates an analysis grid vector to the observation locations. The cost $J(\delta x)$ is minimised using the NEMOVAR variational data assimilation scheme (Mogensen et al., 2009).

The background field is generated by taking the anomalies between the previous day's analysis and its climatology, reducing them in size by 7.4% and then adding them to the current day's climatology. By doing this, the background will relax to climatology in the absence of observations.

Further details can be found in Good et al. (2020)

4.9.1 Uncertainty propagation and estimation

The optimal interpolation equation for the error variance of an analysis (assuming the optimal set of weights has been found) is (Bouttier and P. Courtier, 2002):

$$A = (I - KH)$$
<div align="right">Eq. 20</div>

Where **I** is the identity matrix and **K** is:

$$K = BH^{\mathrm{T}}(HBH^{\mathrm{T}} + R)^{-1}$$
<div align="right">Eq. 21</div>

This propagates the uncertainty in both the background and the observations through to the analysis error covariance matrix A. Note that although the observation error covariance matrix does not appear explicitly in the equation, it is represented in the expression for K.

It is not possible to evaluate this equation directly using NEMOVAR. Instead an analysis quality method is adopted. It uses an analysis where all observations are set to one and the background field to zero. For each analysis grid point, this provides the sum of the K matrix i.e. the sum of the weights that were applied to the observation minus background differences in the SST analysis. This is then used in an expression for the analysis error variance (Donlon et al., 2012)

$$\varepsilon_i^a = \sqrt{B_i[\alpha + \beta(1 - \varepsilon_i^o)]}$$
<div align="right">Eq. 22</div>

In this equation the error standard deviation ε_i^a in the analysis at point i is given by the square root of the background error variance B_i at that point multiplied by one minus the result of the special analysis ε_i^o (which is the sum of the K matrix for that grid point). The correspondence between this theoretical equation, and the optimal interpolation equation for analysis error variance used for more practical implementation, is straightforward to see. Two additional parameters (α and β) are used in the equation to tune the range of uncertainties that are produced. These are set to 0.5 and 4 respectively.

4.9.2 Sea ice concentration interpolation and infilling

The input sea ice concentration data are provided every two days in the early part of the record. In addition, some fields contained artefacts and weren't used. A linear model was used to interpolate temporal data gaps in the sea ice fields in those earlier records. A full detailed description of the procedure, and results of its application, can be found in section 8 of the ESA SST CCI ATBD CDRv1.

The input ice concentration data do not extend as close to land as the land/sea mask used for the L4 analysis. This results in low ice concentrations near the coast. In CDR3, to prevent this being an issue the near-coast ice concentrations are removed and infilled using a nearest neighbour approach from the regions covered by the input data. More details can be found in Good et al. (2020).

4.10 Calibration-Spike Adjustment

Comparing the global mean Level 4 SST against HadSST4 (Kennedy et al., 2019) shows artefacts of 0.1-0.2 K in monthly average during the 1980s and early 1990s (see Figure 9). This is an improvement on the version 2.1 CDR which had monthly artefacts up to 0.6 K (Merchant and Embury, 2020). These "spikes" are due to intermittent problems in the calibration of the early AVHRR sensors and are especially noticeable in the early record which relies on fewer input sensors (see Figure 6).
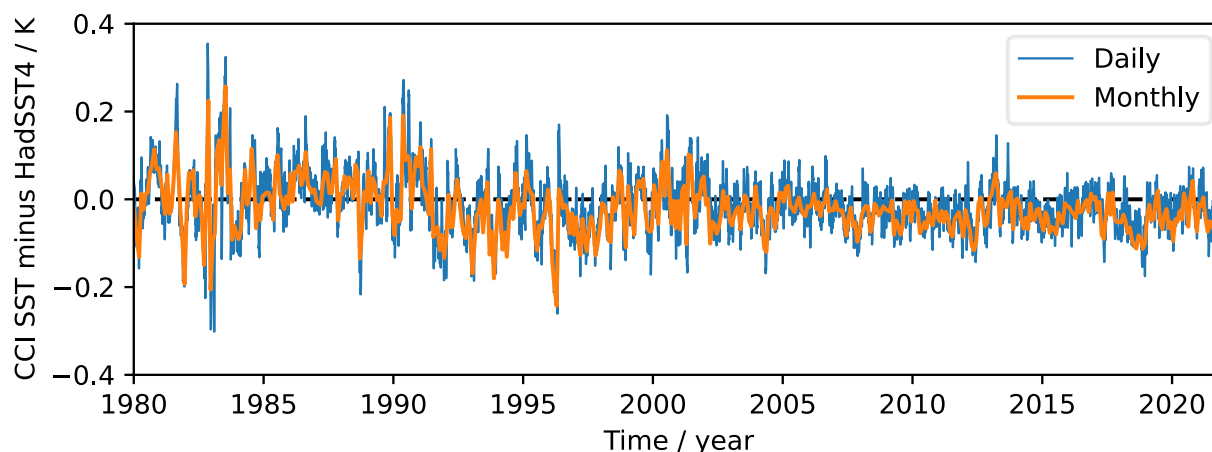

Figure 9: Timeseries of global-mean difference of raw CCI Analysis minus HadSST4. Thin blue line: daily average. Thick orange line: monthly average.

The calibration-spikes are reduced using the method of Merchant and Embury (2020) which calculates an adjustment for the earlier data to match the distribution of differences to the later period. For the data after 1997 the CCI minus HadSST4 differences are distributed near-normally as $\mathcal{N}(-0.02,\ 0.048)$, while the earlier data is more affected by outliers with mean of +0.006 K and standard deviation 0.083 K. An adjustment for the earlier data is estimated by matching the quantiles of the two distributions, resulting in the nearly linear function shown in Figure 10. This is applied to the early-period data, with a linear tapering of the

adjustment from start August 1996 to zero adjustment from end Jan 1997 onwards. Figure 11 shows the timeseries of daily CCI minus HadSST4 differences before and after adjustment.
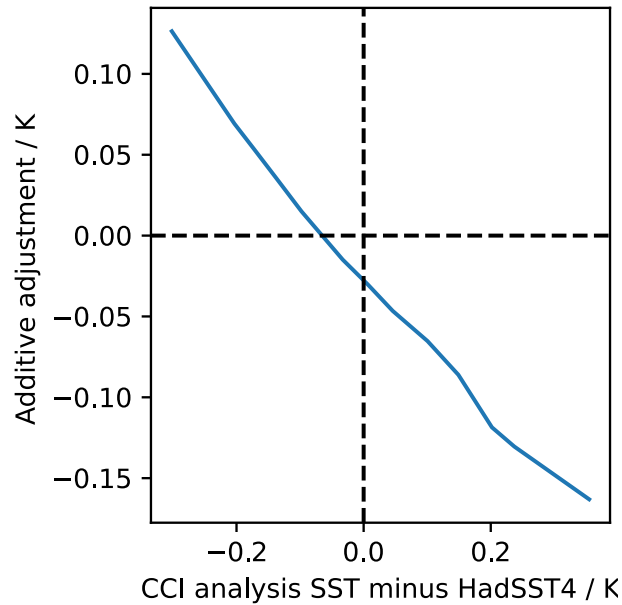

Figure 10: Adjustment function for homogenisation of pre-1997 difference distribution
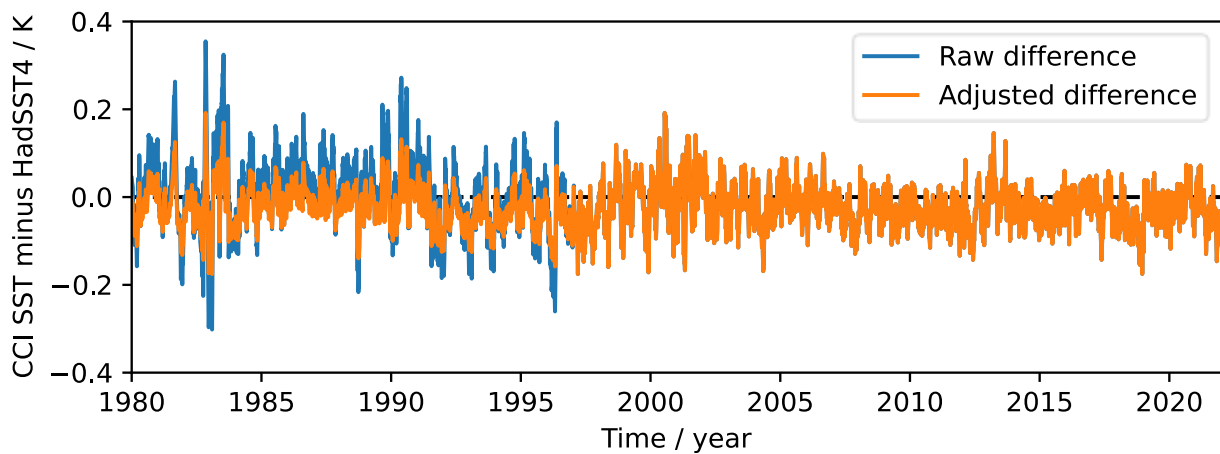

Figure 11: Timeseries of daily global-mean difference of raw CCI Analysis minus HadSST4. Blue line: before adjustment. Orange line: including spike correction adjustment to early data.

5. REFERENCES

Alerskans, E., Høyer, J.L., Gentemann, C.L., Pedersen, L.T., Nielsen-Englyst, P., Donlon, C., (**2020**). Construction of a climate data record of sea surface temperature from passive microwave measurements. *Remote Sensing of Environment* 236, 111485. https://doi.org/10.1016/j.rse.2019.111485

Baran, A.J., Foot, J.S., (**1994**). New application of the operational sounder HIRS in determining a climatology of sulphuric acid aerosol from the Pinatubo eruption. *J. Geophys. Res.* 99, 25673–25679. https://doi.org/10.1029/94JD02044

Block, T., Embacher, S., Merchant, C.J., Donlon, C., (**2018**). High-performance software framework for the calculation of satellite-to-satellite data matchups (MMS version 1.2). *Geosci. Model Dev.* 11, 2419–2427. https://doi.org/10.5194/gmd-11-2419-2018

Bouttier, F., P. Courtier, (**2002**). Data Assimilation Concepts and Methods (Meteorological Training Course Lecture Series). ECMWF.

Bozzo, A., Benedetti, A., Flemming, J., Kipling, Z., Rémy, S., (**2020**). An aerosol climatology for global models based on the tropospheric aerosol scheme in the Integrated Forecasting System of ECMWF. *Geosci. Model Dev*. 13, 1007–1034. https://doi.org/10.5194/gmd-13-1007-2020

Bulgin, C.E., Embury, O., Corlett, G., Merchant, C.J., (**2016a**). Independent uncertainty estimates for coefficient based sea surface temperature retrieval from the Along-Track Scanning Radiometer instruments. *Remote Sensing of Environment* 178, 213–222. https://doi.org/10.1016/j.rse.2016.02.022

Bulgin, C.E., Embury, O., Merchant, C.J., (**2016b**). Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data. *Remote Sensing of Environment* 177, 287–294. https://doi.org/10.1016/j.rse.2016.02.021

Bulgin, C.E., Mittaz, J.P.D., Embury, O., Eastwood, S., Merchant, C.J., (**2018**). Bayesian Cloud Detection for 37 Years of Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) Data. *Remote Sensing* 10, 97. https://doi.org/10.3390/rs10010097

Carboni, E., Grainger, R., Walker, J., Dudhia, A., Siddans, R., (**2012**). A new scheme for sulphur dioxide retrieval from IASI measurements: application to the Eyjafjallajökull eruption of April and May 2010. *Atmos. Chem. Phys*. 12, 11417–11434. https://doi.org/10.5194/acp-12-11417-2012

Cordoba, M., Dance, S.L., Kelly, G.A., Nichols, N.K., Waller, J.A., (**2017**). Diagnosing atmospheric motion vector observation errors for an operational high-resolution data

assimilation system: AMV Observation Errors for Data Assimilation. *Q.J.R. Meteorol. Soc*. 143, 333–341. https://doi.org/10.1002/qj.2925

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., (**2011**). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc*. 137, 553–597. https://doi.org/10.1002/qj.828

Desroziers, G., Berre, L., Chapnik, B., Poli, P., (**2005**). Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc*. 131, 3385–3396. https://doi.org/10.1256/qj.05.108

Donlon, C.J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., Wimmer, W., 2012. The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sensing of Environment*, Advanced Along Track Scanning Radiometer(AATSR) Special Issue 116, 140–158. https://doi.org/10.1016/j.rse.2010.10.017

Embury, O., Merchant, C.J., (**2012**). A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: A new retrieval scheme. *Remote Sensing of Environment*, Advanced Along Track Scanning Radiometer(AATSR) Special Issue 116, 47–61. https://doi.org/10.1016/j.rse.2010.11.020

Embury, O., Merchant, C.J., Corlett, G.K., (**2012a**). A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: Initial validation, accounting for skin and diurnal variability effects. *Remote Sensing of Environment*, Advanced Along Track Scanning Radiometer(AATSR) Special Issue 116, 62–78. https://doi.org/10.1016/j.rse.2011.02.028

Embury, O., Merchant, C.J., Filipiak, M.J., (**2012b**). A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: Basis in radiative transfer. *Remote Sensing of Environment*, Advanced Along Track Scanning Radiometer(AATSR) Special Issue 116, 32–46. https://doi.org/10.1016/j.rse.2010.10.016

Fairall, C.W., Bradley, E.F., Godfrey, J.S., Wick, G.A., Edson, J.B., Young, G.S., (**1996**). Cool-skin and warm-layer effects on sea surface temperature. *Journal of Geophysical Research: Oceans* 101, 1295–1308. https://doi.org/10.1029/95JC03190

Fiedler, E.K., Mao, C., Good, S.A., Waters, J., Martin, M.J., (**2019**). Improvements to feature resolution in the OSTIA sea surface temperature analysis using the NEMOVAR assimilation scheme. *Quarterly Journal of the Royal Meteorological Society* 145, 3609–3625. https://doi.org/10.1002/qj.3644

GHRSST Science Team, 2012. The Recommended GHRSST Data Specification (GDS) 2.0. Zenodo. https://doi.org/10.5281/zenodo.4700466

Good, S., Fiedler, E., Mao, C., Martin, M.J., Maycock, A., Reid, R., Roberts-Jones, J., Searle, T., Waters, J., While, J., Worsfold, M., (**2020**). The Current Configuration of the OSTIA System for Operational Production of Foundation Sea Surface Temperature and Ice Concentration Analyses. *Remote Sensing* 12, 720. https://doi.org/10.3390/rs12040720

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N., (**2020**). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146, 1999–2049. https://doi.org/10.1002/qj.3803

Holl, G., Mittaz, J.P.D., Merchant, C.J., (**2019**). Error Correlations in High-Resolution Infrared Radiation Sounder (HIRS) Radiances. *Remote Sensing* 11, 1337. https://doi.org/10.3390/rs11111337

Horrocks, L.A., Candy, B., Nightingale, T.J., Saunders, R.W., O'Carroll, A., Harris, A.R., (**2003**). Parameterizations of the ocean skin effect and implications for satellite-based measurement of sea-surface temperature. *Journal of Geophysical Research: Oceans* 108. https://doi.org/10.1029/2002JC001503

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J.J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., Suttie, M., (**2019**). The CAMS reanalysis of atmospheric composition. *Atmos. Chem. Phys*. 19, 3515–3556. https://doi.org/10.5194/acp-19-3515-2019

Kalman, R.E., 1960. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering 82, 35–45. https://doi.org/10.1115/1.3662552

Kantha, L.H., Clayson, C.A., (**1994**). An improved mixed layer model for geophysical applications. *Journal of Geophysical Research: Oceans* 99, 25235–25266. https://doi.org/10.1029/94JC02257

Kennedy, J.J., Rayner, N.A., Atkinson, C.P., Killick, R.E., (**2019**). An Ensemble Data Set of Sea Surface Temperature Change From 1850: The Met Office Hadley Centre HadSST.4.0.0.0 Data Set. *J. Geophys. Res. Atmos.* 124, 7719–7763. https://doi.org/10.1029/2018JD029867

Kilic, L., Prigent, C., Aires, F., Boutin, J., Heygster, G., Tonboe, R.T., Roquet, H., Jimenez, C., Donlon, C., (**2018**). Expected Performances of the Copernicus Imaging Microwave Radiometer (CIMR) for an All-Weather and High Spatial Resolution Estimation of Ocean and Sea Ice

Parameters. *J. Geophys. Res. Oceans* 123, 7564–7580. https://doi.org/10.1029/2018JC014408

Koner, P., Harris, A., (**2016**). Sea Surface Temperature Retrieval from MODIS Radiances Using Truncated Total Least Squares with Multiple Channels and Parameters. *Remote Sensing* 8, 725. https://doi.org/10.3390/rs8090725

Koner, P.K., Harris, A., Maturi, E., (**2015**). A Physical Deterministic Inverse Method for Operational Satellite Remote Sensing: An Application for Sea SurfaceTemperature Retrievals. *IEEE Trans. Geosci. Remote Sensing* 53, 5872–5888. https://doi.org/10.1109/TGRS.2015.2424219

McGarragh, G.R., Poulsen, C.A., Thomas, G.E., Povey, A.C., Sus, O., Stapelberg, S., Schlundt, C., Proud, S., Christensen, M.W., Stengel, M., Hollmann, R., Grainger, R.G., (**2018**). The Community Cloud retrieval for CLimate (CC4CL) - Part 2: The optimal estimation approach. *Atmospheric Measurement Techniques* 11, 3397–3431. https://doi.org/10.5194/amt-11-3397-2018

Merchant, C.J., Block, T., Corlett, G.K., Embury, O., Mittaz, J.P.D., Mollard, J.D.P., (**2020a**). Harmonization of Space-Borne Infra-Red Sensors Measuring Sea Surface Temperature. *Remote Sensing* 12, 1048. https://doi.org/10.3390/rs12061048

Merchant, C.J., Embury, O., (**2020**). Adjusting for Desert-Dust-Related Biases in a Climate Data Record of Sea Surface Temperature. *Remote Sensing* 12, 2554. https://doi.org/10.3390/rs12162554

Merchant, C.J., Embury, O., Bulgin, C.E., Block, T., Corlett, G.K., Fiedler, E., Good, S.A., Mittaz, J., Rayner, N.A., Berry, D., Eastwood, S., Taylor, M., Tsushima, Y., Waterfall, A., Wilson, R., Donlon, C., (**2019**). Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci Data* 6, 223. https://doi.org/10.1038/s41597-019-0236-x

Merchant, C.J., Embury, O., Rayner, N.A., Berry, D.I., Corlett, G.K., Lean, K., Veal, K.L., Kent, E.C., Llewellyn-Jones, D.T., Remedios, J.J., Saunders, R., (**2012**). A 20 year independent record of sea surface temperature for climate from Along-Track Scanning Radiometers. *Journal of Geophysical Research: Oceans* 117. https://doi.org/10.1029/2012JC008400

Merchant, C.J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C.E., Corlett, G.K., Good, S., McLaren, A., Rayner, N., Morak-Bozzo, S., Donlon, C., (**2014**). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). *Geoscience Data Journal* 1, 179–191. https://doi.org/10.1002/gdj3.20

Merchant, C.J., Harris, A.R., Maturi, E., Maccallum, S., (**2005**). Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval. *Q. J. R. Meteorol. Soc*. 131, 2735–2755. https://doi.org/10.1256/qj.05.15

Merchant, C.J., Le Borgne, P., Marsouin, A., Roquet, H., (**2008**). Optimal estimation of sea surface temperature from split-window observations. *Remote Sensing of Environment* 112, 2469–2484. https://doi.org/10.1016/j.rse.2007.11.011

Merchant, C.J., Le Borgne, P., Roquet, H., Legendre, G., (**2013**). Extended optimal estimation techniques for sea surface temperature from the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI). *Remote Sensing of Environment* 131, 287–297. https://doi.org/10.1016/j.rse.2012.12.019

Merchant, C.J., Le Borgne, P., Roquet, H., Marsouin, A., (**2009**). Sea surface temperature from a geostationary satellite by optimal estimation. *Remote Sensing of Environment* 113, 445–457. https://doi.org/10.1016/j.rse.2008.10.012

Merchant, C.J., Saux-Picart, S., Waller, J., (**2020b**). Bias correction and covariance parameters for optimal estimation by exploiting matched in-situ references. *Remote Sensing of Environment* 237, 111590. https://doi.org/10.1016/j.rse.2019.111590

Mogensen, K., Alonso-Balmaseda, M., Weaver, A., Martin, M., Vidard, A., (**2009**). NEMOVAR: A variational data assimilation system for the NEMO ocean model. ECMWF Newsletter 120, 17–21. https://doi.org/10.21957/3YJ3MH16IQ

Poulsen, C.A., Siddans, R., Thomas, G.E., Sayer, A.M., Grainger, R.G., Campmany, E., Dean, S.M., Arnold, C., Watts, P.D., (**2012**). Cloud retrievals from satellite data using optimal estimation: evaluation and application to ATSR. *Atmos. Meas. Tech*. 5, 1889–1910. https://doi.org/10.5194/amt-5-1889-2012

Roberts-Jones, J., Bovis, K., Martin, M.J., McLaren, A., (**2016**). Estimating background error covariance parameters and assessing their impact in the OSTIA system. *Remote Sensing of Environment* 176, 117–138. https://doi.org/10.1016/j.rse.2015.12.006

Rodgers, C.D., (**2000**). Inverse Methods for Atmospheric Sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics. WORLD SCIENTIFIC. https://doi.org/10.1142/3171

Rodgers, C.D., (**1976**). Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Reviews of Geophysics* 14, 609–624. https://doi.org/10.1029/RG014i004p00609

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., Lupu, C., (**2018**). An update on the RTTOV fast radiative transfer model (currently at version 12). *Geosci. Model Dev*. 11, 2717–2737. https://doi.org/10.5194/gmd-11-2717-2018

SST CCI Phase II, (**2019**). Algorithm Theoretical Basis Document (v2 Reprocessing) SST_CCI-ATBD-UOR-203.   https://climate.esa.int/en/projects/sea-surface-temperature/SST-key-documents/

Thomas, G.E., Poulsen, C.A., Sayer, A.M., Marsh, S.H., Dean, S.M., Carboni, E., Siddans, R., Grainger, R.G., Lawrence, B.N., (**2009**). The GRAPE aerosol retrieval algorithm. *Atmos. Meas. Tech.* 2, 679–701. https://doi.org/10.5194/amt-2-679-2009

Waller, J., Ballard, S., Dance, S., Kelly, G., Nichols, N., Simonin, D., (**2016**). Diagnosing Horizontal and Inter-Channel Observation Error Correlations for SEVIRI Observations Using Observation-Minus-Background and Observation-Minus-Analysis Statistics. *Remote Sensing* 8, 581. https://doi.org/10.3390/rs8070581

APPENDIX A. BIAS-AWARE OPTIMAL ESTIMATION

A.1   Sea surface temperature retrieval

Optimal estimation (OE) is a widely used inverse method in geophysics and remote sensing (Rodgers, 2000). OE has been applied to retrieval of sea surface temperature from space-based imagery (Kilic et al., 2018; Merchant et al., 2009, 2008) as well as for many problems in atmospheric sounding (Carboni et al., 2012; McGarragh et al., 2018; Poulsen et al., 2012; Thomas et al., 2009). The most common formulation for SST retrieval jointly estimates the radiometric skin temperature, x, and the total column water vapour (TCWV), w. This is a reduced state vector, $\boldsymbol{z} = [x, w]^{\mathrm{T}}$, since the clear-sky brightness temperatures observed by the satellite, $\boldsymbol{y}$, depend in general on a much longer list of state variables, $\boldsymbol{x}$, including the vertical distribution of atmospheric temperature and humidity, the sea state and atmospheric aerosols.

OE requires a physical model of the observations, which in this case means a radiative transfer model (RTM) that operates on the full state vector, $\boldsymbol{F}(\boldsymbol{x}_a)$. The subscript a here indicates *a priori*, meaning that the RTM is run on the prior estimate of the surface temperature and atmospheric state before retrieval. The prior state generally is obtained from a numerical weather prediction forecast or analysis. The physical model is also differentiated with respect to the retrieved variables:

$$K = \frac{\partial F}{\partial z} \qquad\qquad \text{Eq. 23}$$

The retrieved state, $\hat{\boldsymbol{z}}$, corresponds to the prior state, $\boldsymbol{z}_a$, modified by the difference between the satellite observations and the expected brightness temperatures as simulated by the RTM on the prior state. This difference is transformed from the observation space to the state space by a gain matrix that accounts for the sensitivity of the brightness temperatures to the state variables (Eq. 23) and the relative uncertainty of the observations and prior state:

$$\hat{\boldsymbol{z}} = \boldsymbol{z}_a + (\boldsymbol{K}^{\mathrm{T}}\boldsymbol{S}_\epsilon^{-1}\boldsymbol{K} + \boldsymbol{S}_a^{-1})^{-1}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{S}_\epsilon^{-1}(\boldsymbol{y} - \boldsymbol{F}) \qquad\qquad \text{Eq. 24}$$

Here, $\boldsymbol{S}_\epsilon$ characterises the error distribution in the observations relative to the simulations as an error covariance matrix. The uncertainty in each brightness temperature is given by the square root of the diagonal values of this matrix, while the off-diagonals indicate the strength of covariance (and, thus, correlation) of those errors. Similarly, the error covariance matrix of the prior information is $\boldsymbol{S}_a$, whose first diagonal term is the square of the uncertainty in the prior estimate of SST. If the error distributions of $\boldsymbol{z}_a$ and $\boldsymbol{y} - \boldsymbol{F}$ are zero-mean, the retrieved value will have zero mean expectation of error. If the error distributions are gaussian and correctly characterised by $\boldsymbol{S}_\epsilon$ and $\boldsymbol{S}_a$, the retrieved values will be an optimal solution and their uncertainty will be accurately characterised by the retrieval error covariance matrix $\boldsymbol{S} = (\boldsymbol{K}^{\mathrm{T}}\boldsymbol{S}_\epsilon^{-1}\boldsymbol{K} + \boldsymbol{S}_a^{-1})^{-1}$, which may be calculated as part of the evaluation of Eq. 24.

To obtain optimal estimates, therefore, the prior must be unbiased, the satellite calibration and radiative transfer simulations need to be negligibly biased with respect to each other and good estimates of two error covariance matrices need to be available. To meet these conditions for a system of n_y observations and n_z retrieved variables, we need good estimates of the following: $n_y + n_z$ bias correction parameters (more if there are systematic tendencies which are fitted rather than a single offset); $n_y + n_z$ uncertainty estimates (which again may vary according to the circumstances of the retrieval); and $\frac{1}{2}n_y(n_y + 1) + \frac{1}{2}n_z(n_z + 1)$ error correlation coefficients. For even the minimal system of retrieving SST and TCWV using three infrared brightness temperatures, this corresponds to 14 retrieval parameters if no parameter dependency is fitted. Some relevant information to constrain these parameters is generally available, such as error correlation being zero between independent sources of information, or estimates of instrument noise from sensor specifications or the onboard calibration system. Nonetheless, in past implementations, e.g. Merchant et al. (2009), simplifications such as diagonal error covariance matrices (Rodgers, 2000) have been used along with judgements about many parameters. This element of judgement has been strongly criticized (Koner and Harris, 2016; Koner et al., 2015) and alternative inverse methods have been proposed with fewer parameters. The alternative response, pursued here, is to exploit a means of systematically evaluating the retrieval parameters, to which we refer as "bias-aware" optimal estimation.

A.2   Bias parameter estimation

Bias-aware optimal estimation (BAOE) combines the following elements: the insight in Rodgers (2000) that uncertain retrieval parameters can themselves be retrieved across a large number of retrieval instances; the idea that anchoring the system to some reference data that are taken to be unbiased can help disambiguate biases arising from different sources; and the use of expressions derived as diagnostics in data assimilation (Desroziers et al., 2005) as means of objectively estimating error covariances (Cordoba et al., 2017; Waller et al., 2016) after bias-correction. Merchant et al. (2020b) (hereafter M20) presented a BAOE approach with reference to in situ references, and in Merchant et al. (2020a) an approach for cross-satellite harmonisation of sea surface temperature is demonstrated. In this work, in situ references are again used, but the derivation of the bias parameters is reformulated (relative to M20) to reduce complexity and the amount of radiative transfer modelling involved.

Observational biases are generally present and mean that errors $\boldsymbol{y} - \boldsymbol{F}$ do not have zero mean over a large ensemble of retrievals. The retrieved value is sensitive to the bias in this difference, irrespective of whether the source of bias is in the satellite calibration or in the forward model. We formulate the bias parameters, $\boldsymbol{\beta}$, as corrections to be added to the forward model, but this choice does not imply that the forward model is the only or main source of systematic errors.

The prior estimate of the state may also have non-zero mean error across many instances of retrieval. Previous results (M20) suggest that clear-sky areas of infrared imagery have lower TCWV than prior NWP humidity profiles; since these are all-sky profiles, this is physically plausible. Corrections for prior bias, $\boldsymbol{\gamma}$, are defined such that $\boldsymbol{z}_a + \boldsymbol{\gamma}$ will be unbiased.

The $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ parameters are included in the extended state (retrieved) vector, $\tilde{\boldsymbol{z}}$. To anchor the estimate of observational biases, reference data (see section A.4) are included as additional observations in the observation vector, $\widetilde{\boldsymbol{y}}$. (This is the point of difference between the approach here and M20. In M20, the forward model was run using the reference data for the SST, rather than using the usual prior SST source for the retrieval as here. This is a simplification in procedure since simulations for the latter source are naturally available from the satellite data processing chain.)

The error covariance matrices are initially specified using expert knowledge and any information available on error sources such as instrument noise. The bias-correction parameters will be best estimated if the error covariances matrices are well estimated. Conversely, the error covariance matrices need to characterise the uncertainties after bias corrections have been applied. The iterative update of the error covariance matrices is therefore necessary and is discussed in section A.3.

Using the initial estimates of error covariance matrices, the extended optimal estimator is solved across a large sample of cases for which in situ reference data are available. These cases are hereafter referred to as 'matches' since matching of the satellite data and in situ data is required, as described in section A.4. The matches are randomised, to ensure independence of consecutive iterations with respect to geophysical parameters and errors in individual reference values. Updated bias parameters (and their uncertainties) are passed between iterations. The equation for the ith iteration thus uses the bias parameters retrieved in the $i-1^{\text{th}}$ retrieval. For the case where the reference data are measured values of SST, x_r having estimated uncertainty u_r, and the bias in the prior TCWV, γ, is to be estimated, the extended OE equation is:

Algorithm Theoretical Basis Document D2.1 v3

$$\tilde{z}_i = \tilde{z}_a + \left(\widetilde{K}^{\mathrm{T}}\tilde{S}_\epsilon^{-1}\widetilde{K} + \tilde{S}^{-1}\right)^{-1}\widetilde{K}^{\mathrm{T}}\tilde{S}_\epsilon^{-1}\left(\tilde{y} - \left(\widetilde{F}(\tilde{z}_a) + \gamma_{i-1}\,\partial\widetilde{F}/\partial w + \widetilde{\beta}_{i-1}\right)\right)$$

$$\tilde{z}_a = \begin{bmatrix} x_a \\ w_a + \gamma_{i-1} \\ \gamma_{i-1} \\ \boldsymbol{\beta}_{i-1} \end{bmatrix}$$

$$\tilde{y} = \begin{bmatrix} \boldsymbol{y} \\ x_r \end{bmatrix}$$

$$\widetilde{F}(\tilde{z}_a) = \begin{bmatrix} \boldsymbol{F}(\boldsymbol{z}_a) \\ x_a \end{bmatrix}$$

$$\widetilde{\beta}_{i-1} = \begin{bmatrix} \boldsymbol{\beta}_{i-1} \\ 0 \end{bmatrix}$$

$$\widetilde{K} = \begin{bmatrix} \partial\boldsymbol{F}/\partial x & \partial\boldsymbol{F}/\partial w & \partial\boldsymbol{F}/\partial w & \boldsymbol{I} \\ 1 & 0 & 0 & \boldsymbol{0} \end{bmatrix}$$

Eq. 25

$$\tilde{S} = \begin{bmatrix} \boldsymbol{S}_a + \begin{bmatrix} 0 & 0 \\ 0 & u_{\gamma_{i-1}}^2 \end{bmatrix} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & u_{\gamma_{i-1}}^2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{S}_{\beta_{i-1}} \end{bmatrix}$$

$$\tilde{S}_\epsilon = \begin{bmatrix} \boldsymbol{S}_\epsilon & 0 \\ 0 & u_r^2 \end{bmatrix}$$

$$\tilde{z}_i = \begin{bmatrix} \hat{x}_i \\ \widehat{w}_i \\ \gamma_i \\ \boldsymbol{\beta}_i \end{bmatrix}$$

The form of this extended retrieval is identical to Eq. 24, and Eq. 25. is written explicitly to show how extension of the state and observation vectors leads to extended forward model, partial derivative and error covariance matrices. After the retrieval of the extended state vector for the ith match, the updated parameters are passed to the next randomly selected match, and the parameters are thus progressively updated. The principle is analogous to model parameter estimation using Kalman filtering (Kalman, 1960) but without any concept of continuity in time. Graphs of the evolution of the iterative $\boldsymbol{\beta}$ and γ parameter estimates are inspected to ensure convergence, and $O(10^5)$ iterations are required for a given pass.

Eq. 25 is written as if $\boldsymbol{\beta}$ and γ are constants for a given bias-corrected quantity. This would not be adequate representation of the systematic dependencies observed. For example, the correction of a given brightness temperature may need to depend on factors such as satellite zenith angle, slant path integrated water vapour and/or instrumental parameters, for example. Context-dependent corrections are estimated by deriving different parameter

values from only those matches falling within sub-ranges of auxiliary quantities and using these values to define a piecewise linear correction function.

For example, the correction of the prior total column water vapour, w_a, has been derived as a function of its own uncorrected value. The range of w_a in matched data typically ranges from close to zero (dry, high-latitude locations) to ~60 kg m⁻² (regions of convergence in the tropics). This range is split into n_γ sub-ranges (hereafter, "bins", index $j \in \{1 \ldots n_\gamma\}$). In a given iteration, i, of Eq. 25, the prior water vapour is interrogated to identify the value, J, of j (i.e., in which bin w_a lies). γ actually comprises n_γ parameter values, γ_j, one for each bin, and the iteration i updates $\gamma_{j=J}$, while the parameter values $\gamma_{j \neq J}$ are not modified. The mean prior water vapour in each bin, $\overline{w}_j$ is also calculated. The full γ correction is defined as the piecewise linear interpolation with respect to w_a between the (final iterated) values $\gamma(\overline{w}_j) = \gamma_j$. Extrapolation is not used, and for $w_a < \overline{w}_1$ the correction is fixed at $\gamma(w_a) = \gamma_1$, and similarly for $w_a > \overline{w}_{n_\gamma}$ the correction is fixed at $\gamma(w_a) = \gamma_{n_\gamma}$.

Piecewise linear correction is also used for the adjustment of brightness temperature for each infrared channel in use. To obtain a correction that accounts for more than one dependency, cumulative piecewise linear corrections are calculated. If the n_q auxiliary quantities for the brightness temperature corrections (e.g., satellite zenith angle, etc) are $q_k : k \in \{1 \ldots n_q\}$ and $L_{c,k}(q_k)$ is the piecewise linear function defined by interpolated between the derived parameter values $\beta_{c,k}(\overline{q}_{k,j})$, then the total brightness temperature correction is:

$$\beta_c\left(q_1, \ldots, q_{n_q}\right) = \sum_{k=1}^{n_q} L_{c,k} \qquad \text{Eq. 26}$$

for the given channel, c. The choice of auxiliary quantities could in principle differ between channels according to insights into any channel-specific problems, but in the present implementation the same auxiliary quantities are used for all infrared channels of a given sensor.

To obtain $\boldsymbol{\beta}$ values that add cumulatively (Eq. 26), they are derived sequentially. Correction parameters $\beta_{c,1}(\overline{q}_{1,j})$ with respect to auxiliary quantity q_1 are obtained by iterative application of Eq. 25 with no other corrections applied. Then, applying the corresponding correction $L_{c,1}(q_1)$ to $\boldsymbol{F}$ for each iteration i, the iterative procedure is applied again for the auxiliary quantity q_2, and so on.

A.3 Covariance matrix estimation

The estimation of the bias correction parameters described in the previous section is initially done using observation and prior error covariance matrices, $\boldsymbol{S}_\epsilon$ and $\boldsymbol{S}_a$, that are obtained by experience, expert judgement and information such as sensor noise estimates (from the literature or from the onboard calibration processes). It has been noted even in the earliest implementations of OE for SST retrieval (Merchant et al., 2008) that it would be preferable to

have more objective means of determination, and BAOE provides this by adapting expressions originally brought together in Desroziers et al. (2005).

The observation error covariance (which accounts for uncertainty in satellite calibration, noise and radiative transfer simulation) is first re-estimated after the bias corrections for the brightness temperatures are applied. The first step is to undertake joint retrieval, $\hat{\boldsymbol{z}}$, of the state using the extended observation vector (corrected brightness temperatures and the reference data) across the full set of matches. The statistics of the result are then used to update $\boldsymbol{S}_\epsilon$ using:

$$\widehat{\boldsymbol{S}}_\epsilon = \begin{bmatrix} \widehat{\boldsymbol{S}}_\epsilon & \boldsymbol{D} \\ \boldsymbol{D}^{\mathrm{T}} & \hat{u}_r^2 \end{bmatrix} = \frac{1}{2} \langle \boldsymbol{d}_r^o \boldsymbol{d}_a^{o\,\mathrm{T}} + \boldsymbol{d}_a^o \boldsymbol{d}_r^{o\,\mathrm{T}} \rangle$$

$$\boldsymbol{d}_r^o = \widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{F}}(\hat{\boldsymbol{z}}) - \langle \widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{F}}(\hat{\boldsymbol{z}}) \rangle \qquad \text{Eq. 27}$$

$$\boldsymbol{d}_a^o = \widetilde{\boldsymbol{y}} - \left(\widetilde{\boldsymbol{F}}(\hat{\boldsymbol{z}}) - \widetilde{\boldsymbol{K}}_{\hat{z}}(\hat{\boldsymbol{z}} - \boldsymbol{z}_a) \right) - \langle \widetilde{\boldsymbol{y}} - \left(\widetilde{\boldsymbol{F}}(\hat{\boldsymbol{z}}) - \widetilde{\boldsymbol{K}}_{\hat{z}}(\hat{\boldsymbol{z}} - \boldsymbol{z}_a) \right) \rangle$$

Eq. 27 is formulated explicitly to show use of de-meaned residuals and how to make the estimated matrix symmetric. $\widehat{\boldsymbol{S}}_\epsilon$ provides improved observation uncertainty information on its diagonal, and information about error correlations via the off-diagonal terms. Satellite noise is not expected to be correlated between channels, although there are exceptions, e.g. Holl et al. (2019). However, calibration and simulation errors may involve cross-channel correlations, which previously have been poorly known. This also provides an updated estimate, $\hat{u}_r$ for the reference data uncertainty, which is useful since it is not necessarily well quantified up front, particularly further back in time.

Instrumental uncertainty expressed as noise in brightness temperatures are typically scene dependent, because of non-linearity of the radiance-temperature relationship (which is the channel-integrated Planck function). The component of "observation error" that comes from the radiative transfer simulation is expected to be variable: the approximations of fast radiative transfer simulations are more uncertain when the optical path length from surface to sensor increases. It is reasonable, therefore, to estimate $\boldsymbol{S}_\epsilon$ as a piecewise linear function of a quantity that correlates somewhat with both brightness temperature and optical path length of a quantity that correlates usefully with brightness temperature and infrared optical path length. We use the slant-path integrated water vapour (hereafter "WV path") for this purpose (i.e., $w \sec(\theta)$, where θ is the satellite zenith angle). Eq. 27 is applied to subsets of the matches falling within bins of WV path to obtain the piecewise linear dependence.

Having obtained an improved estimate of $\boldsymbol{S}_\epsilon$ the next step is to update the prior error covariance, $\boldsymbol{S}_a$. For this, a new set of SST retrievals is made, now incorporating the new estimates of the observation error covariance as well as the brightness temperature corrections. The differences of the retrieved and prior state are used to estimate $\boldsymbol{S}_a$.

The diagonal of S_a contains the squares of the uncertainty in prior SST and prior TCWV. The uncertainty in each depends on the source of prior information, is likely to increase for earlier times and may be geographically variable. The possibility of time-dependence is incorporated in our implementation only indirectly by estimating S_a separately for different sensors. To capture to some degree the possibility of large-scale variations in these uncertainties, S_ϵ is estimated for bins of prior SST itself, which varies with latitude and correlates with TCWV. This is done by applying Eq. 28 to matches falling within sub-ranges of the SST distribution.

$$\widehat{S}_a = \frac{1}{2}\langle(K^{\mathrm{T}}K)^{-1}K^{\mathrm{T}}\left(d_a^r d_a^{o\,\mathrm{T}} + d_a^o d_a^{r\,\mathrm{T}}\right)K(K^{\mathrm{T}}K)^{-1}\rangle$$

$$d_a^r = K(\hat{z} - z_a) - \langle K(\hat{z} - z_a)\rangle$$

Eq. 28

We also have strong expectations for S_a. The value of the prior SST uncertainty should, when combined in quadrature with the reference SST uncertainty, correspond closely to the standard deviation between the prior and the reference. In our implementation, we constrain this to be the case. The off-diagonals should be zero, since we do not expect error correlation between SST and TCWV, and in our implementation the implied correlation coefficients are inspected to verify they are small (they are generally <0.25 in magnitude) and then set to zero. In contrast, the TCWV uncertainty embodied in the estimate of S_a is not otherwise well quantified and is used unmodified thereafter.

Some practical measures to stabilize the calculation of Eq. 27 and Eq. 28 are implemented. In evaluating $\langle d_r^o d_a^{o\,\mathrm{T}} + d_a^o d_r^{o\,\mathrm{T}}\rangle$, a trimmed mean is used to avoid undue influence of a small fraction of outliers. $K^{\mathrm{T}}K$ is inverted in Eq. 28, and is occasionally (typically for <0.5% of matches) ill-conditioned. In our implementation, matches for which the condition number of $K^{\mathrm{T}}K$ exceeds 10^6 are excluded from the evaluation.

A.4   Reference data

To estimate the observational biases and error covariance matrices we need to select a reference SST dataset. For the MetOp AVHRR sensors we can use the SST observations from the dual-view ATSR and SLSTR sensors as described in Merchant et al. (2020a). Compared to the earlier work the multi-sensor matchup dataset (Block et al., 2018) has been regenerated for both MetOp-A and MetOp-B AVHRRs, adding SLSTR-B as a further reference sensor, and switching to ERA-5 (Hersbach et al., 2020) as the NWP source. The bridging datasets used for MetOp AVHRRs are listed in Table 13.

AVHRR	Reference	Start	End	Matches / million
MetOp-A	AATSR	2007-06-01	2012-03-31	34.39
	SLSTR-A	2019-02-01	2021-06-30	1.65
	SLSTR-B	2020-01-01	2021-06-30	1.46
MetOp-B	SLSTR-A	2019-02-01	2021-06-30	2.30
	SLSTR-B	2019-05-01	2021-06-30	3.08

Table 13: Inter-sensor bridging datasets used to harmonise MetOp AVHRR BAOE retrievals.

Although the later NOAA AVHRRs do overlap with the ATSR sensors, the earlier NOAA AVHRRs can only be harmonised against in situ data. Furthermore, ERA5 NWP data were not available for the 1980s through early 2000s when the NOAA AVHRR matchup datasets were being generated. Therefore, in order to ensure consistency in the processing of NOAA AVHRRs all of the NOAA AVHRRs were harmonised against in situ reference using ERA-interim (Dee et al., 2011) NWP.

A.5   Other aspects of implementation

As described in (Merchant et al., 2020a) the sequence of bias correction and updating covariance matrices is cycle through iteratively (two complete cycles followed by the final bias correction step). This is needed so that the bias corrections and covariance parameters have the opportunity to respond to each other during optimization.

The process is coded semi-automatically. Expert intervention arises in terms of choosing the auxiliary quantities used for each sensor. This is based upon inspection of the patterns of residual biases seen in untuned OE results obtained before the BAOE process.