remote sensing

MDPI

Article

Presenting a Long-Term, Reprocessed Dataset of Global Sea Surface Temperature Produced Using the OSTIA System

Mark Worsfold [1,*], Simon Good [1], Chris Atkinson [1] and Owen Embury [2]

[1] Met Office, Exeter EX1 3PB, UK; simon.good@metoffice.gov.uk (S.G.); chris.atkinson@metoffice.gov.uk (C.A.)
[2] Department of Meteorology, University of Reading, Reading RG6 6UR, UK; o.embury@reading.ac.uk
* Correspondence: mark.worsfold@metoffice.gov.uk

Abstract: Over the past few decades, the oceans have stored the majority of the excess heat in the climate system resulting from anthropogenic emissions. An accurate, long-term sea surface temperature (SST) dataset is essential for monitoring and researching the changes to the global oceans. A variety of SST datasets have been produced by various institutes over the years, and here, we present a new SST data record produced originally within the Copernicus Marine Environment Monitoring Service (which is therefore named CMEMS v2.0) and assess: (1) its accuracy compared to independent observations; (2) how it compares with the previous version (named CMEMS v1.2); and (3) its performance during two major volcanic eruptions. By comparing both versions of the CMEMS datasets using independent in situ observations, we show that both datasets are within the target accuracy of 0.1 K, but that CMEMS v2.0 is closer to the ground truth. The uncertainty fields generated by the two analyses were also compared, and CMEMS v2.0 was found to provide a more accurate estimate of its own uncertainties. Frequency and vector analysis of the SST fields determined that CMEMS v2.0 feature resolution and horizontal gradients were also superior, indicating that it resolved oceanic features with greater clarity. The behavior of the two analyses during two volcanic eruption events (Mt. Pinatubo and El Chichón) was examined. A comparison with the HadSST4 gridded in situ dataset suggested a cool bias in the CMEMS v2.0 dataset versus the v1.2 dataset following the Pinatubo eruption, although a comparison with sparser buoy-only observations yielded less clear results. No clear impact of the El Chichón eruption (which was a smaller event than Mt. Pinatubo) on CMEMS v2.0 was found. Overall, with the exception of a few specific and extreme events early in the time series, CMEMS v2.0 possesses high accuracy, resolution, and stability and is recommended to users.

Keywords: sea surface temperatures; OSTIA; CMEMS v2.0; climate data record

check for updates

Citation: Worsfold, M.; Good, S.; Atkinson, C.; Embury, O. Presenting a Long-Term, Reprocessed Dataset of Global Sea Surface Temperature Produced Using the OSTIA System. *Remote Sens.* **2024**, *16*, 3358. https://doi.org/10.3390/rs16183358

Academic Editor: Chung-Ru Ho

Received: 1 August 2024
Revised: 27 August 2024
Accepted: 29 August 2024
Published: 10 September 2024

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. The Purpose of the Reprocessed Dataset

Over the past few decades, the oceans have stored more than 93% of the excess heat in the climate system resulting from anthropogenic emissions (Section 4.1.1 of [1]). Obtaining spatially complete global sea surface temperatures (SSTs) is therefore essential to our understanding of the past, present, and future climate [1]. However, the research community is faced with the challenge of generating such globally complete SST datasets from observations of SST that naturally have gaps in their spatial coverage. In situ observations, such as those from ships or buoys, can only provide an SST measurement from their immediate location. Satellites have far wider coverage, but have the complication of temporal and spatial gaps in their coverage due to, for example, cloud cover. Data producers have developed techniques to merge these multiple data sources together and fill in any remaining gaps. The result is a global map of SST values (also called a Level 4 (L4) analysis [2]).

One such system capable of producing an L4 analysis is the Met Office Operational Sea Surface Temperature and Ice Analysis (OSTIA). It can produce a daily global SST and lake temperature L4 analysis in near real time (NRT). This analysis is used as input to the Numerical Weather Prediction (NWP) system of the European Centre for Medium-Range Weather Forecasting (ECMWF) [3]. It was also used in the Met Office NWP system until its transition to a coupled ocean–atmosphere system in May 2022 (although the OSTIA lake temperatures remain in use). Its versatility means that it can also be reconfigured to generate an L4 analysis using historical observations for users who require a consistently processed, long-term climate data record (CDR) of SST.

The Copernicus Marine Environment Monitoring Service (CMEMS) and its predecessor, MyOcean, have funded the development of OSTIA reprocessed products. The first version of this covered the period of 1985–2007 (the latest version of which we will refer to as CMEMS v1.2).

The OSTIA analysis system has been continuously upgraded throughout its lifetime, and in 2018, the OSTIA system was used to produce a completely new CDR using a wider range of input observation datasets (CMEMS v2.0). The first tranche of data covering the period of 1982–2018 was made publicly available in December 2019. It has been extended further to May 2022, with further extensions planned. Information on how to obtain the data can be found in Section 11. The production of extensions to the dataset is currently funded by the UK Marine and Climate Advisory Service (UKMCAS).

The OSTIA system has been extensively upgraded since CMEMS v1.2, so one of the main questions that this paper will attempt to address is whether CMEMS v2.0 represents SSTs more accurately (lower bias and standard deviation) over the whole time period compared to v1.2.

In addition, we also wish to analyze the uncertainties inherent in the CMEMS v2.0 analysis compared to v1.2, as well as the effective resolution of the dataset.

A longer-term CDR would also allow researchers to investigate events further back into our past and closer into our present with a higher degree of confidence.

This paper is structured as follows: First, in Section 1.2, the two versions are briefly compared with other equivalent long-term products to give a perspective on how different data records can be used for different research purposes. In Section 2, an overview is given of the OSTIA system and the configuration used to generate CMEMS v2.0, followed by Section 3, which describes the input datasets used. Section 4 describes the methodology behind analyzing the accuracy of the two CMEMS products using various statistical techniques, and Section 5 presents the results. Further investigations into the products' uncertainties and feature resolution are presented in Sections 6 and 7, respectively. An investigation into how stable the CDR was during two major volcanic eruptions is presented in Section 8. Section 9 contains the summary and conclusions, and Section 10 contains a list of acronyms used in this paper.

1.2. Overview of Currently Available Long-Term Daily SST Products

There are a wide variety of long-term SST L4 products available; however, there are only four daily SST L4 products that extend to 1982 or earlier, which are summarized in Table 1. A study was conducted to compare the accuracies of eight different L4 SST products (including the four long-term products explored here), and the relative accuracies were determined to be very similar across the eight datasets [3].

The previous version of the CMEMS analysis (v1.2) is also included in Table 1 for comparison purposes.

The development of the four different data records is the result of different practical and research requirements. Perhaps the most fundamental difference between their outputs is which point in the water column their SST values represent. The term SSTdepth is used to denote the SST values at different points in the water column. A measurement from a specific depth is expressed as SSTzm (where z is the depth in meters).

Table 1. Available SST L4 CDRs and their characteristics.

Analysis Name	SST Type	Time Span	Grid Res (Degrees)	Satellite Datasets Used	In situ Dataset
MDGSST	SST depth (1 m)	1 January 1982–Present	0.25° (approx. 27.78 km at the equator)	AVHRR and MetOp-A, WINDSAT, AMSR2/E	Ships and buoys
NOAA DOISST v2.1	0.2 m nominal	1 September 1981–Present		AVHRR 7-19 and MetOp-A/B AVHRR	Ships, moored and drifter buoys, and Argo
ESA SST CCI v2.1	Daily average SST depth (0.2 m)	1 September 1981 to 31 December 2016 2017–2022 from C3S	0.05° (approx. 5.55 km at the equator)	AVHRR, AATSR and ATSR1/2 SLSTR in C3S extension	Does not use in situ
CMEMS analysis v2.0	Foundation SST	1 October 1981-Present (with a 6-month delay)		AVHRR 7-19, MetOp-A, ATSR1/2, AATSR, GOES-13/16, SEVIRI, SLSTR-A/B, AMSR2/E, and GMI	Moored and drifter buoys
Previous version: CMEMS analysis v1.2		1 January 1985–31 December 2007		AVHRR, AATSR and ATSR1/2	Ships, moored and drifter buoys

During the day, solar radiation heats the surface of the ocean, which can result in the formation of a near-surface warm layer, which cools down at night. Consequently, SSTs undergo a diurnal cycle. The temperature from which the diurnal cycle grows during the day is termed the foundation SST (SSTfnd) [2]. The exception to this is where there is strong wind (defined as when the speed is $\geq$6 m/s) which causes mixing of the water layers, in which case the SSTs nominally in the warm layer can be assumed to be equal to SSTfnd through the day [4,5]. Additionally, the effects of diurnal heating and cooling are reduced with depth, and SSTfnd can be approximated by temperature values recorded between 3 and 5 m [6].

The variations in water temperatures throughout the water column have implications for interpreting satellite datasets since most sensors are only capable of measuring a very thin layer at the surface of the ocean (SSTskin). Due to heat transfer from the water to the atmosphere (by the mechanism of evaporation and diffusion), SSTskin tends to be cooler than the water beneath it. Research has shown that, where the wind speed is greater than 6 m/s, the average difference in temperature between the SSTskin and the SST at a depth of 0.2–1 m is -0.17 [4,5]. This correction (sometimes referred to as the skin to bulk correction) can therefore be applied to satellite datasets that represent SSTskin to convert them into SSTdepth measurements.

The CMEMS products investigated in this paper represent foundation SSTs, whilst the others represent depth SSTs from different points in the water column.

MGDSST:

The merged satellite and in situ data global daily sea surface temperature (MGDSST) is a daily 1-meter-depth SST product produced by JAXA ([7,8]) which uses optimal interpolation data assimilation.

In common with CMEMS v2.0, it uses data from passive microwave (PMW) instruments. Unlike infra-red (IR) instruments, PMW instruments are able to observe the surface of the ocean and, thus, retrieve SSTs in the presence of clouds at the expense of lower resolution [9]. The implications of using microwave instruments are expanded upon further in Section 3.1.4. Information on the location of this dataset can be found in Section 11.

NOAA Daily Optimum Interpolation Sea Surface Temperature (DOISST) v2.1:

The NOAA daily OISST v2.1 is the latest version in a series of long-term data records produced by NOAA. The original products produced by NOAA were the OI.v1 and OI.v2

datasets, which were 1° resolution, optimally interpolated (OI) SST 0.5 m depth products produced weekly.

OI.v1 ran from January 1985 onwards [10]. A daily version of OI.v1 was also developed as an operational SST product for NOAA's own use.

It was later upgraded to OI.v2, which spanned November 1981 to 29 January 2023 and implemented improvements to reduce bias [11]. Both of these versions were often referred to as the "Reynolds" CDR or simply "OISST" in earlier research papers and reports.

Daily OISST v1, v2, and v2.1 are newer daily 0.25° optimally interpolated SST (OISST) products. There has been some confusion between the old weekly products and the new daily products because some older research papers also referred to the older weekly OI.v1 and OI.v2 datasets as OI SST v1 and v2.

The Daily OISST v1 spanned the period of 04/01/1985 to the present day [12]. The Daily OISST v2 was later developed and expanded further backwards in time to 1 September 1981 [13]. It was later improved further with the addition of a greater number of in situ observations, the addition of Argo data, and a more sophisticated sea ice concentration to SST conversion algorithm, resulting in v2.1 [14]. This dataset is bias-corrected using drifting buoys as the baseline, and since they measure SSTs at a depth of 0.2 m. OISST v2.1 is, therefore, a 0.2 m nominal depth product [14]. Information on the location of this dataset can be found in Section 11.

European Space Agency SST Climate Change Initiative (ESA SST CCI) v2.1:

The ESA SST CCI project produced climate data records (CDRs) at different data levels: L2, L3, and finally L4. Level 2 data files contain SSTs in the form of a strip along the satellite's path. There were two types of L3 data provided: L3U (uncollated), which consisted of L2 files mapped to a latitude/longitude grid without combining data from multiple orbits, and L3C (collated), where multiple passes from the same sensor at different times are merged [2]. Historical satellite records were generated at levels 2 and 3 for the (A)ATSR and AVHRR series of sensors. These products include the skin SST as observed by the satellite and an estimate of the temperature at 20 cm depth, adjusted to the nearest 10.30 am or pm local mean solar time (chosen because the SST at 10.30 am/pm is a good approximation to the daily mean SST) [15]. The L4 CDR product was generated using the OSTIA system. Despite the fact that OSTIA is the same system used to generate the CMEMS v2.0 dataset, the resulting products differ, as ESA SST CCI L4 is a satellite-only product representing the daily average SST at 20 cm depth and uses a restricted set of satellite input data to achieve consistency over the data record [15].

The single-sensor ESA SST CCI level 3 products were also used as inputs for the CMEMS v2.0 analysis (Table 2). Before assimilation, the Donlon skin to bulk correction was applied and the data were filtered by wind speed [4,5,16].

An extension of the ESA SST CCI CDR was maintained by the Copernicus Climate Change Service (C3S) up to 2022.

A new version of the ESA SST CCI dataset (version v3) is now available covering 1980 to 2021 [17]. This was extended by C3S to 2022. Further extensions are being funded by the UKMCAS and the UK Earth Observation Climate Information Service (UKEOCIS). This ensures that there is a fully consistent CDR at levels 3 and 4 with daily updates less than one month behind the present. More information on the locations of these datasets can be found in Section 11.

Copernicus Marine Environment Monitoring Services (CMEMS) v2.0:

CMEMS v2.0 is a foundation SST product (in common with the NRT OSTIA product) and is therefore suitable for users who want a long-term, consistent data record with the same characteristics as the OSTIA product. Although CMEMS v2.0 uses data from the AVHRR/ ATSR instrument families like CMEMS v1.2 (Table 1), it uses newer versions of these datasets. Like MDGSST, CMEMS v2.0 contains PMW data, but also incorporates a wider range of satellites than any of the other products, including some that have never been used in a CDR before (as far as the authors are aware). One such example is the GMI instrument, which is designed to improve coverage in the tropics, an area which suffers

from extensive cloud cover. Further information about the input satellite datasets used and the GMI instrument is provided in Sections 3.1.1 and 3.1.4, respectively.

Table 2. Satellite and associated sensor information.

Instrument	Start/End Date	Platform	Dataset Provider	Spectrum and (Orbit)	Resolution of Input Dataset	File Format
AVHRR 7	24 August 1981–18 February 1985	NOAA-7	ESA SST CCI-CDR version 2.0. (Not to be confused with the L4 CDR dataset).	IR (LEO)	0.05 deg	L3U
AVHRR 9	4 January 1985–28 October 1988	NOAA-9				
AVHRR 11	12 October 1988–1 September 1994	NOAA-11				
AVHRR 12	16 September 1991–14 December 1998	NOAA-12				
AVHRR 14	19 January 1995–31 December 1999	NOAA-14				
AVHRR 15	1 January 1999–31 December 2001	NOAA-15				
AVHRR 16	1 June 2003–31 December 2006	NOAA-16				
AVHRR 17	1 December 2002–31 December 2009	NOAA-17				
AVHRR 18	5 June 2005–31 December 2009	NOAA-18				
AVHRR 19	22 February 2009–31 December 2016	NOAA-19				
AVHRR MTA	21 November 2006–31 December 2016	MetOp-A				
AVHRR 19	1 January 2017–31 December 2018	NOAA-19	C3S			L3C
AVHRR MTA	1 January 2017–30 September 2021	MetOp-A				
AVHRR MTB	1 October 2021–Present	MetOp-B				
ATSR1	1 November 1991–1 January 1996	ERS-1	ESA SST CCI			L3U
ATSR2	1 August 1995–22 June 2003	ERS-2				
AATSR	1 December 2002–31 December 2012	EnviSat				
ABI-13	1 April 2010–14 December 2017	GOES-13	EUMETSAT OSI SAF	IR (GEO)		L3C
ABI-16	17 December 2017–Present	GOES-16				
SEVIRI	1 June 2004–Present	MetoSat 8-11				
SLSTR A	1 January 2017–Present	Sentinel-3a	C3S	IR (LEO)		
SLSTR B	15 December 2019–Present	Sentinel-3b				
AMSR 2	1 December 2013–Present	GCOM-W	REMSS	PMW (LEO)	25 km	L2P
GMI	4 March 2014–Present	GPM_Core				L3U

For clarity, the datasets are loosely grouped into families of instruments and their source institutions.

Information on how to obtain the CMEMS v2.0 analysis can be found in Section 11.

CMEMS v1.2:

The previous CMEMS product, in addition to using fewer sensors and having a shorter timespan than CMEMS v2.0, was also produced using an older version of OSTIA which used an Optimum Interpolation-like data assimilation system [18]. The current version of OSTIA used to generate both the ESA SST CCI v2.1 and CMEMS v2.0 datasets uses a variational data assimilation system called NEMOVAR [19], which was developed alongside the Nucleus for European Modeling of the Ocean (NEMO) project [19–21].

2. Description of the OSTIA Climate Configuration

A detailed explanation of the current version of the OSTIA NRT processing system is provided in [16]. This version of OSTIA was also used for processing CMEMS v2.0. While there have been upgrades to the NRT OSTIA configuration since, these only affected the satellite inputs to the NRT analyses. Rather than repeat the contents of [16], we instead focus on where the OSTIA climate configuration differs from the NRT configuration described in the paper.

2.1. Assimilation of Input Datasets

For each day of the OSTIA NRT analysis, the input SST observations were taken from 1800 UTC on the previous day up to 0600 UTC on the day after the analysis. However, if they were included in the previous analysis, they were not reused. These observations were extracted from the MetDB, the Met Office's observations database, which stores data received in NRT.

In the CMEMS v2.0 configuration, all the satellite data and in situ datasets for the analysis day, plus a 24 h overlap on each side, were assimilated. Consequently, we used each observation three times, but the uncertainty for the input data was increased for the data on the days on either side of the analysis day. These observations datasets were not extracted from the MetDB, but rather from files provided by their source institution; more information is given in Section 3.

2.2. Satellite Dataset Quality Control (QC)

In common with the NRT configuration, the CMEMS v2.0 configuration only accepts SSTs which have the two highest quality levels (4–5) assigned to them and filters by wind speed.

However, in addition to the quality level for each SST, the input data files themselves have a file quality level attribute which is set to a number from zero to three, with three being the highest [2]. The CMEMS v2.0 configuration will only accept files with a quality level of three, with the exception of ATSR2 during the time period of 7 February 2001–6 July 2001, where the file quality level was degraded due to a gyroscope failure on the ERS2 satellite (Section 3.1.3). This exception was made because there are few other satellite datasets available at the same time, and hence, the ATSR2 data were still valuable regardless of the potentially reduced quality.

2.3. In Situ Dataset QC

In the NRT system configuration, the in situ data were taken from the MetDB, which stores in situ data received in NRT via the Global Telecommunication System (GTS). The NRT QC system involves the application of a reject list, which removes poor-quality in situ instruments from the input dataset. The reject list is updated each month by comparing recent in situ data to the OSTIA analyses. If the instrument data improves later, it can be removed from the reject list. For more information, see Section 2.1.3 of [16].

In the climate configuration, the input in situ datasets were taken from the Met Office Hadley Centre Integrated Ocean Database version 1.2.0.0 (HadIOD.1.2.0.0) [22,23]. More information is given in Section 3.3. The buoy data used in the assimilation had Met Office Hadley Centre QC applied, which helped to ensure their consistency over time. The QC suite comprises, for example, bias corrections, basic sanity checks, climatology checks, and buddy checks.

As described in [16], uncertainty is assigned to in situ data by combining two components in quadrature:

1. A platform-dependent uncertainty.
2. A geographically varying uncertainty, which can be pre-calculated [24–26].

In [26] it was found that the platform-dependent error variances (component 1) for moored and drifting buoys were 0.16 K^2 and 0.04 K^2, respectively. The geographically varying uncertainty (component 2) was adjusted to include the drifting buoy error variance.

An error variance value of 0.12 K^2 (0.16–0.04 K^2) was then added to moored buoy data before it was assimilated into the OSTIA analysis (component 1). However, due to concerns about over-fitting drifting buoy data, it was subsequently decided to add 0.12 K^2 to both the drifting and moored buoys' error variances before assimilation.

Therefore, the in situ pre-processing can be summarized in the following points:

1. Reject points with a QC value greater than one (one is a pass and four is a fail).
2. Add HadIOD bias corrections to the in situ SST values (in HadIOD.1.2.0.0, the correction values for buoys are zero, but this step is good practice).
3. For each observation, an error standard deviation value is assigned by combining, in quadrature, a platform-dependent error variance (0.12 K) with a geographic uncertainty value extracted from the matching location in the spatial error variance data file.

The uncertainty values calculated in step 3 were used by the NEMOVAR data assimilation scheme when combining the various satellite and in situ datasets.

2.4. Selection of Reference Satellite Sensor

OSTIA compensates for bias in the satellite data it uses by defining a reference dataset consisting of in situ buoy data and high-quality satellite data. The reference satellite data were chosen from the most accurate instrument(s) available. The reference sensors used in the NRT OSTIA system were the two VIIRS instruments currently on board the SNPP and NOAA-20 satellites and the two SLSTRs on Sentinel 3A and B. In the case of the CMEMS v2.0 processing, a variety of reference satellite sensors were used throughout, as no single sensor covers the entire record. Where available, the (A)ATSR and SLSTR sensors were used. Outside the period in which they were available, the reference sensor was chosen through expert judgement from the available options. Information about which sensors were used is given in Figure 1 and Table 3. The implications of choosing the correct reference sensor are expanded upon in Section 5.2. An explanation of the terms and acronyms is given in the glossary (Section 10).

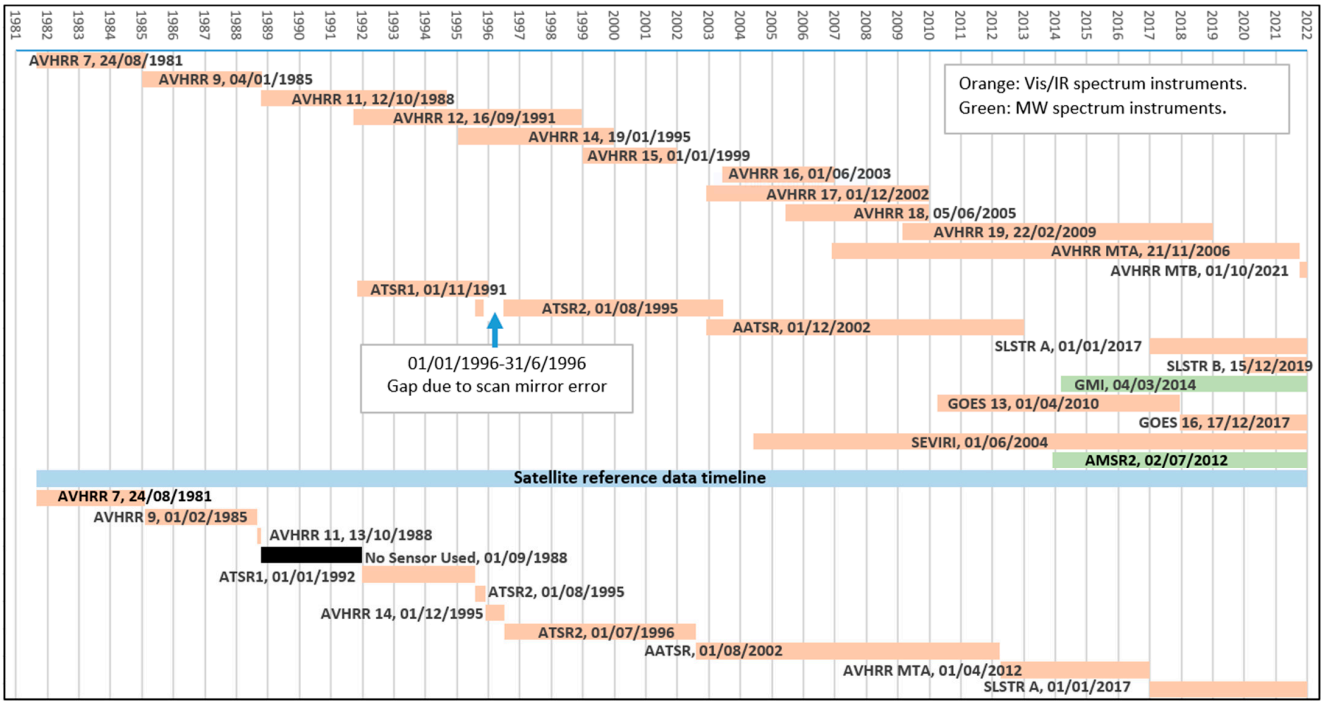


Figure 1. Timeline of satellite and reference sensors. The reference sensor timeline is below the blue line.

Table 3. Reference sensor dataset timeline.

Sensor	Start Date	End Date
AVHRR 7	24 August 1981	31 January 1985
AVHRR 9	1 February 1985	31 August 1988
No Instrument	1 September 1988	12 October 1988
AVHRR 11	13 October 1988	31 December 1991
ATSR1	1 January 1992	31 July 1995
ATSR2	01 August 1995	30 November 1995
AVHRR 14	1 December 1995	30 June 1996

Table 3. *Cont.*

Sensor	Start Date	End Date
ATSR2	1 July 1996	31 July 2002
AATSR	1 August 2002	31 March 2012
AVHRR MTA	1 April 2012	31 December 2016
SLSTR A	1 January 2017	31 December 2021

3. Input Datasets

3.1. Satellite Datasets

3.1.1. Sensors Used and Their Timeline

As outlined in Section 1.2, CMEMS v2.0 uses a greater range of observation sources than other similar datasets that are currently available.

Figure 1 shows the timeline of the satellite sensors and reference sensors used in the analyses. Information about the datasets themselves is displayed in Table 2, which includes loose groupings according to the instrument types and families. The reference sensor dataset timeline is given in Table 3. An explanation of the terms and acronyms is given in the glossary (Section 10).

Although one of the aims of CMEMS v2.0 was to use as many available satellite SST data sources as possible, it was not possible to incorporate data from the VIIRS sensor as originally intended, since assimilating a dataset this size alongside the already large volumes of data from other sources resulted in the OSTIA processing workflow suffering from memory overflow errors. Given the wide range of satellites already assimilated, it was decided to exclude VIIRS from the inputs for the time being. CMEMS v2.0 continues to be extended to the present day with a 6-month delay, and the input datasets are under continuous review to ensure the best-quality analysis.

3.1.2. Sources of Satellite Datasets

The ESA SST CCI CDR v2.1 release provided a record of L3U AVHRR/(A)ATSR satellite datasets from 1981 to December 2016 [15]. These datasets were the basis of the L4 CDR already discussed in Section 1.2. Later on, the Copernicus Climate Change Service (C3S) instigated a project to extend the AVHRR data series from 2017 (as L3C files) with a short delay to real time (a month's delay) using the same software and algorithms used to generate the ESA SST CCI dataset [27]. The same project also added the SLSTR-A/B instruments to the repertoire of datasets (Table 2).

Since the production of CMEMS v2.0, ESA SST CCI v3 has become available, it will be used for extending CMEMS from 2023 and beyond [17,28,29].

All the geostationary satellite data used in the CDR were produced by EUMETSAT's Ocean and Sea Ice Satellite Application Facility (OSI-SAF). Initially, the GOES-EAST location, covering the east coast of the USA, was monitored by the Geostationary Operational Environmental Satellite (GOES-13) satellite, but was later replaced by GOES-16 in 2017. The Spinning Enhanced Visible and InfraRed Imager (SEVIRI) dataset is a continuous data series compiled from Metosats 8–11 [30].

All the PMW datasets used in CMEMS v2.0 originated from Remote Sensing Systems (REMSS), which is also the only agency that has generated a SST dataset from the GMI (GPM Microwave Imager) instrument carried aboard the Global Precipitation Measurement (GPM) satellite [31–33]. REMSS also generated data for the Advanced Microwave Scanning Radiometer collected from the EOS satellite (AMSR-E) [34,35], and also AMSR-2 (from the GCOM-W satellite) [35], using the same process for each. However, AMSR-E was not included in the reprocessing since its L2 dataset was assigned an erroneously low data quality flag, although this has been corrected in the newer L3U version of this dataset [36].

All of the input satellite datasets used in the CDR were produced in the format dictated by the Group for High-Resolution SST (GHRSST) [2].

3.1.3. Satellite Instrument Issues

The Earth Remote Sensing (ERS)-2 satellite carrying the ATSR2 instrument was launched in April 1995 [37]. The 1 January 1996–1 July 1996 gap in the ATSR2 dataset (Figure 1) was due to a scan mirror error in the instrument where its rotation rate deviated from the optimum value of 6.7 Hz. This led to intermittent issues with the dataset's quality, and the instrument was shut down temporarily as a result of the mechanism overheating. The performance since the shutdown has been judged to be good [38].

By the year 2000, all but one of the ERS-2 six navigation gyroscopes had failed. In January 2001, the team began a project to reconfigure the satellite software to use other instruments for navigating without needing gyroscopes. They were eventually successful in developing a Zero Gyro Mode (ZGM) in June 2001, allowing ERS-2 to resume normal operations [39] with similar accuracy to before ZGM was implemented, until the failure of the onboard tape recorder in June 2003 ended the mission [37,40].

When the NOAA platforms (which carry AVHRR 7 to 19) are first launched, their orbits cross the equator near 8 am or 2 pm, allowing the satellite to spend half an orbit in Earth's shadow/daylight side. However, they do not use fuel to maintain a constant orbit; therefore, their equator-crossing times slowly drift towards the twilight times (6 am–6 pm) [15]. This results in one side of the satellite receiving more sunlight, leading to the instruments heating and a degradation in their quality. Therefore, some AVHRR data feeds were halted even though the satellite was still functional and producing data.

This degradation in AVHRR-9 is the reason that it was not designated as a reference sensor for 41 days between 1 September 1988 and 12 October 1988. Only in situ data were used for bias correction until AVHRR-11 became available.

3.1.4. Comparison of Infra-Red (IR) and Passive Microwave (PMW) Datasets

IR instruments have approximately 1 km spatial resolution (although many of the datasets that are used in CMEMS v2.0 were provided at 4 km resolution) [41], but they are unable to observe the Earth through clouds, and therefore obtain SSTs for typically only 10–15% of their swath coverage [9]. This limitation particularly affects observations of the cloudy tropics. As stated in Høyer et al. (2019), while PMW instruments have a lower resolution (approximately 50 km) and are more susceptible to contamination from coastlines and radio frequency interference (RFI), they have the great advantage of being resistant to aerosol contamination and can work in all weather conditions except rain [9].

Out of all the datasets which extend back to 1981 (Table 1), CMEMS v2.0 is the only one to use data from the GMI instrument [42]. Despite its lower accuracy compared to other PMW instruments [43], it remains a valuable dataset due to the fact that the GPM satellite orbit is inclined at 65 deg (rather than the typical ~90 deg), and therefore, it provides better tropical coverage at the expense of polar coverage [44].

3.2. Sea Ice Datasets

The sea ice data ingested into the CMEMS v2.0 production system were produced by EUMETSAT OSI-SAF. Table 4 shows the periods for which different OSI-SAF sea ice dataset versions were used in the analysis, as well as the satellites that the data originated from.

The OSI-430-b [45] sea ice dataset is an extension of the OSI-450 dataset [46,47] with the same processing chain and algorithms. However, OSI-430-b was not available for 2016–2018 when generating the analysis (although it is available now), and therefore, OSI-430 was used to fill in the gap between OSI-450 and OSI-430-b.

Table 4. OSI-SAF dataset types and satellites used for the sea ice dataset used in CMEMS v2.0.

OSI-450			OSI-430			OSI-430-b		
Satellite	**Start**	**End**	**Satellite**	**Start**	**End**	**Satellite**	**Start**	**End**
Nimbus 7 SMMR	1 January 1979	1 August 1987	DMSP F17 SSMIS	1 January 2016	1 December 2018	DMSP F16 SSMIS	1 January 2016	Present
DMSP F8 SSM/I	1 July 1987	1 December 1991	DMSP F18 SSMIS			DMSP F17 SSMIS		
DMSP F10 SSM/I	1 January 1991	1 November 1997				DMSP F18 SSMIS		

Table 4. *Cont.*

Satellite	OSI-450 Start	End	Satellite	OSI-430 Start	End	Satellite	OSI-430-b Start	End
DMSP F11 SSM/I	1 January 1992	1 December 1999						
DMSP F13 SSM/I	1 March 1995	1 December 2008						
DMSP F14 SSM/I	1 May 1997	1 August 2008						
DMSP F15 SSM/I	1 February 2000	1 July 2006						
DMSP F16 SSMIS	1 November 2005	1 December 2013						
DMSP F17 SSMIS	1 December 2006	1 December 2015						
DMSP F18 SSMIS	1 March 2010	1 December 2015						

DMSP = Defense Meteorological Satellite Program. SMMR = scanning multichannel microwave radiometer. SSM/I = special sensor microwave/imager. SSMIS = special sensor microwave imager/sounder.

3.3. In Situ Datasets

The drifting and moored buoy observations for 1982–2021 were taken from HadIOD version 1.2.0.0 [22,23]. HadIOD is a database of global historical in situ ocean temperature and salinity measurements extending from 1850 to the present (with a few months' delay). It merges observations from surface-only platforms (like ships and buoys) and observations from sub-surface ocean profiling platforms (like profiling floats and bathythermographs) and supplements these (where possible) with additional metadata, including quality flags, duplicate flags, bias corrections, and estimates of measurement uncertainty. Data were extracted from the HadIOD database into various formats for different users. To generate the analysis, we made use of data extracted into netCDF SST feedback file format [48], which is designed for compatibility with the NEMOVAR data assimilation system used by OSTIA. From these files, we selected only the drifting and moored buoy data. Initially, the coverage for both datasets was sparse, but became much more extensive over time.

HadIOD is an amalgamation of various other observational datasets and studies. For further details, see the HadIOD.1.2.0.0. user guide [23], which describes the first version of HadIOD.

4. Methodology Used to Analyze the Accuracy of CMEMS v2.0

4.1. Source of Data

The HadIOD database (Section 3.3) is not only the source for the in situ data used to generate the analysis, but also provides the reference data used to validate the analyses. However, while the CMEMS v2.0 analysis used HadIOD data extracted into a NEMOVAR-specific netCDF file format (also known as feedback file format), the validation code (Section 4.2) required HadIOD.1.2.0.0 data extracted into a netCDF format originally created for validating ESA SST CCI and, later, C3S satellite data. This is referred to as the SST CCI Independent Reference Dataset (SIRDS) [49]. The chief differences between the SIRDS files and the feedback format files were the file format (SIRDS files contain monthly SST data with separate files for different platform types) and some of the data selected for inclusion. For example, the SIRDS mooring data included Global Tropical Moored Buoy Array (GTMBA) data sourced from the NOAAs Pacific Marine Environmental Laboratory for higher native sampling frequency.

Validation of the L4 analysis was carried out using Argo and buoy SIRDS data. The buoy data were formed by combining standard and GTMBA moored buoy data with drifting buoy data originally sourced from the International Comprehensive Ocean-Atmosphere Dataset (ICOADS) release 2.5.1 [50] prior to 2017 and from CMEMS after 2017. The SIRDS dataset also contained two types of Argo data: the standard dataset, spanning 2000 to the present day, and a shorter near-surface subset containing only measurements from modified floats that continued sampling closer to the surface. The standard dataset was used here.

4.2. Methodology

The SIRDS validation code used for the analysis (see Section 4.1) is a modified version of a verification code available via GitHub [51] and is written in Python and shell script. The validation process was as follows:

1. Each daily analysis data file contains data from midnight to midnight for that day.
2. The SIRDS dataset is a monthly dataset containing observation data for the entire month as well as their latitude/longitude locations.
3. Any SIRDS observation where its timestamp is outside the analysis dataset's time window is removed.
4. Any SIRDS observation with a quality flag is removed.
5. For each analysis grid cell, the SIRDS observations within that cell are extracted. Both SIRDS and analysis values are then saved as a matchup dataset.
6. The matchup dataset is then used for calculating the mean and standard deviation of the analysis minus in situ differences, which can then be used for validation of the analysis.

In Section 4.1, we compare the two analyses (CMEMS v1.2 and v2.0) versus data collected from Argo floats. Argo data are not used in the analyses, and therefore represent an independent dataset. However, the Argo dataset is only available from 2000 onwards and is very sparse until 2003, as shown by the time series of the monthly count of SIRDS observations used for validating each CMEMS analysis (Figure 2).

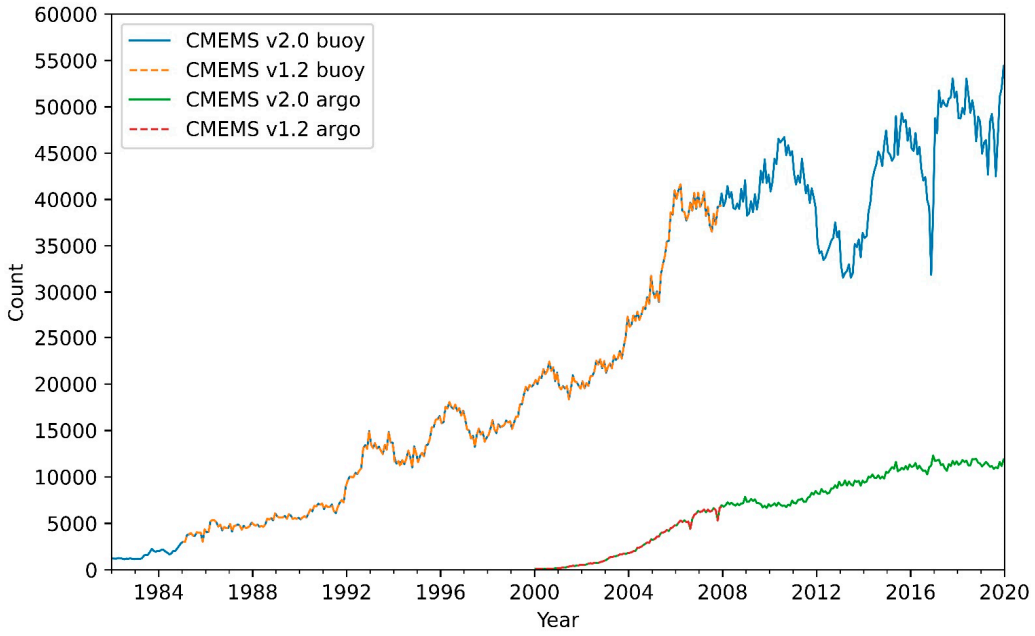


Figure 2. Monthly count of the number of SIRDS observations used for validating each of the analyses. There are slight differences between the two because of differences in the locations of the edges of the ice sheet.

In order to verify the analysis before 2003, drifting and moored buoy data were also used. Although these enabled validation of the analyses during the earlier part of the CDR, they had the disadvantage that they were not independent since they were already used in the generation of the analyses. To mitigate this issue, a temporal offset was used. While Argo observations were matched to the analysis day on which the observation was recorded, the methodology used for drifting buoy data was different. As stated in Section 2, both CMEMS v1.2 and v2.0 assimilated in situ data within a three-day assimilation window and, therefore, used buoy data from the days either side of the analysis day. Therefore, to increase the independence of the buoy dataset, data from analysis day +2 were used for verification instead. The non-contemporaneous nature of the buoy validation data and

analyses impacted on the validation statistics. Figures 3 and 4 show that the buoy data for the period of 2003 onwards were very similar to the results for the Argo data from that same time period. This indicates that, even though the buoy dataset was offset by two days, it remained a representative validation dataset that could be used for assessing the quality of the analyses prior to the Argo era.

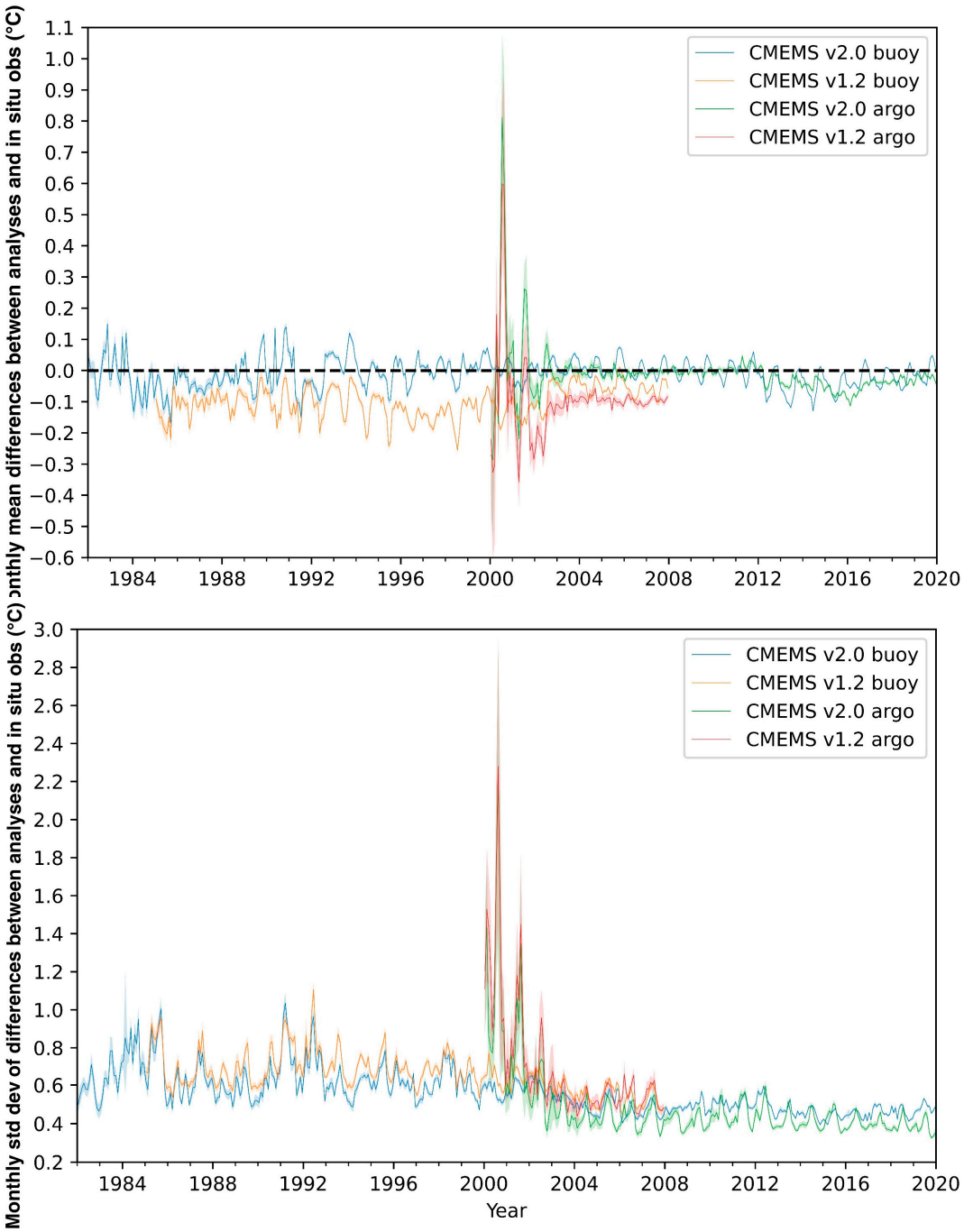


Figure 3. Analysis vs. in situ statistics plots of CMEMS v1.2 and v2.0. (**top**) Analysis–in situ mean difference; (**bottom**) analysis–in situ standard deviation. Shaded areas are 5th and 95th percentiles.

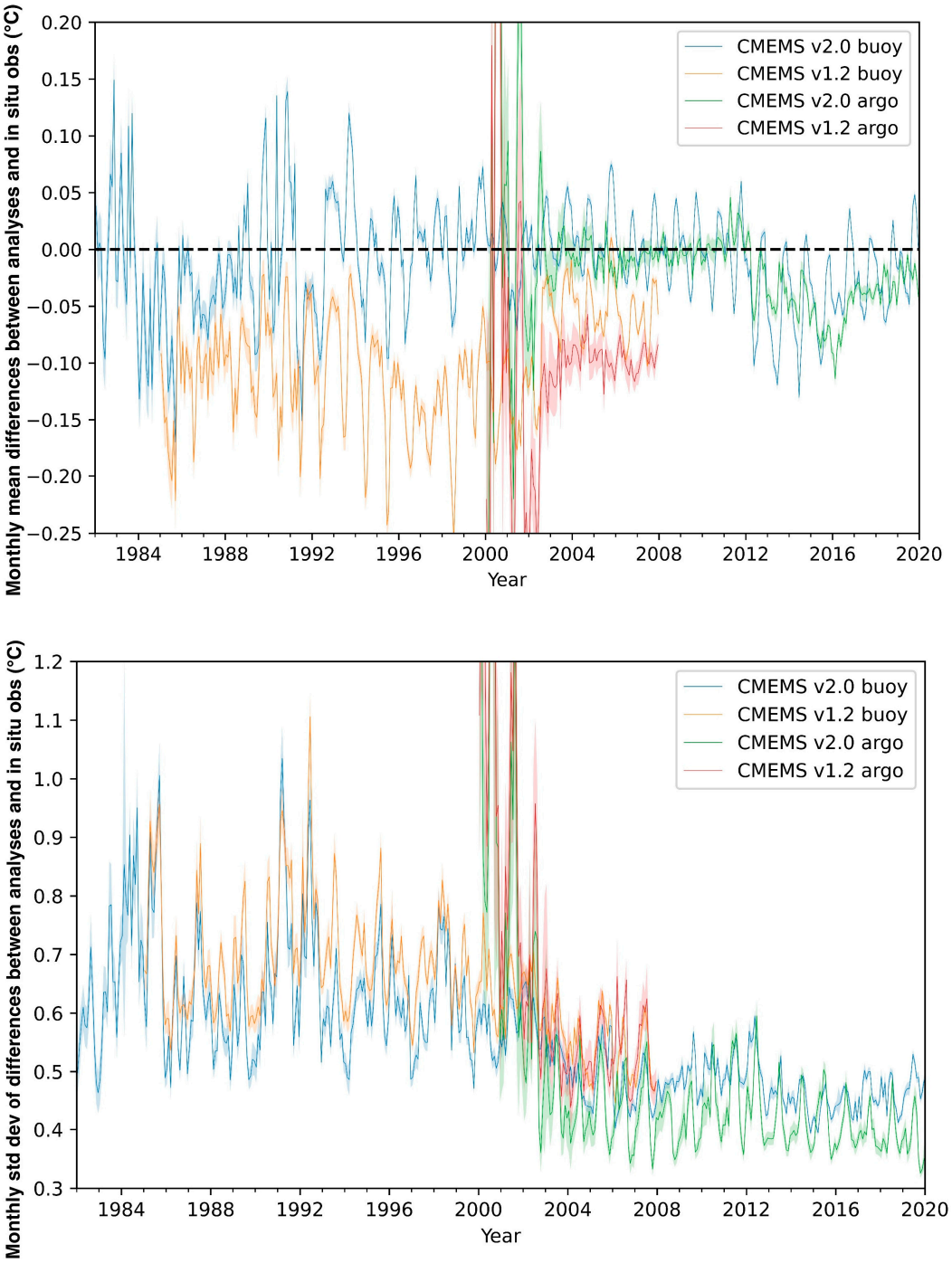


Figure 4. Analysis vs. in situ statistics (same as Figure 3, but with reduced scale) plots of CMEMS v1.2 and v2.0. (**top**) Analysis–in situ mean difference; (**bottom**) analysis–in situ standard deviation. Shaded areas are 5th and 95th percentiles.

4.3. Use of Bootstrapping to Calculate Confidence Intervals for Bias and Standard Deviation Statistics

Each in situ observation had an uncertainty value associated with it (as described in Section 3.3). The OSTIA system also calculated an uncertainty value for each cell of the L4 analysis dataset. However, in the case of buoys, it was not straightforward to propagate these analysis vs. in situ uncertainties into confidence intervals due to two factors:

- The L4 analyses were calculated using a variety of inputs, including in situ buoy data (see Section 2). This effect was minimized by using in situ data from outside the data

assimilation window. Nevertheless, errors in the analysis were not truly independent of the in situ buoy data.
- The uncertainty calculated for each in situ measurement did not consider the error caused by the offset of the analysis day (since this is very difficult to quantify).

In addition, there were two other error sources common to both Argo and buoy measurements:

- Representativity error: the difference that occurred due to the mismatch between the analysis cell size and the location of the in situ data point within this pixel.
- Uncertainty arising from the Argo and drifting buoys not measuring the foundation SST directly.

As a result, it was decided that the best method of estimating the confidence intervals of these statistics was to first calculate the in situ vs. analysis statistics as normal and then, using the well-established bootstrap method with 10,000 iterations, determine the 5th and 95th percentiles of the distribution in the statistics. The following references provide a detailed explanation of the bootstrap method [20,52], and a brief explanation of the methodology is provided below:

1. Take an original dataset A containing η number of data points.
2. Randomly draw η number of data points from A to create a resampled dataset B. Due to the random nature of the selection, B will contain some data points more than once.
3. Save the resampled dataset B.
4. Repeat the resampling process for the number of chosen iterations (i = 10,000 in this case).
5. For each of the iterations, the statistic x can be calculated (mean and std dev in this case).
6. The result is a bootstrap distribution indicating the variability in x.
7. The confidence intervals are then calculated from this bootstrap distribution (5th and 95th percentile in this case).

5. Assessment of the Accuracy of CMEMS v2.0

5.1. Results

By comparing the monthly average Argo and buoy statistics for both analyses (Figures 3 and 4, which contains the same information as Figure 3 but with a smaller scale for clarity), some key results are:

1. From 2003 onwards, the buoy and Argo validation statistics are similar. This implies that, even though the buoy dataset is offset by two days, it still remains a representative validation dataset that can be used for assessing the quality of the analyses prior to the Argo era.
2. The Argo dataset is initially very noisy from 2000 to 2003 due to the sparse network. The noise reduces as the network matures.
3. The analysis minus in situ standard deviation values for CMEMS v2.0 are slightly lower than those of v1.2.
4. The analysis minus in situ mean difference values show that the SSTs produced by CMEMS v2.0 are closer to zero (less biased) relative to the in situ temperatures than CMEMS v1.2.
5. CMEMS v2.0 shows a cool bias between 2012 and 2020; this is investigated further in Section 5.2.
6. The validation statistics display a clear seasonal variability.
7. The shaded areas that represent the 5th and 95th percentiles of the dataset are extremely small, indicating the high precision of the datasets. Consequently, they are difficult to observe in the plots.

Table 5 gives global statistics for the two analyses versus in situ data for different time periods. The first two columns compare CMEMS v1.2 and v2.0 for the two sets of validation data (buoy and Argo). The final two columns contain statistics for CMEMS v2.0 only. The

bias values for both analyses were below the required measurement bias of 0.1 K stated by [2], and CMEMS v2.0 also outperformed v1.2.

Table 5. Globally averaged analysis − situ statistics for both analyses.

	Buoy Time Period 1985–2007		Argo Time Period 2000–2007		Buoy Time Period 1982–2020		Argo Time Period 2000–2020	
	Analysis − Buoy Bias	Analysis − Buoy Std Dev	Analysis − Argo Bias	Analysis − Argo Std Dev	Analysis − Buoy Bias	Analysis − Buoy Std Dev	Analysis − Argo Bias	Analysis − Argo Std Dev
CMEMS v1.2	−0.09018	0.63064	−0.09884	0.56086				
CMEMS v2.0	0.00334	0.56782	−0.00668	0.46151	−0.01039	0.51653	−0.02664	0.41992

5.2. In Situ vs. Analysis—Investigation into 2012–2020 Bias Increase

It can be seen in Figure 3 that the average difference values for CMEMS v2.0 decreased slightly in 2012 before gradually returning towards zero by 2020. This indicates that the analysis was biased slightly cold compared to in situ observations. The most likely explanation for this is that the reference sensors used (AVHRR-MTA starting from 1 April 2012 and then SLSTR-A starting from 1 January 2017) over the period were themselves slightly biased (Table 3 and Figure 1). To test this, the analysis was rerun from 2012 onwards using an identical configuration and dataset except that AMSR2 was used as the reference sensor. AMSR2 was chosen since this instrument dataset spanned the period under study.

The analysis versus Argo statistics for the two configurations (Figure 5) showed that the standard deviation time series (red and green line) were very similar, and, in some places, the same, as indicated by the overlapping confidence intervals. Figure 5 also shows that using AMSR2 as the reference sensor resulted in the analysis becoming warm relative to in situ observations (blue line). A map of analysis vs. Argo bias showed that this warmth was most noticeable in the tropics and high latitudes (Figure 6b), which was further confirmed by the longitudinal average plot (Figure 6d). In contrast, the original analysis (which uses AVHRR-MTA and SLSTR-A) showed a more neutral bias field across the globe (Figure 6a,c).

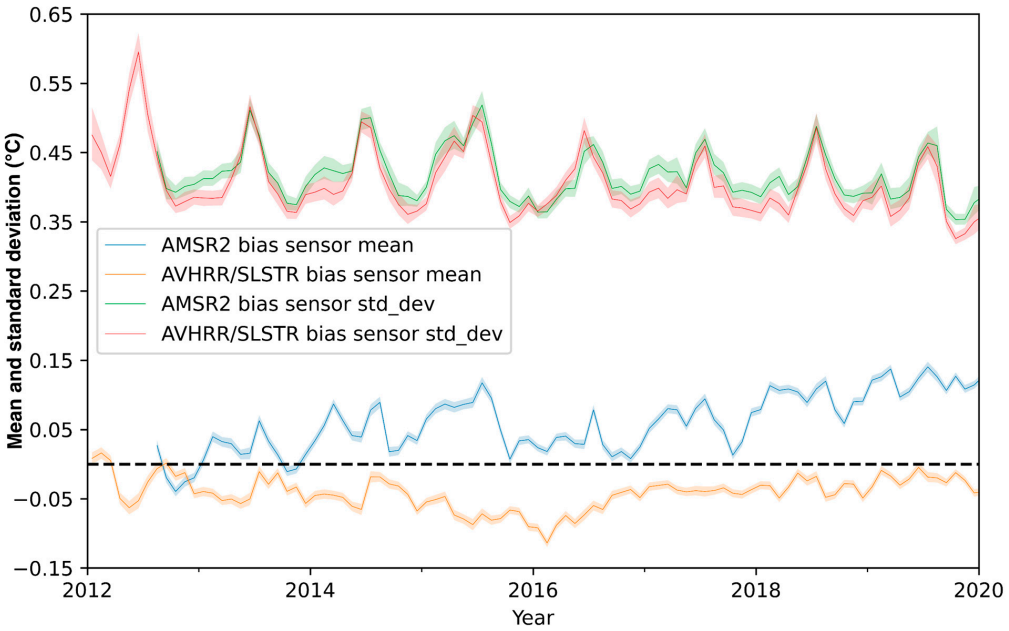


Figure 5. Analysis − Argo in situ mean difference and standard deviation for trial using AMSR2 as a reference sensor versus the AVHRR-MTA and SLSTR-A reference sensors used in CMEMS v2.0. Shaded areas are 95th-percentile confidence intervals calculated using a bootstrap method. The dashed line is zero.

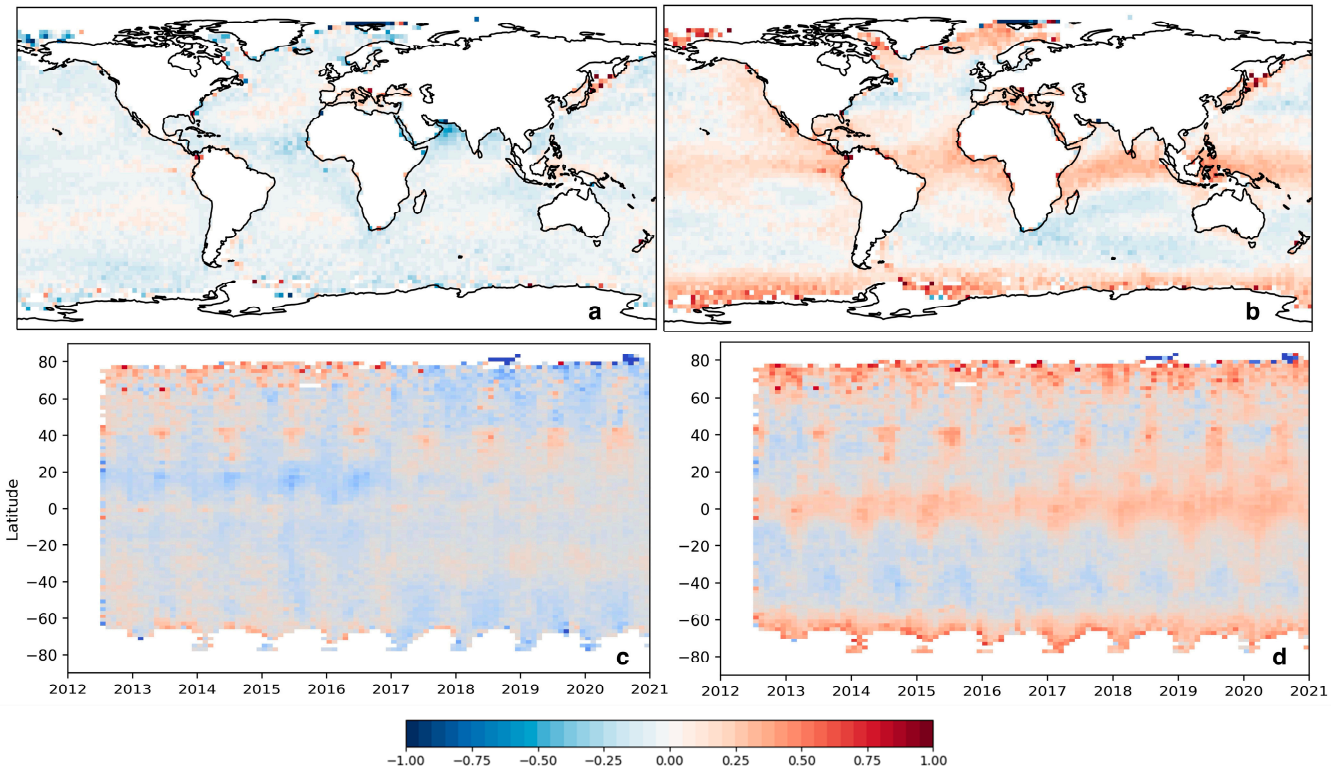


Figure 6. Binned analysis: in situ mean difference, averaged over August 2012 to December 2020. Top row—spatial bias maps with 2.5° bins: (**a**): CMEMS v2.0. (**b**): AMSR2 as reference sensor. Bottom row—longitudinal average plots with 2° bins: (**c**): CMEMS v2.0. (**d**): AMSR2 as reference sensor.

The bias in the original configuration was cooler (Figure 5, amber line), indicating that AVHRR-MTA and then SLSTR-A were the likely sources of the cool bias. However, the bias was also closer to zero than the AMSR2 trial, vindicating the decision to use AVHRR-MTA and SLSTR-A as reference sensors.

In practice, any sensor that is chosen as a reference for bias correction is likely to be slightly biased itself. Despite this, the number of different input satellite observations indicates that reference sensors are essential for ensuring the stability of the analysis.

6. Evaluation of Analysis Uncertainties

6.1. Mathematical Background

As outlined in Section 2, both sets of L4 analysis files included an uncertainty value for each grid cell. As described in [16], the OSTIA processing system generated an observation influence analysis for the purpose of determining how much the observational data contributed to the analysis at each location. A value of zero would mean that the observations had contributed nothing to the analysis at that location, and a value of one would mean that observations had a significant impact. Through a combination of this observation influence analysis with the prescribed background error standard deviation field, an analysis uncertainty value was calculated for each pixel.

It was important to evaluate whether the uncertainty values were realistic, i.e., whether the uncertainty values were consistent with the calculated discrepancies between the analyses and the in situ reference data. An effective tool for this was generating measure-of-discrepancy plots as a function of the analysis uncertainty [53]. These were produced by the SIRDS validation code used for the analysis validation.

Note that, since outliers can skew the comparison of the measure of discrepancy to the uncertainties, in this section, we use an outlier tolerant statistic, the robust standard

deviation (RSD) [54,55]. In order to calculate the robust standard deviation (RSD) statistic, first, the median absolute deviation (or MAD) was calculated using Equation (1) [55]:

$$MAD = median\left(\left|X_i - \widetilde{X}\right|\right) \quad (1)$$

The variables in the MAD are:

The analysis minus in situ SST difference values: $X_i = analysisSST_i - insituSST_i$.

The median of the whole dataset: $\widetilde{X} = median(X)$.

The MAD was then multiplied by a scale factor k to yield the robust standard deviation (RSD) (Equation (2)). For normally distributed data, the k value was 1.4826.

$$RSD = k \cdot MAD \quad (2)$$

Therefore, the RSD gave a robust estimate of the variation in analysis minus in situ discrepancies. Similarly, the robust standard error (RSE) was calculated using Equation (3):

$$Robust\ Standard\ Error = \frac{RSD}{\sqrt{NumDataPoints}} \quad (3)$$

6.2. Analysis Uncertainty Evaluation Plots—Methodology

The analysis uncertainty evaluation plots were calculated as follows:

1. The in situ observations were matched up with the corresponding SST analysis grid cell.
2. The differences between the matched-up in situ observations and the SST analysis values were binned according to the associated analysis uncertainty values. Twenty bins of 0.05 K (ranging from zero to one) were used.
3. If there were more than 100 matchups in a bin, the following statistics were calculated and plotted (steps 4–7).
4. The RSD, plotted on the y-axis with cyan vertical lines (Equation (2)).
5. The median values, plotted with red crosses.
6. The robust standard error (Equation (3)) was calculated for the data points in each bin and plotted with red vertical bars.
7. The green dashed line signifies the expected RSD envelope, which was calculated using Equation (4) using the expected in situ uncertainty.

$$ExpectedRSD = \sqrt{BinVal^2 + InsituUncertainty^2} \quad (4)$$

where *BinVal* is the analysis uncertainty bin value (which forms the x-axis of the plot). *InsituUncertainty* was set to 0.2 K for buoy data and 0.005 K for Argo data [56].

The value chosen for drifting buoy accuracy is a subject of debate, since there have been many studies regarding their accuracy using various techniques, often producing similar but different uncertainty values around 0.2 K; a review is provided by [57]. Therefore, 0.2 K was used in this analysis as a good approximation of their overall accuracy [58–60].

6.3. Results

The methodology detailed above was used to produce the plots in Figure 7. These compare the uncertainty calculated by the analysis with the *expected* uncertainty calculated from the observed variation in the analysis minus in situ values. We can use this to assess whether the uncertainties calculated by the analysis are reliable. An analysis which calculates its uncertainties correctly will have the cyan vertical lines matching the envelope in the discrepancy plots. Where the vertical lines are short of the envelope, this signifies that the analysis uncertainties are overestimated. From Figure 7a,c, we can see that not only did the v1.2 analysis consistently overestimate its uncertainties for all bins, the cyan bars

are also very similar for each bin. This indicates that the estimated analysis uncertainties have little relationship with the actual analysis quality.

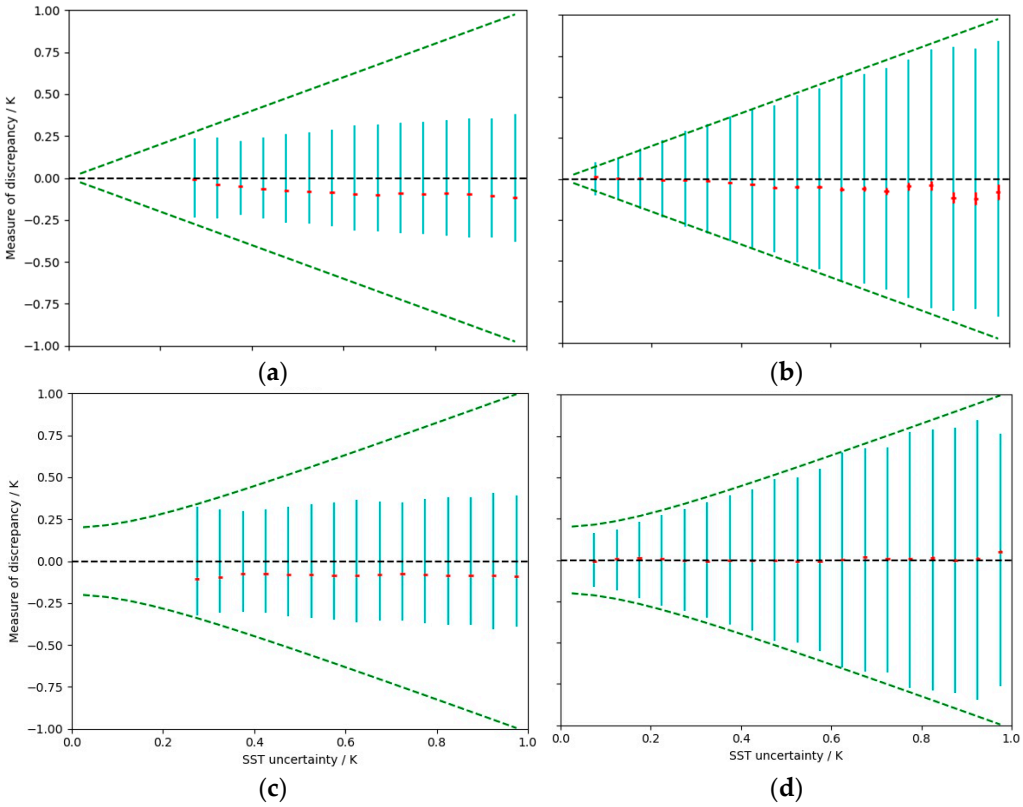


Figure 7. Top row—global Argo data, 2000–2007. (**a**) CMEMS v1.2, (**b**) CMEMS v2.0. Bottom row—global buoy data, 1985–2007. (**c**) CMEMS v1.2 (**d**) CMEMS v2.0.

In contrast, the outputs show that the analysis uncertainties for v2.0 are representative (Figure 7b,d), indicating that the new version of OSTIA is more effective at quantifying its own uncertainties than the older version.

7. Investigation of Feature Resolution Using Power Spectra and Horizontal Gradients

7.1. Introduction to Concept

The bias and mean statistics analyzed in Section 5 are an effective measurement for how well the analysis SSTs corresponds overall to ground truth. These statistics do not, however, reveal how well the analyses resolve features such as currents, and so the feature resolution needs to be analyzed. To clarify, by feature resolution, we do not mean the size of the dataset grid cells (which are 0.05 degrees in this case), but rather, the resolution of the analysis itself and, hence, its ability to resolve oceanic features.

Three regions of Interests (ROIs) were chosen for this investigation since they have high eddy-energy [61], and so we aimed provide a good test of how well the code worked.

1. The Gulf Stream.
2. Agulhas current retroreflection.
3. Kuroshio current.

7.2. Feature Resolution—Methodology

The method employed to measure the feature resolution was the power spectral density plot, which was calculated via the following steps:

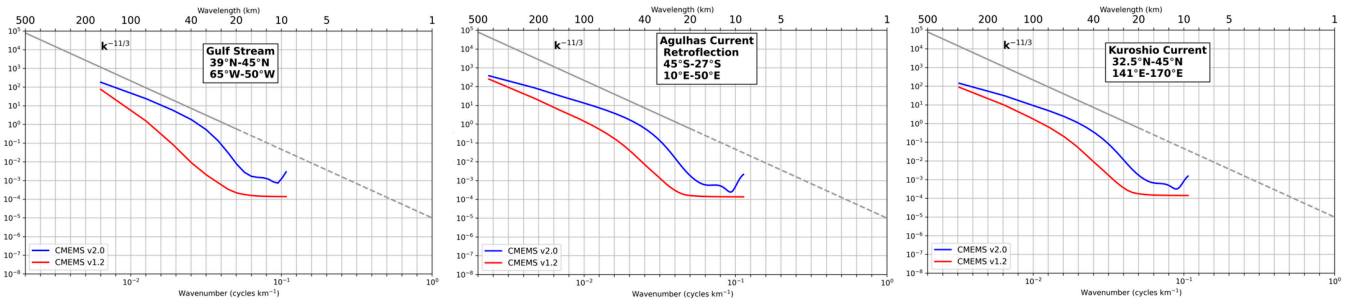
1. Loop through each day in the L4 dataset and extract a region of interest (ROI).
2. For each daily ROI dataset, loop though the latitude coordinates.
3. For each latitude, calculate a cross section using the procedure below:

- Extract all points on the longitude.
- Detrend the dataset (by taking the result of a linear least-squares fit to the dataset and then subtracting this from the dataset).
- Calculate a spectral analysis.
- Store this value.

4. At the end of the latitude loop, the mean is taken of all the values from step 3 to give the spectrum of that entire ROI for that day, which is then saved into an array.
5. The spectrum values for each day are then averaged and plotted.

However, the concept behind steps 3–5 is complex and deserves further explanation. A typical power spectrum plot is an analysis of a signal wave that varies over *time*, and therefore, the x-axis is often the period or frequency. However, in this case, the plot is of the spectral power over *distance* (the longitude in this case). If the SSTs within a horizontal cross section (step 3) followed a regular pattern, which would be expected if the L4 dataset had correctly resolved an oceanic feature such as a current, then the power spectral analysis of this L4 dataset would return a series of distances (in other words, wavelengths), each with an associated power. These would then be averaged for each day (step 4) and plotted, with the wavelength (or wavenumber k) on the *x*-axis and the power on the y-axis (step 5). For more detail regarding the mathematics behind the spectral analysis calculations (step 3), see [6].

The expected output of step 5 is key to the interpretation of resolution. Previous research investigating the power spectral analysis of a sea surface height (SSH) dataset discovered that its spectral slope followed a $k^{-11/3}$ gradient, which agreed with the theory that this phenomenon was best described by surface quasi-geostrophic (SQG) turbulence [61]. It was known previously that SSH and SST share the same magnitude of cascade [62], and therefore, a spectral analysis of an SST field would also follow a $k^{-11/3}$ gradient (grey line in Figure 8).



(**a**) (**b**) (**c**)

Figure 8. Power spectrum plots of three regions of Interest (ROI) for the time period 1985–2007. (**a**) Gulf Stream. (**b**) Agulhas Current Retroreflection. (**c**) Kuroshio Current.

As a corollary, if the cross section of the L4 dataset (step 3) only contained noise (which would be the case where the analysis was unable to resolve a feature), then the analysis would return a constant power at all wavelengths for that cross section. Therefore, the gradient of the wave power spectra line would flatten. Additionally, a power spectral analysis can only resolve signals above twice the grid size. Therefore the 0.05-degree resolution (~5.5 km at the equator) of the analysis datasets means that the spectral analysis will not detect any features smaller than ~11 km.

To summarize, the plot of a "perfect" L4 SST analysis dataset would take the form of a line on the $k^{-11/3}$ gradient and only flatten (the signal turns to noise) at an 11 km wavelength.

In practice, however, the L4 analyses displayed less power (coarser features), signified by a line below the $k^{-11/3}$ gradient, and flattened at larger wavelengths. Nonetheless, the magnitude of these characteristics provided valuable information about how well the L4 analysis can resolve signals.

7.3. Feature Resolution—Results and Implications

In Figure 8, the plot for each ROI indicates that CMEMS v2.0 (blue line) is closest to the ideal $k^{-11/3}$ gradient (grey line). In addition, the gradient for CMEMS v2.0 flattens at a wavelength of approximately 16 km (close to the theoretical maximum), whereas CMEMS v1.2 flattens at approximately 20 km. These two observations support the case that CMEMS v2.0 is able to resolve finer features than v1.2.

Power spectrum plots are essentially a measure of the *variation* in the dataset, and it does not necessarily follow that this variation is representative of what is actually happening. For example, if a varying (but incorrect) feature were added to the dataset, an apparently more favorable power spectra plot would be the result. Therefore, the power spectra must be considered together with other information before it can be concluded that it represents an improved resolution. To this end, investigating the horizontal gradients in the datasets gives further valuable information about whether the variability measured by the power spectra is the result of oceanic features, rather than an incorrect signal being added to the data.

7.4. Horizontal Gradients—Methodology

The horizontal gradient was defined in [19] as "The total horizontal SST gradient [which is] the vector sum of north-south and east-west differences for each grid point". It was calculated using the following process:

1. For each grid cell, check if any of the surrounding pixels are land or ice, and if so, ignore it.
2. Find the X gradient (*Xgrad*). This is the SST difference between the pixels on either side of the grid cell (step 1) divided by the distance between those points.
3. Find the Y gradient (*Ygrad*). This is the SST difference between the pixels above and below the grid cell (step 1) divided by the distance between those points.
4. The horizontal gradient for grid cell (step 1) is then calculated using Equation (5):

$$HorizonalGrad = \sqrt{Xgrad^2 + Ygrad^2} \quad (5)$$

These values are then averaged over time and then plotted. Ocean regions with defined features (such as currents) have larger gradient values. An analysis that struggled to resolve such features would have smaller gradients and a more uniform appearance in their horizontal resolution plots.

7.5. Horizontal Gradients—Results

The horizontal gradient plot for all the ROIs in 2007 (Figure 9) illustrates that CMEMS v2.0 was better able to resolve the current boundaries. The same plots were generated for each year between 1985 and 2007 for the three regions under investigation (not displayed) and consistently showed that CMEMS v2.0 had larger horizontal gradients than v1.2. The horizontal gradient and power spectral analyses together indicate that the feature resolution of CMEMS v2.0 was much higher than v1.2.

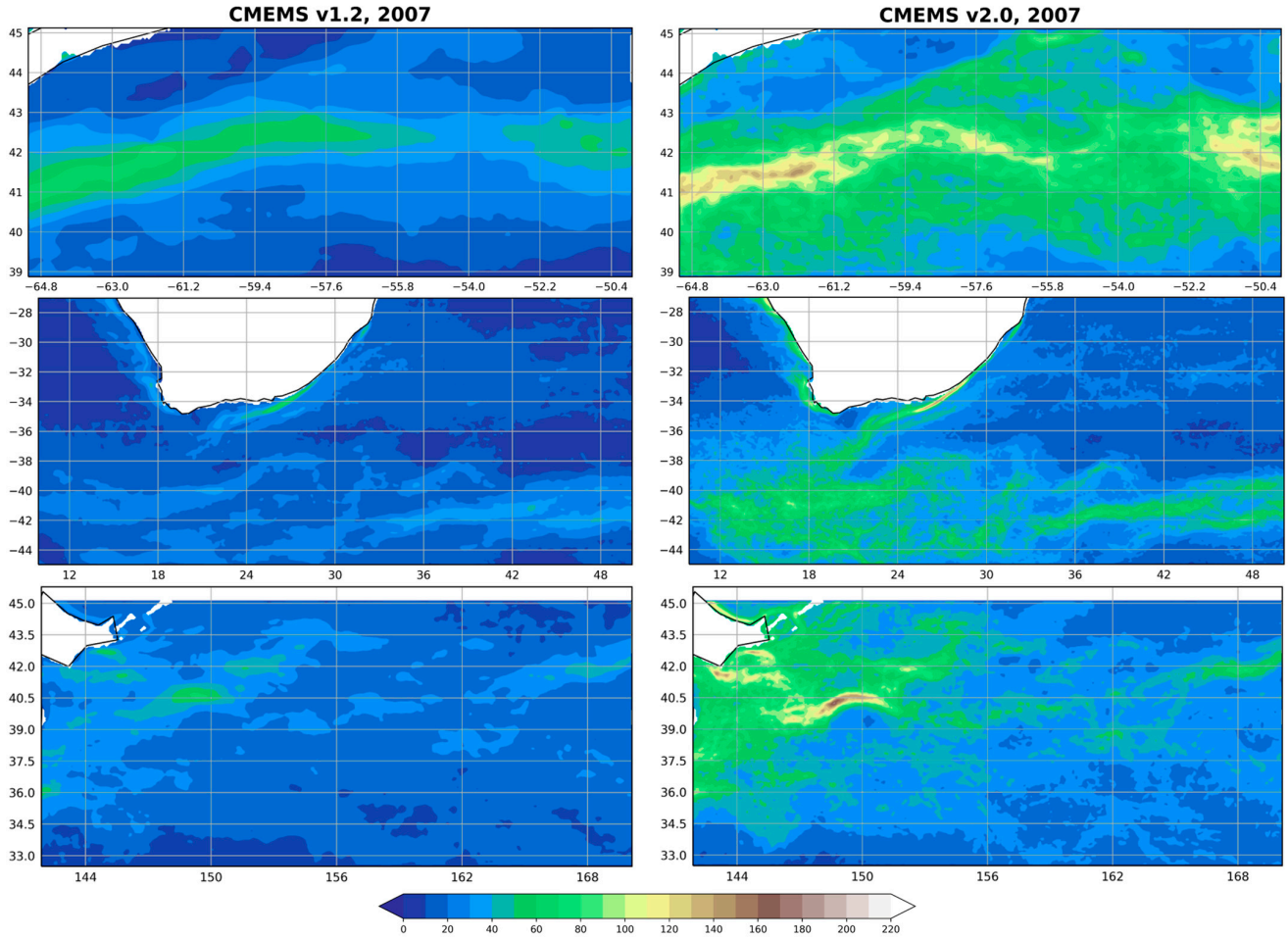


Figure 9. ROI horizontal gradients for 2007. Units: mk/km. (**Left**) CMEMS v1.2. (**Right**) CMEMS v2.0. (**Top**) Gulf Stream. (**Middle**) Agulhas Current Retroreflection. (**Bottom**) Kuroshio Current.

8. Volcanic Eruption Case Studies

8.1. Introduction

Large volcanic eruptions produce clouds of ash and sulfur dioxide (SO_2), which reflect solar radiation back into space [63] and can lead to brief cooling of the climate. In addition, they interfere with our ability to measure the climate from space. Aerosols absorb infra-red radiation emitted from the surface, which is then re-emitted at the lower temperature of the aerosol, resulting in satellite temperature measurements being cooler than the reality [63]. While normal aerosols such as water vapor or dust can be compensated for [28], large, isolated events such as volcanic eruptions provide a greater challenge for any climate analysis. In this section, the impact of the 15 June 1991 Pinatubo eruption on both versions of the CMEMS CDRs is examined. In addition, it is investigated whether the smaller March/April 1982 El Chichón eruption had any impact on the CMEMS v2.0 CDR.

8.2. Mt. Pinatubo

8.2.1. Introduction and Limitations of Validation Datasets

The June 1991 Mt. Pinatubo eruption was one of the largest eruptions in living memory; it propelled aerosols 20–30 km high. These encircled the Earth a mere 21 days after the event, but were predominantly confined to 20°S to 30°N latitudes [64]. By October 1991, the aerosols had spread further south to 50°S, and eventually returned to pre-eruption levels by October 1993 [65].

CMEMS v1.2 used two main input datasets, the NASA/NOAA Pathfinder v5.0 AVHRR and the (A)ATSR multi-mission series v2.0 [18]. Later research discovered that the

Pathfinder dataset suffered from a cold bias during this time period [37,66]. The ATSR1 instrument was operational during the eruption period, which is significant since it is less sensitive to aerosols due to its dual view operation mode, which is where two measurements are taken—one of the ocean directly below the satellite and a second of the ocean ahead of the satellite (in the case of ATSR1). The additional information provided by two views allows for better compensation for the effect of the atmosphere on the signals received by the satellite [67].

The effect of aerosols on the satellite datasets used in CMEMS v1.2 can be observed by plotting the global CMEMS v1.2 AVHRR SST data minus the analysis background field (Figure 10). These plots can be used as a proxy for biases in the satellite data (assuming that the in situ data are unbiased) because satellite data were corrected to the in situ data when the analyses were generated. Therefore, the analysis backgrounds should also be relatively unbiased [18]. It can be seen that the AVHRR data were increasingly cool in the tropics compared to the background data as time went on. For ATSR1 data, which started in August 1991, the volcanic aerosol had less impact on than on the AVHRR retrievals (Figure 10).

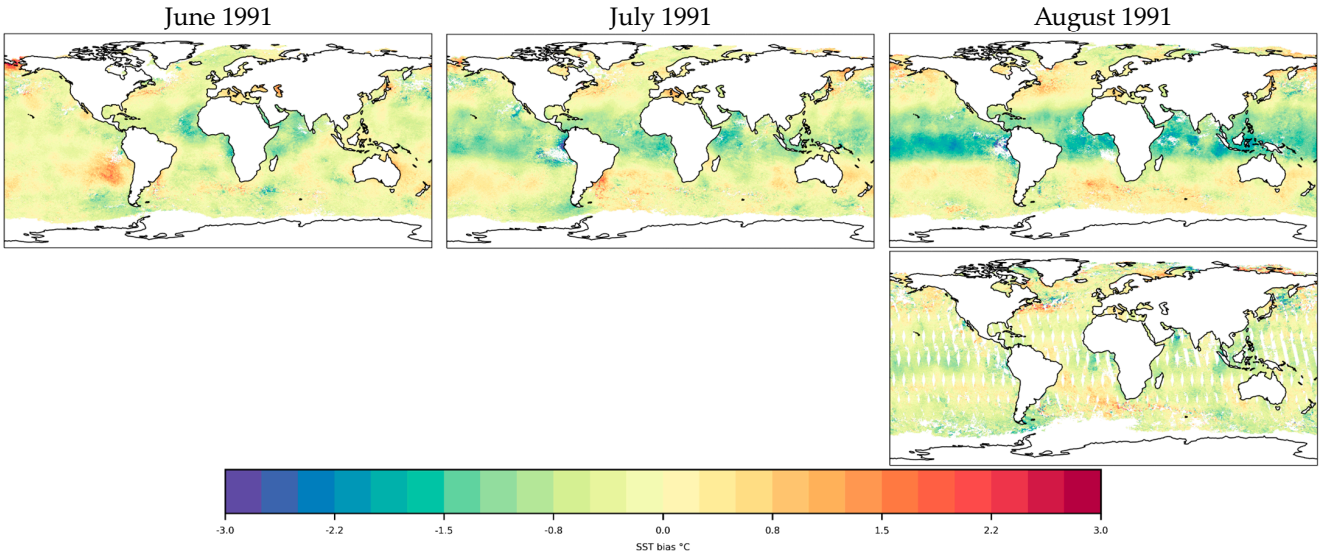


Figure 10. CMEMS v1.2 monthly average obs-background field bias plots for June–August 1991. AVHRR (**top row**); ATSR-1: (**bottom row**).

The ESA SST CCI AVHRR v2.1 dataset used for the CMEMS v2.0 analysis had an aerosol correction applied for the El Chichón and Mt. Pinatubo eruption periods (1981–1986 and 1990–1995, respectively) which used information from the high-resolution infrared radiation sounder (HIRS) dataset [15,65]. Therefore, it was expected to have been less affected by aerosol than the data used in CMEMS v1.2. The equivalent plots to Figure 10 were generated for CMEMS v2.0 (Figure 11). However, since CMEMS v1.2 only uses in situ data (ship and buoy data) for bias correction, whereas CMEMS v2.0 uses buoy and satellite data (Section 2), Figure 11 is not able to show the same effect as Figure 10, but it is included for the sake of completeness. In addition, it was not possible to make a plot for ATSR1 as it was excluded from CMEMS v2.0 until November 1991 because the data did not meet the minimum required file quality level of 3 (see Section 2 for more information), whereas CMEMS v1.2 used the slightly lower-quality data from ATSR1.

In addition to the satellite SST minus background field plots, the analysis versus drifting and moored buoy statistics for the latitudes 20°S to 30°N and the period of June 1991 to June 1992 were calculated using the same technique used to produce the results in Section 5.1 (Table 5). The percentiles were calculated using the bootstrap technique described in Section 4.3. This study period was chosen due to the most severe effects occurring

in the first year [63]. This suggests that CMEMS v2.0 had a lower mean difference (less biased) and standard deviation than CMEMS v1.2 (Table 6) during the Mt Pinatubo eruption.

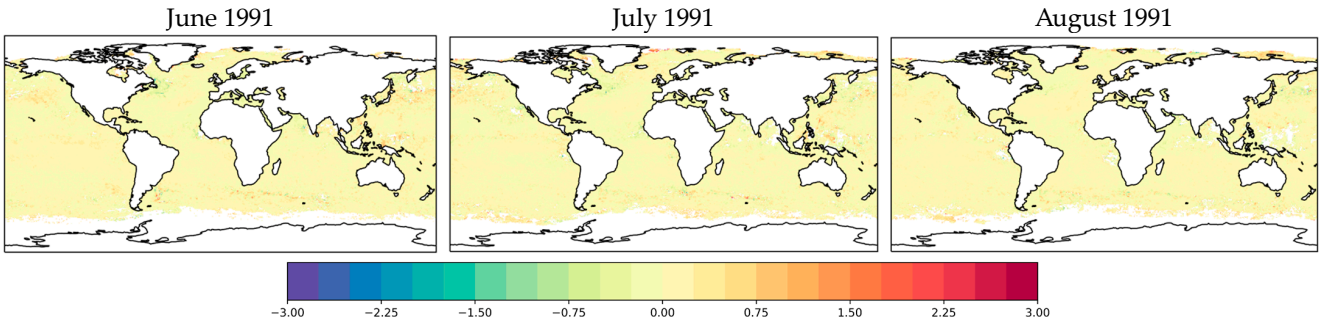


Figure 11. CMEMS v2.0 AVHRR monthly average obs-background bias plots. ATSR1 data were not used in CMEMS v2.0 during this time period.

Table 6. Analysis − insitu statistics for latitudes 20°S to 30°N over the period of June 1991–June 1992.

		Analysis − Obs Mean Percentile			Analysis − Obs Std Dev Percentile	
	Analysis − Obs Mean	**5th**	**95th**	**Analysis − Obs Std Dev**	**5th**	**95th**
CMEMS v1.2	−0.136164	−0.141289	−0.130924	0.630929	0.612625	0.650864
CMEMS v2.0	−0.100210	−0.104891	−0.095304	0.580533	0.559674	0.602480

However, drifting/moored buoy coverage (Figure 12) for the chosen region and time period under investigation was sparse, with an average of 106.7 observations per day. Therefore, a further investigation was carried out using the Met Office Hadley Centre's Sea Surface Temperature dataset version 4 (HadSST4) [68], which combines ship and buoy data to improve in situ coverage. Here, we use HadSST.4.0.1.0.

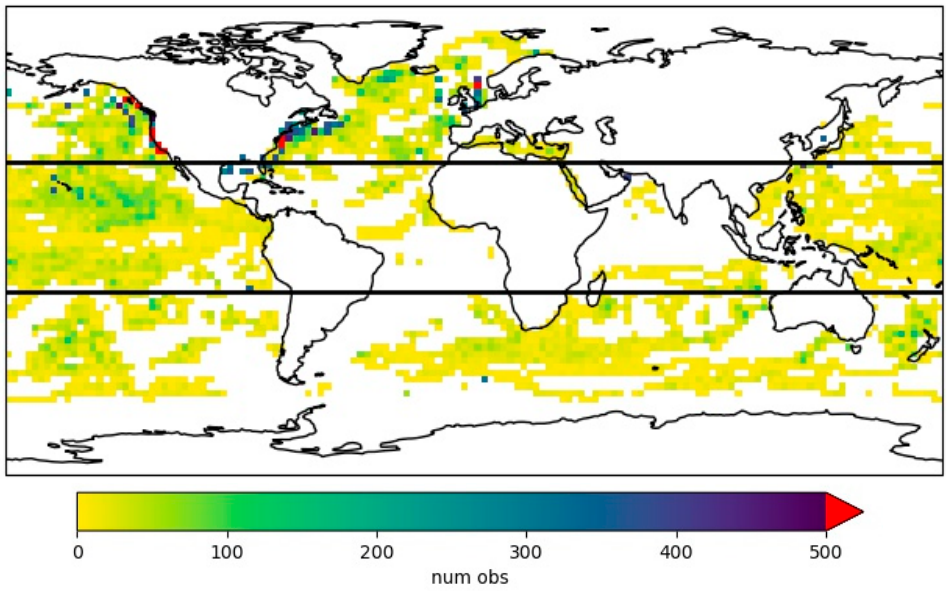


Figure 12. Number of drifting and moored buoy observations for 1991 in 2.5-degree bins. Black lines are 20°S and 30°N latitudes to demonstrate the Mt. Pinatubo study region.

8.2.2. Validation Dataset and Methodology

HadSST4 [68] is a gridded in situ SST dataset which is presented on a monthly global 5° grid from 1850 to the present day. Although its grid and temporal resolution are

much coarser than the analysis resolution, it is adequate for the assessment made here of widespread volcanic effects over several months.

Prior to the 1990s, ship observations were the chief source of in situ observations; however, these are known to suffer from persistent biases, and HadSST4 applies corrections to remove these. The uncertainty of the bias corrections is complex and is explored via an ensemble of 200 datasets, with each ensemble member providing a separate realization of the SST bias corrections.

Additional HadSST4 fields also estimate the uncertainties associated with other measurement and sampling errors. HadSST4, therefore, provides us with a bias-corrected in situ reference and its estimated uncertainty (reliability). Some caution is still needed when making comparisons to HadSST4 in case of errors not captured by the current error model.

The process by which we compared the HadSST4 ensemble to the various L4 analyses is detailed below:

1. Take the L4 analysis dataset for the region and time period (in the case of Mt. Pinatubo, this is latitude −20°S to 30°N and the time period is January 1990 to December 1992).
2. Calculate the monthly average of the L4 data (step 1).
3. Take the HadSST4 median actual SST for the region and time period (HadSST4 is already a monthly dataset).
4. The HadSST4 dataset can have gaps in its coverage. Mask out the L4 analysis data from step 2 wherever there are gaps in the HadSST4 dataset.
5. Take the spatial average of the L4 analysis (Step 4) and subtract by the spatial average of the HadSST4 dataset (Step 3) and plot.

The process by which the HadSST4 uncertainties were calculated is as follows:

1. Loop through each of the HadSST4 ensemble members.
2. For each of the 200 ensemble members, extract the region and time period.
3. Take the monthly spatial average of these data and save.
4. Take the 200 ensemble members from step 3 and calculate the standard deviation for each month.
5. Calculate the monthly HadSST4 measurement and sampling error uncertainty standard deviations.
6. Combine the monthly ensemble standard deviations (step 4) in quadrature with the monthly measurement/sampling uncertainties (step 5).
7. Multiply step 6 by two and use this to plot the ±2-sigma uncertainty range.

8.2.3. Results

The HadSST4 ensemble comparisons were repeated for each of the L4 analyses under study. The results are shown in Figure 13. The HadSST4 uncertainties are expressed in the figure in the form of lines at zero, with shading to indicate the confidence intervals (Figure 13—yellow line). This approach reduces plot clutter; the alternative is to add confidence intervals to each of the five analyses. The HadSST4 median anomaly (relative to 1961–1990) was also plotted to help us assess whether any changes were related to sudden changes in HadSST4 values (Figure 13—black line). Note that a separate assessment made of HadSST4 versus other commonly used SST datasets (ERSSTv5, [69]; COBE2, [70]; HadISST1.1, [71]) for this region and time showed differences comparable to the HadSST4 uncertainties; thus, the use of a different reference SST dataset to HadSST4 would not be expected to materially affect the results presented here.

The CMEMS v2.0 dataset (Figure 13—blue line) was relatively cool in the middle of 1991, taking a sudden dip during the eruption period. This apparent cooling soon recovered back to normality, possibly due to the addition of AVHRR-12 on 16 September 1991 and ATSR1 on 1 November 1991. Although the ATSR1 dataset is available for August–October 1991, it was not used by CMEMS v2.0 due to its rejection by more stringent QC checks. Therefore, until AVHRR-12 came online, there had only been one operational satellite (AVHRR-11) during the eruption period.

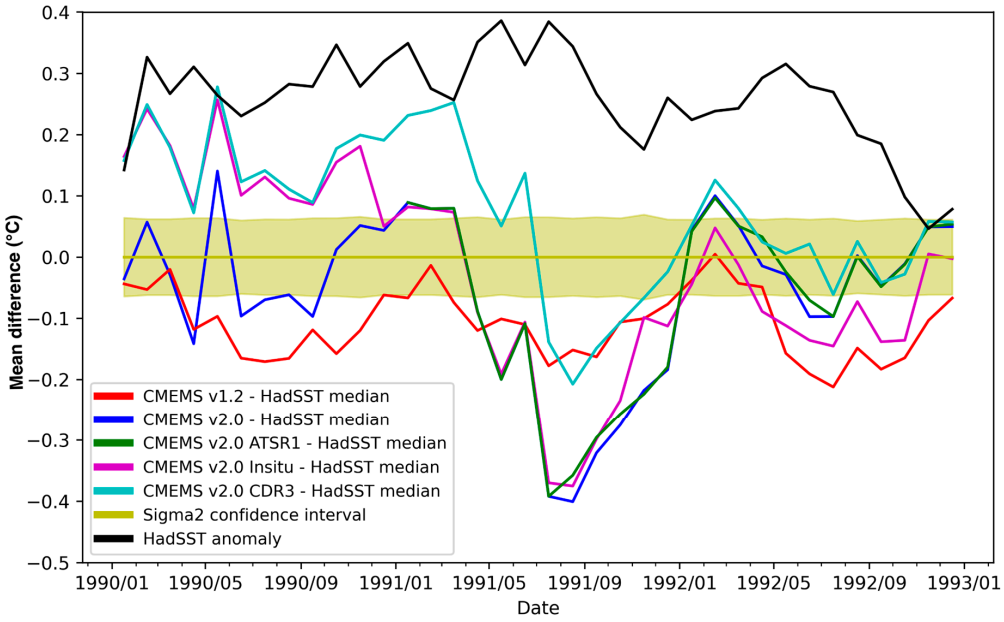


Figure 13. Spatially averaged L4 analysis minus median HadSST4 in situ difference for latitudes 20°S to 30°N.

This is in contrast to the relative stability exhibited by CMEMS v1.2 (Figure 13—red line), which used both ATSR1 and AVHRR-11 during the eruption period.

With the sparse dataset that existed in the early 1990s, in certain cases, having more numerous observations was more important than higher quality ones. To determine whether this was the cause of the slightly poorer results exhibited by CMEMS v2.0, a follow-up investigation was carried out in which CMEMS v2.0 was rerun with the same configuration except with relaxed QC checks for ATSR1 data (Figure 13—green line). The results show that the addition of the poorer-quality ATSR1 dataset resulted in only slightly reduced cooling during the Pinatubo event. Therefore, the cooling exhibited by CMEMS v2.0 cannot be explained as a result of a sparser dataset.

Another distinction between CMEMS v1.2 and v2.0 lies in the fact that v1.2 uses only in situ data for bias correction, as opposed to the satellite bias correction scheme employed by CMEMS v2.0. To investigate whether this could explain the cooling, CMEMS v2.0 was run using only in situ data for bias correction (Figure 13—Magenta line). The results were very similar to the other CMEMS v2.0 results.

CMEMS v2.0 uses the AVHRR datasets taken from the ESA SST CCI v2.1 dataset (see Section 3.1.2). Since the production of CMEMS v2.0, ESA SST CCI v3 has become available. Although the v3 dataset's treatment of stratospheric volcanic sulfate aerosol has not changed from v2.1, there have been a wide range of improvements in other significant areas [28,29].

1. A new version of Radiative Transfer (RTTOV) software (v12.3) used for cloud detection and SST retrieval.
2. Use of bias-aware Optimal Estimation (OE) retrieval for AVHRR. In v2.1, the bias-corrections were a post hoc adjustment of the SSTs.
3. Improved processing to reduce AVHRR data gaps.
4. Inclusion of tropospheric dust aerosol in the model, which should fix the large cold dust biases seen in v2 SSTs.
5. More sophisticated quality filtering allows more usable data to be obtained from the AVHRR-12 dataset.

To determine whether the v3 dataset could improve the bias statistics during the Pinatubo event, a fourth analysis was carried out where CMEMS v2.0 was rerun with AVHRR v3 dataset in place of the AVHRR v2.1 data (Figure 13—cyan line). The results

show that this version was relatively warmer, but very closely mirrored the shape of the time series from the other CMEMS v2 investigations.

Overall, a comparison with HadSST4 suggests that, during the period of June 1991 to January 1992, the CMEMS v2.0 analysis was more affected (biased cool) by the Mt. Pinatubo eruption than CMEMS v1.2. Experiments suggested that the effects seen in CMEMS v2.0 versus CMEMS v1.2 were not obviously methodological (i.e., related to the QC of ATSR1 data or the bias correction made to in situ data), and instead may have been related to the differing input datasets used.

8.3. El Chichón

8.3.1. Introduction

El Chichón erupted three times, on 29th March and on 3rd and 4th April 1982, which together ejected 7 megatons (Mt) of sulfur dioxide (SO_2) into the atmosphere. El Chichón was a smaller event compared to Mt Pinatubo, which emitted 20 Mt of SO_2 [72]. Despite it being a smaller event, within three weeks, its ash cloud had already encircled the Earth. The event was followed by a strong El Niño, which led to speculation that it had triggered it, although subsequent research concluded that this was coincidence [73]. The same research also determined that much of the aerosol cloud was located in the eastern Pacific Ocean from 0 to 30°N and from 60 to 150°W, and so it was decided to conduct a regional analysis vs. in situ validation for the period of 1982–1984.

Similarly to before, due to the sparsity of buoy observations for this period and region, HadSST4 measurements were used for the in situ validation using the same methodology employed for the Mt. Pinatubo investigation.

As an SST dataset, HadSST4 also includes El Niño effects; therefore, it is a good validation dataset to use to analyze phenomena occurring within an El Niño year.

8.3.2. Results

The results (Figure 14—blue line) show that there were no drops at the start of the event (the cooler temperatures in May merely represent the continuation of a trend that had started earlier). The time series also showed several large, unexplained drops and peaks over this period.

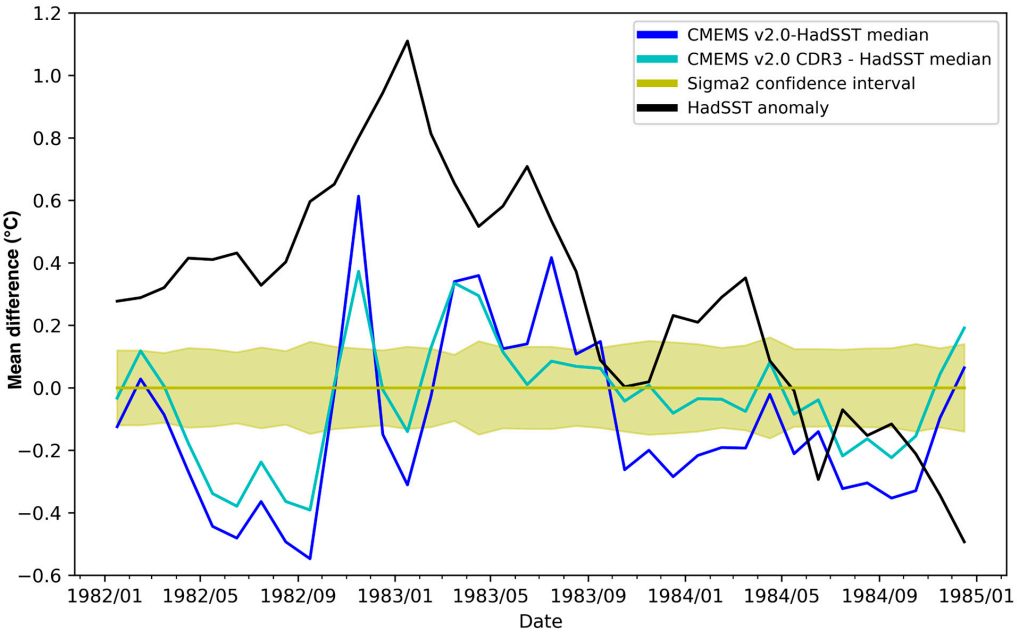


Figure 14. Spatially averaged L4 analysis minus median HadSST4 in situ difference for latitudes 0–30°N and longitudes 60–150°W.

The only input dataset available during the eruption period was the ESA SST CCI AVHRR-7 dataset (Table 2). Merchant et al. [15] noted that the ESA SST CCI v2.1 dataset exhibited unexplained biases in May 1982, October to December 1982, early August 1983, and late September 1983. These correspond roughly to the peaks/troughs observed in the bias analysis (Figure 14—Blue line). Where only one instrument was available, OSTIA was sensitive to satellite calibration issues. Therefore, it appeared that the behavior of the SST values over time did not appear to correspond to volcanic activity, but this was more likely the result of calibration errors within the AVHRR-7 instrument.

Again, it was decided to run the CMEMS v2 analysis using the ESA SST CCI v3 AVHRR dataset. As with the Pinatubo investigation, the results showed similar variability and a slight relative warmth (Figure 14—cyan line).

In conclusion, there was no clear impact of the El Chichón eruption on the CMEMS v2.0 dataset. This may be related to the smaller aerosol cloud than that of Pinatubo, aerosol corrections applied to the AVHRR dataset used, and instability in the AVHRR-7 data at this time, which together may have helped to mitigate or obscure any volcanic signal.

9. Conclusions

In this paper, the CMEMS v2.0 dataset is presented, which is a significant upgrade to the previous version (v1.2). In addition to its greater temporal span, there have been many upgrades to the OSTIA analysis used to generate it. Chief among them was the use of the NEMOVAR data assimilation scheme. Although NEMOVAR has already been proven to be superior to the optimal interpolation method used in CMEMS v1.2 [20], the results presented herein further verify its effectiveness.

The analysis versus in situ verification carried out in Section 5.2 showed that the SSTs produced by CMEMS v2.0 were in better agreement with matched in situ data than CMEMS v1.2. Section 6 showed that this method also quantified its uncertainties more accurately than CMEMS v1.2, allowing users to have greater confidence in the reliability of the CMEMS v2.0 dataset. The investigation carried out in Section 7 showed that CMEMS v2.0 can also resolve features more clearly.

In Section 5.2, it was demonstrated that a small cool bias observed from 2012 onwards can be explained by the use of the AVHRR/SLSTR instruments chosen as the reference sensors at this time. Despite this, the use of these sensors was still justified due to their accuracy and reliability, and the analysis was still well within the target accuracy of 0.1 K [2].

An investigation into the robustness of the new dataset to two volcanic eruptions (Mt. Pinatubo and El Chichón) yielded mixed results (Section 8). A comparison with the HadSST4 dataset suggested a cool bias in the CMEMS v2.0 dataset versus the v1.2 dataset following the Pinatubo eruption (a comparison with sparser buoy-only observations yielded less clear results). No clear impact of the El Chichón eruption on CMEMS v2.0 was found.

The study of marine heatwaves (MHW) and marine cold spells (MCS) is becoming increasingly important to climate science, but most research into the topic has used the NOAA OISST surface SST dataset (Table 1) [74,75]. One of the rare exceptions is a report in the 7th edition of the Copernicus Ocean State Report, which used OSTIA CMEMS v2.0 to investigate the trends in MHW and MCS [76]. However, more research should be conducted regarding whether a foundation SST product would be more appropriate for studying MHW/MCS events than a surface SST, and CMEMS v2.0 would be a good candidate dataset to use for this investigation.

In conclusion, although this study was unable to prove CMEMS v.20's robustness to volcanic eruptions, the performance of CMEMS v2.0 was evaluated and has been shown to be overall generally superior to the previous version. Coupled with the fact that it is the only long-term daily foundation SST data record, it should, therefore, be useful to users from a wide range of disciplines.

10. List of Acronyms

Acronym	Expansion
AMSR	Advanced Microwave Scanning Radiometer
ATSR	Along-Track Scanning Radiometer
AVHRR	Advanced Very-High-Resolution Radiometer
C3S	Copernicus Climate Change Services
CDR	Climate Data Record
CMEMS	Copernicus Marine Environment Monitoring Service
DMSP	Defence Meteorological Satellite Program
DOISST	Daily Optimum Interpolation Sea Surface Temperature
ECMWF	European Centre for Medium-Range Weather Forecasting
ERS	Earth Remote Sensing Satellite
ESA SST CCI	European Space Agency SST Copernicus Climate Initiative
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GAC	Global Area Coverage
GCOM-W	Global Change Observation Mission—Water
GCOS	Global Climate Observing System
GHRSST	Group for High-Resolution SST
GMI	GPM Microwave Imager
GOES	Geostationary Operational Environmental Satellite
GPM	Global Precipitation Measurement
GTMBA	Global Tropical Moored Buoy Array
GTS	Global Telecommunication System
HadIOD	Met Office Hadley Centre Integrated Ocean Database
HadSST4	Hadley Centre's Sea Surface Temperature Dataset
HIRS	High-Resolution Infrared Radiation Sounder
ICOADS	International Comprehensive Ocean–Atmosphere Dataset
IR	Infra-Red
JAXA	Japan Aerospace Exploration Agency
L2P	Level 2 Preprocessed
L3C	Level-3 Collated
L3U	Level-3 Uncollated
L4	Level 4
MAD	Median Absolute Deviation
MCS	Marine Cold Spell
MetDB	Met Office Observations Database
MGDSST	The Merged Satellite and In Situ Data Global Daily Sea Surface Temperature
MHW	Marine Heatwave
NEMO	Nucleus for European Modeling of the Ocean
NOAA	National Oceanic and Atmospheric Administration
NRT	Near Real Time
NWP	Numerical Weather Prediction
OE	Optimal Estimation
OI	Optimal Interpolation
OSI-SAF	Ocean and Sea Ice Satellite Application Facility hosted by EUMETSAT
OSTIA	Ocean Sea Temperature and Ice Analysis
PMW	Passive MicroWave
PODAAC	Physical Oceanography Distributed Active Archive Centre
QC	Quality Control
REMSS	Remote Sensing Systems
RFI	Radio Frequency Interference
ROI	Region of Interests
RSD	Robust Standard Deviation
RSE	Robust Standard Error

RTTOV	Full expansion: RTTOV = Radiative Transfer for TOVS. TOVS = TIROS Operational Vertical Sounder. TIROS = Television Infrared Observation Satellite.
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SIRDS	SST CCI Independent Reference Dataset
SLSTR	Sea and Land Surface Temperature Radiometer
SMMR	Scanning Multichannel Microwave Radiometer
SNPP	Suomi National Polar-Orbiting Partnership
SO2	Sulfur Dioxide
SQG	Surface Quasi-Geostrophic
SSH	Sea Surface Height
SSM/I	Special Sensor Microwave/Imager
SSMIS	Special Sensor Microwave Imager/Sounder
SST	Sea Surface Temperature
UKEOCIS	UK Earth Observation Climate Information Service
UKMCAS	UK Marine and Climate Advisory Service
VIIRS	Visible Infrared Imaging Radiometer Suite
ZGM	Zero Gyro Mode

11. Dataset Locations

This information is accurate as of this paper's publication. The authors accept no responsibility if the links expire or become unsafe.

CMEMS v2.0:

Data held on Physical Oceanography Distributed Active Archive Centre (PODAAC).

PODAAC product name: OSTIA-UKMO-L4-GLOB-REP-v2.0
DOI: https://doi.org/10.5067/GHOST-4RM02

Citation: [77]

MGDSST:

Data source: NEAR-GOOS RRTDB
Webpage URL: https://www.data.jma.go.jp/gmd/goos/data/database.html (accessed on 28 August 2024).
Download location: https://www.data.jma.go.jp/gmd/goos/data/pub/JMA-product/mgd_sst_glb_D/Readme_mgd_sst_glb_D (accessed on 28 August 2024).

DOISST v2.1:

Data held on Physical Oceanography Distributed Active Archive Centre (PODAAC).

PODAAC product name: AVHRR_OI-NCEI-L4-GLOB-v2.1
DOI: https://doi.org/10.5067/GHAAO-4BC21
Citation: [78]

ESA SST CCI v2.1:

Webpage URL: https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_REP_OBSERVATIONS_010_024/description (accessed on 28 August 2024).
Dataset DOI: https://doi.org/10.48670/moi-00169
Citations: [15,79]

ESA SST CCI v3:

Webpage URL: https://catalogue.ceda.ac.uk/uuid/4a9654136a7148e39b7feb56f8bb02d2 (accessed on 28 August 2024).
Dataset DOI: https://dx.doi.org/10.5285/4a9654136a7148e39b7feb56f8bb02d2
Citation: [17,80]

Author Contributions: Conceptualization, M.W. and S.G.; methodology, M.W.; software, M.W. and O.E.; validation, M.W., O.E., and C.A.; formal analysis, M.W.; writing—original draft preparation, M.W.; writing—review and editing, M.W. and S.G.; visualization, M.W. and O.E. All authors have read and agreed to the published version of the manuscript.

Funding: This work benefited from funding from the UK government/DSIT Earth Observation Investment Package (https://www.gov.uk/government/publications/earth-observation-investment/projects-in-receipt-of-funding, webpage accessed on 28 August 2024) and the Copernicus Marine Environment Monitoring Service (CMEMS; 78-CMEMS-TAC-SST).

Data Availability Statement: See Section 11 for relevant dataset locations. HadIOD.1.2.0.0 NetCDF feedback format data are available at the Met Office and are © British Crown Copyright, Met Office, 2022, provided under a Non-Commercial Government Licence http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/ (accessed on 28 August 2024). HadIOD SIRDS data are available from the Met Office and are © British Crown Copyright, Met Office, 2022, provided under an Open Government License, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ (accessed on 28 August 2024). HadSST.4.0.1.0 data were obtained from http://www.metoffice.gov.uk/hadobs/hadsst4/data (accessed on 28 August 2024) in September 2022. © British Crown Copyright, Met Office, 2022, provided under an Open Government License, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ (accessed on 28 August 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. GCOS-200. The Global Observing System for Climate: Implementation Needs. 2016. Available online: https://library.wmo.int/idurl/4/55469 (accessed on 28 August 2024).
2. GHRSST Science Team. *The Recommended GHRSST Data Specification (GDS)*; GHRSST: Leicester, UK, 2010; p. 123. [CrossRef]
3. Yang, C.; Leonelli, F.; Marullo, S.; Artale, V.; Beggs, H.; Nardelli, B.; Chin, T.; de Toma, V.; Good, S.; Huang, B.; et al. Sea surface temperature intercomparison in the framework of the Copernicus climate change service (C3S). *Am. Meterological Soc.* **2021**, *34*, 5257–5283. [CrossRef]
4. Donlon, C.; Martin, M.; Stark, J.; Roberts-Jones, J.; Fiedler, E.; Wimmer, W. The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.* **2012**, *116*, 140–158. [CrossRef]
5. Donlon, C.; Minnett, P.J.; Gentemann, C.; Nightingale, T.J.; Barton, I.J.; Ward, B.; Murray, M.J. Toward improved validation of satellite sea surface skin temperature measurements for climate research. *J. Clim.* **2002**, *15*, 353–369. [CrossRef]
6. Fiedler, E.; McLaren, A.; Banzon, V.; Brasnett, B.; Ishizaki, S.; Kennedy, J.; Rayner, N.; Roberts-Jones, J.; Corlett, G.; Merchant, C.; et al. Intercomparison of long-term sea surface temperature analyses using the GHRSST Multi-Product Ensemble (GMPE) system. *Remote Sens. Environ.* **2019**, *222*, 18–33. [CrossRef]
7. Sakurai, T.; Yukio, K.; Kuragano, T. Merged satellite and in-situ data Global Daily SST. *Int. Geosci. Remote Sens. Symp.* **2005**, *4*, 2606–2608. [CrossRef]
8. Yukio, K.; Japan Aerospace Exploration Agency (JAXA), Tokyo, Japan. Personal communication, January 2022.
9. Høyer, J.; Alerskans, E.; Nielsen-Englyst, P.; Worsfold, M.; Good, S.; Pearson, K.; Embury, O.; Merchant, C.; Donlon, C. Passive Microwave SST Production and Impact Assessment. 2019. Available online: https://climate.esa.int/media/documents/SST_CCI-WP90-Final_Report_Issue-1_signed.pdf (accessed on 28 August 2024).
10. Reynolds, R.; Smith, T. Improved global sea surface temperature analyses using optimum interpolation. *J. Clim.* **1994**, *7*, 929–948. [CrossRef]
11. Reynolds, R.; Rayner, N.; Smith, T.; Stokes, D.; Wang, W. An Improved In Situ and Satellite SST Analysis for Climate. *J. Clim.* **2002**, *15*, 1609–1625. [CrossRef]
12. Reynolds, R.; Smith, T.; Liu, C.; Chelton, D.B.; Casey, K.; Schlax, M.G. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *J. Clim.* **2007**, *20*, 5473–5496. [CrossRef]
13. Banzon, V.; Smith, T.; Chin, T.; Liu, C.; Hankins, W.; Ave, P. A long-term record of blended satellite and in situ sea-surface temperature for climate monitoring, modeling and environmental studies. *Earth Syst. Sci. Data* **2016**, *8*, 165–176. [CrossRef]
14. Huang, B.; Liu, C.; Banzon, V.; Freeman, E.; Graham, G.; Hankins, B.; Smith, T.; Zhang, H.M. Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. *J. Clim.* **2021**, *34*, 2923–2939. [CrossRef]
15. Merchant, C.J.; Embury, O.; Bulgin, C.E.; Block, T.; Corlett, G.K.; Fiedler, E.; Good, S.; Mittaz, J.; Rayner, N.; Berry, D.; et al. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data* **2019**, *6*, 223. [CrossRef] [PubMed]
16. Good, S.; Fiedler, E.; Mao, C.; Martin, M.; Maycock, A.; Reid, R.; Roberts-Jones, J.; Searle, T.; Waters, J.; While, J.; et al. The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sens.* **2020**, *12*, 720. [CrossRef]
17. Embury, O.; Merchant, C.; Good, S.; Rayner, N.; Høyer, J.; Atkinson, C.; Block, T.; Alerskans, E.; Pearson, K.; Worsfold, M.; et al. Satellite-based time-series of sea-surface temperature since 1980 for climate applications. *Sci. Data* **2024**, *11*, 326. [CrossRef] [PubMed]
18. Roberts-Jones, J.; Fiedler, E.; Martin, M. Daily, global, high-resolution SST and sea ice reanalysis for 1985–2007 using the OSTIA system. *J. Clim.* **2012**, *25*, 6215–6232. [CrossRef]

19. Mogensen, K.S.; Balmaseda, M.A.; Weaver, A.; Martin, M.; Vidard, A. NEMOVAR: A variational data assimilation system for the NEMO ocean model. *ECMWF Newsl.* **2009**, *120*, 17–21. [CrossRef]
20. Fiedler, E.; Mao, C.; Good, S.; Waters, J.; Martin, M. Improvements to feature resolution in the OSTIA sea surface temperature analysis using the NEMOVAR assimilation scheme. *Q. J. R. Meteorol. Soc.* **2019**, *145*, 3609–3625. [CrossRef]
21. Mogensen, K.; Alonso Balmaseda, M.; Weaver, A. *The NEMOVAR Ocean Data Assimilation System as Implemented in the ECMWF Ocean Analysis for System 4*; ECMWF: Reading, UK, 2012. [CrossRef]
22. Atkinson, C.; Rayner, N.; Kennedy, J.; Good, S. An integrated database of ocean temperature and salinity observations. *J. Geophys. Res. Ocean.* **2014**, *119*, 7139–7163. [CrossRef]
23. Atkinson, C. HadIOD.1.2.0.0 User Guide. 2020. Available online: https://www-hc/~catkinso/hadiod/webpages/hadiod/HadIOD.1.2.0.0_Product_User_Guide_[1.0].pdf (accessed on 28 August 2024).
24. Waters, J.; Lea, D.J.; Martin, M.J.; Mirouze, I.; Weaver, A.; While, J. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Q. J. R. Meteorol. Soc.* **2015**, *141*, 333–349. [CrossRef]
25. Mirouze, I.; Blockley, E.W.; Lea, D.J.; Martin, M.; Bell, M.J. A multiple length scale correlation operator for ocean data assimilation. *Tellus Ser. A Dyn. Meteorol. Oceanogr.* **2016**, *68*, 29744. [CrossRef]
26. Roberts-Jones, J.; Bovis, K.; Martin, M.; McLaren, A. Estimating background error covariance parameters and assessing their impact in the OSTIA system. *Remote Sens. Environ.* **2016**, *176*, 117–138. [CrossRef]
27. Embury, O.; Good, S. C3S SST Product User Guide and Specification. 2020. Available online: https://datastore.copernicus-climate.eu/documents/satellite-sea-surface-temperature/v2.0/D3.SST.1-v2.2_PUGS_of_v2SST_products_v6.0_APPROVED_Ver1.pdf (accessed on 28 August 2024).
28. Embury, O. *ESA CCI Phase 3 Sea Surface Temperature (SST) Algorithm Theoretical Basis Document D2.1 v3*; University of Reading: Reading, UK, 2023.
29. Embury, O. *ESA CCI Phase 3 Sea Surface Temperature (SST) Product Validation and Inter-Comparison Report D4.1 v2*; University of Reading: Reading, UK, 2023.
30. Geostationary Sea Surface Temperature Product User Manual. 2018. Available online: https://osi-saf.eumetsat.int/lml/doc/osisaf_cdop3_ss1_pum_geo_sst.pdf (accessed on 28 August 2024).
31. Draper, D.; Newell, D.A.; Wentz, F.; Krimchansky, S.; Skofronick-Jackson, G. The Global Precipitation Measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3452–3462. [CrossRef]
32. Meissner, T.; Wentz, F.; Draper, D. GMI Calibration Algorithm and Analysis Theoretical Basis Document. 2012. Available online: https://images.remss.com/papers/rsstech/2012_041912_Meissner_GMI_ATBD_vG.pdf (accessed on 28 August 2024).
33. Wentz, F.; Meissner, T.; Scott, J.; Hilburn, K. Remote Sensing Systems GPM GMI Daily Environmental Suite on 0.25 Deg Grid, Version 8.2a. 2015. Available online: www.remss.com/missions/gmi (accessed on 28 August 2024).
34. Wentz, F.; Gentemann, C.; Hilburn, K. Three Years of Ocean Products from AMSR-E: Evaluation and Applications. 2005. Available online: https://images.remss.com/papers/rssconf/wentz_IGARSS_2005_Seoul_AMSRE.pdf (accessed on 28 August 2024).
35. Wentz, F.; Meissner, T.; Gentemann, C.; Hilburn, K.; Scott, J. *Remote Sensing Systems AQUA AMSR-E Daily Environmental Suite on 0.25 Deg Grid, Version V7*; Remote Sensing Systems: Santa Rosa, CA, USA, 2014. Available online: www.remss.com/missions/amsr (accessed on 28 August 2024).
36. Brewer, M.; Remote Sensing Systems (REMSS): Santa Rosa, CA, USA. Personal communication, February 2023.
37. Merchant, C.; Embury, O.; Rayner, N.; Berry, D.; Corlett, G.; Lean, K.; Veal, K.L.; Kent, E.C.; Llewellyn-Jones, D.T.; Remedios, J.J.; et al. A 20 year independent record of sea surface temperature for climate from Along-Track Scanning Radiometers. *J. Geophys. Res. Ocean.* **2012**, *117*, 1–18. [CrossRef]
38. Mutlow, C. ATSR-1/2 User Guide. 1999; 29p. Available online: http://www.atsr.rl.ac.uk/documentation/docs/userguide/atsr_user_guide_rev_3.pdf (accessed on 28 August 2024).
39. Miranda, N.; Rosich, B.; Santella, C.; Grion, M. *Review of the impact of ERS-2 piloting modes on the SAR Doppler stability*; European Space Agency: Paris, France, 2005; pp. 39–47.
40. Berry, D.; Corlett, G.; Embury, O.; Merchant, C. Stability assessment of the (A)ATSR sea surface temperature climate dataset from the European Space Agency Climate Change Initiative. *Remote Sens.* **2018**, *10*, 126. [CrossRef]
41. Robel, J.; Graumann, A.; Kidwell, K.; Aleman, R.; Ruff, I.; Muckle, B.; Kleespies, T. *NOAA KLM User's Guide with NOAA-N, N Prime, and Metop Supplements*; NOAA NESDIS: Asheville, NC, USA, 2014; p. 2530.
42. Gentemann, C.; Hilburn, K. In situ validation of sea surface temperatures from the GCOM-W1 AMSR2 RSS calibrated brightness temperatures. *J. Geophys. Res. Ocean.* **2015**, *120*, 3567–3585. [CrossRef]
43. Kim, H.-Y.; Park, K.-A.; Chung, S.-R.; Baek, S.-K.; Lee, B.-I.; Shin, I.-C.; Chung, C.-Y.; Kim, J.-G.; Jung, W.-C. Validation of Sea Surface Temperature (SST) from Satellite Passive Microwave Sensor (GPM/GMI) and Causes of SST Errors in the Northwest Pacific. *Korean J. Remote Sens.* **2018**, *34*, 1–15.
44. Skofronick-Jackson, G.; Petersen, W.A.; Berg, W.; Kidd, C.; Stocker, E.F.; Kirschbaum, D.B.; Kakar, R.; Braun, S.A.; Huffman, G.J.; Iguchi, T.; et al. The Global Precipitation Measurement (GPM) Mission for Science and Society. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1679–1695. [CrossRef]
45. *Global Sea Ice Concentration Interim Climate Data Record 2016 Onwards*; EUMETSAT-OSI-SAF: De Bilt, The Netherlands, 2017.

46. Lavergne, T.; Macdonald Sørensen, A.; Kern, S.; Tonboe, R.; Notz, D.; Aaboe, S.; Bell, L.; Dybkjær, G.; Eastwood, S.; Gabarro, C.; et al. Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *Cryosphere* **2019**, *13*, 49–78. [CrossRef]
47. *Global Sea Ice Concentration Climate Data Record 1979–2015*; EUMETSAT-OSI-SAF: De Bilt, The Netherlands, 2017. [CrossRef]
48. Atkinson, C. HadIOD.1.2.0.0 Feedback File. 2020. Available online: https://www.metoffice.gov.uk/hadobs/hadiod/feedback_data.html (accessed on 28 August 2024).
49. Atkinson, C. *HadIOD SIRDS Data File*; Met Office: Exeter, UK, 2020.
50. Woodruff, S.D.; Worley, S.J.; Lubker, S.J.; Ji, Z.; Eric Freeman, J.; Berry, D.; Brohan, P.; Kent, E.C.; Reynolds, R.; Smith, S.R.; et al. ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.* **2011**, *31*, 951–967. [CrossRef]
51. Embury, O. *surftemp/c3s-pqar: OSTIA Verification*; CERN: Geneve, Switzerland, 2022. [CrossRef]
52. Fiedler, E.; Mao, C.; McLaren, A. *SST: Results and Recommendations. Euro-Argo Improvements for the GMES Marine Service (E-AIMS)*; Report number: D4.3.3; European Commission: Luxembourg, 2015.
53. Bulgin, C.; Embury, O.; Corlett, G.; Merchant, C. Independent uncertainty estimates for coefficient based sea surface temperature retrieval from the Along-Track Scanning Radiometer instruments. *Remote Sens. Environ.* **2016**, *178*, 213–222. [CrossRef]
54. Huber, P.J. *Robust Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 1981.
55. Embury, O.; Merchant, C.; Corlett, G. A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: Initial validation, accounting for skin and diurnal variability effects. *Remote Sens. Environ.* **2012**, *116*, 62–78. [CrossRef]
56. Oka, E.; Ando, K. Stability of temperature and conductivity sensors of Argo profiling floats. *J. Oceanogr.* **2004**, *60*, 253–258. [CrossRef]
57. Kennedy, J. A review of uncertainty in in situ measurements and data sets of sea surface temperature. *Rev. Geophys.* **2014**, *52*, 1–32. [CrossRef]
58. O'Carroll, A.; Eyre, J.R.; Saunders, R. Three-way error analysis between AATSR, AMSR,E, and in situ sea surface temperature observations. *J. Atmos. Ocean. Technol.* **2008**, *25*, 1197–1207. [CrossRef]
59. Corlett, G.; Merchant, C.; Minnett, P.; Donlon, C. Assessment of long-term satellite derived sea surface temperature records. In *Experimental Methods in the Physical Sciences*; Academic Press: Cambridge, MA, USA, 2014; Volume 47. [CrossRef]
60. Poli, P.; Lucas, M.; O'Carroll, A.; le Menn, M.; David, A.; Corlett, G.; Blouch, P.; Meldrum, D.; Merchant, C.; Belbeoch, M.; et al. The Copernicus Surface Velocity Platform drifter with Barometer and Reference Sensor for Temperature (SVP-BRST): Genesis, design, and initial results. *Ocean. Sci.* **2019**, *15*, 199–214. [CrossRef]
61. Le Traon, P.; Klein, P.; Hua, B.L.; Dibarboure, G. Do altimeter wavenumber spectra agree with the interior or surface quasi-geostrophic theory? *J. Phys. Oceanogr.* **2008**, *38*, 1137–1142. [CrossRef]
62. Fu, L.L. On the Wave Number Spectrum of Oceanic Mesoscale Variability Observed By the Seasat Altimeter. *J. Geophys. Res.* **1983**, *88*, 4331–4341. [CrossRef]
63. Reynolds, R.W. Impact of Mount Pinatubo Aerosols on Satellite-derived Sea Surface Temperatures. *J. Clim.* **1993**, *6*, 768–774. [CrossRef]
64. Stowe, L.L.; Carey, R.M.; Pellegrino, P.P.; Nesdis, N. Monitoring the Mt Pinatubo aerosol layer with NOAA/11 AVHRR data. *Geophys. Res. Lett.* **1992**, *19*, 159–162. [CrossRef]
65. 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.* **1994**, *99*, 25673–25679. [CrossRef]
66. Blackmore, T.; O'Carroll, A.; Fennig, K.; Saunders, R. Correction of AVHRR Pathfinder SST data for volcanic aerosol effects using ATSR SSTs and TOMS aerosol optical depth. *Remote Sens. Environ.* **2012**, *116*, 107–117. [CrossRef]
67. Murray, M.J.; Allen, M.R.; Mutlow, C.; Zavody, A.M.; Jones, M.S.; Forrester, T.N. Actual and potential information in dual-view radiometric observations of sea surface temperature from ATSR. *J. Geophys. Res.* **1998**, *103*, 8153–8165. [CrossRef]
68. Kennedy, J.J.; Rayner, N.A.; Atkinson, C.P.; Killick, R.E. 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.* **2019**, *124*, 7719–7763. [CrossRef]
69. Huang, B.; Thorne, P.W.; Banzon, V.F.; Boyer, T.; Chepurin, G.; Lawrimore, J.H.; Menne, M.J.; Smith, T.M.; Vose, R.S.; Zhang, H.M. Extended reconstructed Sea surface temperature, Version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Clim.* **2017**, *30*, 8179–8205. [CrossRef]
70. Hirahara, S.; Ishii, M.; Fukuda, Y. Centennial-scale sea surface temperature analysis and its uncertainty. *J. Clim.* **2014**, *27*, 57–75. [CrossRef]
71. Rayner, N.; Parker, D.E.; Horton, E.B.; Folland, C.K.; Alexander, L.V.; Rowell, D.P.; Kent, E.C.; Kaplan, A. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res. Atmos.* **2003**, *108*, 4407. [CrossRef]
72. Bluth, G.J.S.; Doiron, S.D.; Schnetzler, C.C.; Krueger, A.J.; Walter, L.S. Global tracking of the SO2 clouds from the June, 1991 Mount Pinatubo eruptions. *Geophys. Res. Lett.* **1992**, *19*, 151–154. [CrossRef]
73. Robock, A.; Taylor, K.E.; Stenchikov, G.L.; Liu, Y. GCM evaluation of a mechanism for El Niño triggering by the El Chichón ash cloud. *Geophys. Res. Lett.* **1995**, *22*, 2369–2372. [CrossRef]

74. Schlegel, R.W.; Darmaraki, S.; Benthuysen, J.A.; Filbee-Dexter, K.; Oliver, E.C.J. Marine cold-spells. *Prog. Oceanogr.* **2021**, *198*, 102684. [CrossRef]
75. Wang, Y.; Kajtar, J.B.; Alexander, L.V.; Pilo, G.S.; Holbrook, N.J. Understanding the Changing Nature of Marine Cold-Spells. *Geophys. Res. Lett.* **2022**, *49*, e2021GL097002. [CrossRef]
76. Peal, R.; Worsfold, M.; Good, S. Comparing global trends in marine cold spells and marine heatwaves using reprocessed satellite data. *State Planet* **2023**, *1-osr7*, 3. [CrossRef]
77. UK Met Office. GHRSST Level 4 OSTIA Global Reprocessed Foundation Sea Surface Temperature Analysis (GDS2). Version 2.0. PO.DAAC, CA, USA. 2023. Available online: https://podaac.jpl.nasa.gov/dataset/OSTIA-UKMO-L4-GLOB-REP-v2.0 (accessed on 28 August 2024).
78. NOAA National Centers for Environmental Information. Daily L4 Optimally Inter-polated SST (OISST) In situ and AVHRR Analysis. Ver. 2.1. PO.DAAC, CA, USA. 2023. Available online: https://podaac.jpl.nasa.gov/dataset/AVHRR_OI-NCEI-L4-GLOB-v2.1 (accessed on 28 August 2024).
79. UK Met Office. ESA SST CCI and C3S reprocessed sea surface temperature analyses. CMEMS, 2019. Available online: https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_REP_OBSERVATIONS_010_024/description (accessed on 28 August 2024).
80. Good, S.A.; Embury, O. ESA Sea Surface Temperature Climate Change Initiative (SST_cci): Level 4 Analysis product, version 3.0. NERC EDS Centre for Environmental Data Analysis, 9 April 2024. 2024. Available online: https://catalogue.ceda.ac.uk/uuid/4a9654136a7148e39b7feb56f8bb02d2/ (accessed on 28 August 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.